



**HAL**  
open science

# SELF-IMPROVEMENT OF THE BAKRY-EMERY CRITERION FOR POINCARÉ INEQUALITIES AND WASSERSTEIN CONTRACTION USING VARIABLE CURVATURE BOUNDS

Patrick Cattiaux, Max Fathi, Arnaud Guillin

► **To cite this version:**

Patrick Cattiaux, Max Fathi, Arnaud Guillin. SELF-IMPROVEMENT OF THE BAKRY-EMERY CRITERION FOR POINCARÉ INEQUALITIES AND WASSERSTEIN CONTRACTION USING VARIABLE CURVATURE BOUNDS. *Journal de Mathématiques Pures et Appliquées*, 2022, 166, pp.1-29. 10.1016/j.matpur.2022.07.003 . hal-02486264

**HAL Id: hal-02486264**

**<https://uca.hal.science/hal-02486264v1>**

Submitted on 20 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SELF-IMPROVEMENT OF THE BAKRY-EMERY CRITERION FOR POINCARÉ INEQUALITIES AND WASSERSTEIN CONTRACTION USING VARIABLE CURVATURE BOUNDS

PATRICK CATTIAUX <sup>♠</sup> , MAX FATHI <sup>♣♣</sup> , AND ARNAUD GULLIN <sup>◇</sup>

<sup>♠</sup> UNIVERSITÉ DE TOULOUSE

<sup>♣</sup> CNRS

<sup>◇</sup> UNIVERSITÉ CLERMONT-AUVERGNE

ABSTRACT. We study Poincaré inequalities and long-time behavior for diffusion processes on  $\mathbb{R}^n$  under a variable curvature lower bound, in the sense of Bakry-Emery. We derive various estimates on the rate of convergence to equilibrium in  $L^1$  optimal transport distance, as well as bounds on the constant in the Poincaré inequality in several situations of interest, including some where curvature may be negative. In particular, we prove a self-improvement of the Bakry-Emery estimate for Poincaré inequalities when curvature is positive but not constant.

*Key words* : Bakry-Emery curvature condition, Poincaré inequality, Wasserstein contraction.

*MSC 2010* : 26D10, 47D07, 60G10, 60J60.

## 1. INTRODUCTION.

Let  $\mu(dx) = Z^{-1} e^{-V(x)} dx$  be a probability measure defined on  $\mathbb{R}^n$ . We assume that  $V$  is a smooth enough function (of  $C^\infty$  class in this introduction). The most classical consequence of the celebrated Bakry-Emery criterion  $CD(\rho, \infty)$  (see [7, Definition 1.16.1]) is that, as soon as the Hessian of  $V$  is uniformly positive definite, i.e. for some  $\rho > 0$ , all  $u \in \mathbb{R}^n$ , all  $x \in \mathbb{R}^n$

$$\langle u, \text{Hess } V(x) u \rangle \geq \rho |u|^2, \tag{1.1}$$

then  $\mu$  satisfies several functional inequalities, including the following:

**Definition 1.1.** A probability measure  $\mu$  satisfies a Poincaré inequality if for all smooth  $f$ , (here and in the sequel we denote by  $\mu(f)$  the integral of  $f$  w.r.t.  $\mu$ ),

$$\text{Var}_\mu(f) := \mu(f^2) - \mu^2(f) \leq C_P(\mu) \mu(|\nabla f|^2). \tag{1.2}$$

It satisfies a logarithmic Sobolev inequality (LSI) if for all smooth  $f$

$$H(f^2|\mu) := \mu(f^2 \ln(f^2)) - \mu(f^2) \ln(\mu(f^2)) \leq C_{LS}(\mu) \mu(|\nabla f|^2). \tag{1.3}$$

---

*Date*: February 20, 2020.

In the sequel,  $C_P$  and  $C_{LS}$  should be understood as the best constants for which the previous inequalities hold, for a given probability measure. We refer to [2, 7, 47] among many others, for a comprehensive introduction to some of the useful consequences of these inequalities, as well as their most important properties.

When curvature is bounded from below by some constant  $\rho > 0$ , then the celebrated Bakry-Emery theorem states that

$$C_P \leq \rho^{-1}; \quad C_{LSI} \leq 2\rho^{-1}. \quad (1.4)$$

An interesting remark is that in the Gaussian case, the Poincaré and logarithmic Sobolev constants obtained through this (seemingly) crude upper bound are in fact optimal. Under this curvature condition, the key element in the usual proofs of these functional inequalities is that the associated semi-group  $P_t = e^{tL}$ , where  $L = \Delta - \nabla V \cdot \nabla$ , satisfies (provided it is well defined), as soon as (1.1) holds, the pseudo commutation property

$$|\nabla P_t f| \leq e^{-\rho t} P_t(|\nabla f|) \quad (1.5)$$

for all smooth  $f$ . The constant  $\rho$  can be seen as some kind of Ricci curvature lower bound for the semi-group. This terminology comes from the fact that, for a Brownian motion on a smooth manifold, this property is actually equivalent to having a lower bound of the form  $\text{Ric} \geq \rho g$ , where  $\text{Ric}$  is the Ricci curvature tensor and  $g$  the metric tensor. Actually (1.5) is true as soon as (1.1) is satisfied even  $\rho \in \mathbb{R}$  is non-positive, but in this case one cannot immediately deduce the inequalities we are interested in for  $\mu$ . The probabilistic interpretation is provided by the associated stochastic process  $X_t^x$  solution of the stochastic differential (integral) equation

$$X_t^x = x + \sqrt{2} B_t - \int_0^t \nabla V(X_s^x) ds \quad (1.6)$$

where  $B$  is a standard Brownian motion, and with the semi group then satisfying the formula

$$P_t f(x) = \mathbb{E}(f(X_t^x)) := \mathbb{E}_x(f(X_t)). \quad (1.7)$$

The notation  $\mathbb{E}_\nu$  will be used when looking at the stochastic process starting from some random initial data  $X_0$  with distribution  $\nu$ . One can check that  $\mu$  is a stationary (and even reversible) measure for this process.

A natural question is to extend (1.5) to the more general situation of a non-constant Ricci curvature lower bound. More precisely, since  $\text{Hess} V(x)$  is a real symmetric matrix, all its eigenvalues  $(\rho_i(x))_{i=1,\dots,n}$  are real and

$$\langle u, \text{Hess} V(x) u \rangle \geq \rho(x) |u|^2, \text{ with } \rho(x) = \rho_m(x) = \min_{i=1,\dots,n} \rho_i(x). \quad (1.8)$$

Notice that  $\rho_m(x)$  is optimal in the previous inequality simply looking at the eigenvectors of the Hessian matrix.

**Definition 1.2.** We shall say that the curvature is bounded from below if (1.8) is satisfied for some  $\rho(x) \geq \rho_0 \in \mathbb{R}$  and any  $x$ .

Equivalently, for  $\lambda > -\rho_0/2$ ,  $\lambda|x|^2 + V(x)$  is uniformly convex. It follows that, defining  $\psi(x) = |x|^2$ ,  $L\psi(x) \leq K\psi(x) + a^2$  for some constant  $a$ . It is then classical that the process is conservative (non explosive) and ergodic with invariant (reversible) measure  $\mu$ . More precisely, thanks to reversibility (i.e. symmetry in  $L^2(\mu)$ ), the spectral theorem allows us to

show that for  $f \in \mathbb{L}^2(\mu)$ ,  $P_t f \rightarrow \mu(f)$  as  $t \rightarrow +\infty$  in  $\mathbb{L}^2(\mu)$  (a direct probabilistic proof is contained in [15]).

One motivation for introducing a variable curvature bound, even in situations where curvature is actually bounded from below by some positive constant, is that the Bakry-Emery theorem is rigid [21, 25, 41]: if equality holds in either bounds of (1.4), then the measure *must* split off a Gaussian factor, and therefore the optimal  $\rho_m$  in (1.8) is constantly equal to  $\rho$ . Therefore for any measure satisfying a Ricci curvature bound with a *variable* optimal curvature lower bound must satisfy a Poincaré inequality and a logarithmic Sobolev inequality with strictly better bounds. An improvement using the harmonic average of the curvature bound was obtained in [44] for the Poincaré constant. See also [23, 22] for some results in that direction, of a different flavor than those we shall obtain here.

We shall first give a simple proof of the following generalization of (1.5) (see section 2 for precise statements)

**Proposition 1.3.** *If (1.8) is satisfied for some regular enough function  $\rho$ , it holds*

$$|\nabla P_t f|(x) \leq \mathbb{E}_x \left( e^{-\int_0^t \rho(X_s) ds} |\nabla f|(X_t) \right).$$

This result is not new. It is shown for instance in the recent work by Braun, Habermann and Sturm [14] in the much more general framework of metric measure spaces. That work also contains numerous results on equivalences between several notions of “variable Ricci curvature”, and was actually the starting point of the present work. In the diffusion case we are looking at, some of us heard about Proposition 1.3 from D. Bakry in informal discussions. A partial result in this direction is also contained in [18, Section 7.4]. Some related gradient bounds using local curvature bounds were investigated in [4].

The natural question is then to understand what consequences we can get from this gradient bound. The goal of the present work is to derive some applications to the study of the long-time behavior of the underlying stochastic dynamic.

We would like to mention another related approach, the so called “intertwining” method. When  $\mu$  is gaussian with covariance matrix  $\rho Id$  (i.e.  $X$  is an Ornstein-Uhlenbeck process), the inequality (1.5) becomes an equality

$$\nabla P_t f(x) = e^{-\rho t} P_t(\nabla f)(x).$$

An intertwining semi-group is a distortion of the gradient such that

$$\nabla P_t f(x) = P_t^A(A \nabla f)(x),$$

for some perturbed semi-group  $P_t^A$ . In some situations, if the perturbation is nice enough one may use it to recover gradient bounds similar to (1.5). The intertwining method has been well known for a long time in the context of stochastic processes. Its application to functional inequalities (Poincaré or Brascamp-Lieb) is more recent. For background about implementation of the intertwining method for diffusion processes, we refer to [3, 12, 13].

Another approach that can sometimes be applied in situations where curvature may be negative is that of F-Y. Wang [46], later extended by E. Milman [40], who showed that if curvature is bounded from below by some uniform, but possibly negative constant, a strong enough Gaussian concentration inequality allows to recover a logarithmic Sobolev inequality, and even an isoperimetric inequality. See also [26] for an alternative proof, and [9] for some

applications in statistical physics. The results obtained via that method and those we shall present here do not seem to be directly comparable.

In order to use Proposition 1.3, consider some 1-Lipschitz function  $f$ . If  $X_0$  is distributed according to  $\nu$ , we denote by  $P_t^* \nu$  the distribution of  $X_t$ . Since the semi-group is symmetric we should omit the  $*$ , but it helps to understand the nature of the various objects. We thus have for probability measures  $\nu$  and  $\beta$ ,

$$|P_t^* \nu(f) - P_t^* \beta(f)| = |\nu(P_t f) - \beta(P_t f)|$$

and

$$\| |\nabla P_t f| \|_\infty \leq \sup_x \mathbb{E}_x \left[ e^{-\int_0^t \rho(X_s) ds} \right].$$

It follows, using its variational expression, that the  $W_1$  Wasserstein distance satisfies

$$W_1(P_t^* \nu, P_t^* \beta) \leq \sup_x \mathbb{E}_x \left[ e^{-\int_0^t \rho(X_s) ds} \right] W_1(\nu, \beta). \quad (1.9)$$

In order to get some decay to 0 for  $W_1(P_t^* \nu, \mu)$ , it thus remains to estimate the sup-norm of  $\mathbb{E}_x \left( e^{-\int_0^t \rho(X_s) ds} \right)$ . This connection between convergence in Wasserstein distance and gradient estimates is a particular instance of the Kuwada duality theorem [33].

Assuming that the process is ergodic (for instance when the curvature is bounded from below), we know that

$$\frac{1}{t} \int_0^t \rho(X_s) ds \rightarrow \mu(\rho) \quad (1.10)$$

as  $t \rightarrow \infty$ ,  $\mathbb{P}_x$  almost surely for all  $x$ .

Hence we should expect that  $e^{-\int_0^t \rho(X_s) ds}$  may behave like  $e^{-\mu(\rho)t}$  for large  $t$ , so that one can expect that, replacing  $\rho = \min_x \rho(x)$  in the Bakry-Emery criterion by  $\mu(\rho)$ , will allow us to derive interesting results as soon as  $\mu(\rho) > 0$ . In particular, we can hope to handle some situations where  $\rho$  is negative in some region of space, possibly even at infinity.

It turns out that  $\mu(\rho_m)$  is often positive. For instance if  $n = 1$ , since  $\rho_m = V''$ , integrating by parts yields,

$$\mu(\rho_m) = \int V'' e^{-V} dx = \int (V')^2 e^{-V} dx > 0,$$

provided the integrals exist and  $V' e^{-V}$  goes to 0 at infinity.

For a general  $n$ , and any  $u \in \mathbb{R}^n$ , it holds

$$\int \langle u, \text{Hess } V u \rangle e^{-V} dx = \int \langle u, \nabla V \rangle^2 e^{-V} dx \geq 0,$$

provided everything makes sense and  $\langle u, \nabla V \rangle e^{-V}$  goes to 0 at infinity. But in order to estimate  $\mu(\rho_m)$  we would have to take the infimum w.r.t.  $u$  under the integral sign, so that (strict) positivity is unclear.

Nevertheless, the optimal lower bound on average curvature is often positive. This means that something else is needed in order to obtain interesting consequences, since the asymptotic behavior of the exponential term in (1.9) cannot be that simple.

Another reason is the following: if we have a uniform contractive estimate of the form

$$\sup_x \mathbb{E}_x \left( e^{-\int_0^t \rho(X_s) ds} \right) \leq e^{-ct},$$

then the semi-group is a contraction in  $W_1$  distance. This property is known to imply that the curvature is bounded from below by a *positive* constant [45].

One weakness in this approach is that (1.9) involves a supremum over  $x$ . Such a uniform control requires strong continuity assumptions for the semi-group (for instance ultra-boundedness). It is thus interesting to try to get a direct expression for  $W_1(P_t^* \nu, P_t^* \beta)$  without using the variational expression of  $W_1$ . This is done in some specific situations in Section 3.

Similarly, one can try to get  $\mathbb{L}^2$  estimates. Recall that

$$\mathrm{Var}_\mu(P_t f) = 2 \int_t^{+\infty} \mu(|\nabla P_s f|^2) ds.$$

Hence if  $f$  is 1-Lipschitz we thus have

$$\mathrm{Var}_\mu(P_t f) \leq 2 \int_t^{+\infty} \mathbb{E}_\mu \left( e^{-\int_0^s 2\rho(X_u) du} \right) ds. \quad (1.11)$$

If we are able to show that

$$\mathbb{E}_\mu \left( e^{-\int_0^s 2\rho(X_u) du} \right) \leq C e^{-cs}$$

for some  $c > 0$  provided  $\mu(\rho) > 0$ , one can expect some decay

$$\mathrm{Var}_\mu(P_t f) \leq \frac{2C}{c} e^{-ct}$$

for all 1-Lipschitz function  $f$ . Recall the following result [20, Lemma 2.12]:

**Lemma 1.4.** *Let  $\mathcal{C}$  be a dense subset of  $\mathbb{L}^2(\mu)$ . Suppose that there exists  $c > 0$ , and, for any  $f \in \mathcal{C}$ , a constant  $c_f$  such that:*

$$\forall t, \quad \mathrm{Var}_\mu(P_t f) \leq c_f e^{-ct}.$$

*Then*

$$\forall f \in \mathbb{L}^2(\mu), \forall t, \quad \mathrm{Var}_\mu(P_t f) \leq e^{-ct} \mathrm{Var}_\mu(f).$$

Using homogeneity and since we are in the reversible situation, we will deduce that for all  $f$ ,  $\mathrm{Var}_\mu(P_t f) \leq e^{-ct} \mathrm{Var}_\mu(f)$ , so that  $\mu$  satisfies a Poincaré inequality with constant less than  $2/c$ .

In order to develop this strategy, what is required is thus to control the rate of convergence in (1.10). Deviation bounds for additive functionals of ergodic diffusion processes have been studied in [17, 27] based on a previous result by Wu ([48]). One can also mention [36] where analytic tools are used. We shall mainly use [27] where a very detailed study is performed. It is also presumably possible to use direct controls as it is done in [37] for the one dimensional situations, though some points in [37] are not totally clear for us.

To end this introduction let us give a flavor of the results we obtained in the case of Poincaré constant (see Th.5.3 for a more complete result)

**Theorem 1.5.** *Assume that  $V$  is  $C^2$  and that  $\rho$  is bounded below by  $\rho_0 > 0$ , then*

$$C_P(\mu) \leq \frac{1}{\rho_0 + \varepsilon'}$$

with

$$\varepsilon' = \left( (\mu(\rho) - \rho_0) + \frac{1}{\rho_0} \text{Osc}^2(\rho) \right) \left( 1 - \sqrt{1 - \frac{(\mu(\rho) - \rho_0)^2}{\left( (\mu(\rho) - \rho_0) + \frac{1}{\rho_0} \text{Osc}^2(\rho) \right)^2}} \right).$$

Remark that we may always choose  $\rho$  to be bounded, eventually losing on  $\mu(\rho)$ . We thus obtain a strict improvement of the Bakry-Emery estimation of the Poincaré constant in the variable curvature case.

The plan of the paper is the following: in Section 2 we prove by differentiation of the flow or by coupling Proposition 1.3. Then we introduce the methodology for the Wasserstein contraction in Section 3 showing that the crucial estimate is the one of  $E_\nu \left( e^{-\int_0^t \rho(X_s) ds} \right)$ . By using transport-information inequalities we obtain controls for such Laplace transform in Section 4, where a particular attention is considered for the effect of the initial measure  $\nu$ , thus obtaining Wasserstein contraction estimates as well as general bounds on the Poincaré constant. Finally, Section 5 considers various examples and applications, e.g. the logconcave case or the uniformly convex case.

## 2. PROOF(S) OF PROPOSITION 1.3.

We shall propose two approaches for proving Proposition 1.3: the first one is based on the theory of stochastic flows in the spirit of [32, 28], the second one on coupling following the ideas in [18, 44]. Since the framework is simpler than that of [14], the proofs are more direct.

### 2.1. Proof via stochastic flows.

First recall the following result which is a consequence of [32, Theorem 3.1 and Theorem 5.4]

**Theorem 2.1.** *Let*

$$X_t^x = x - \int_0^t \nabla V(X_s^x) ds + \sqrt{2} B_t.$$

*Assume that  $V$  is  $C^k$ , and the process is conservative (i.e. for all  $x$  the lifetime is almost surely infinite). Then for all  $t$  the application  $x \mapsto X_t^x$  is a.s.  $C^{k-2}$ , and its derivatives are obtained by formal differentiation. In particular*

$$\partial_i X_t^x := \frac{\partial X_t^x}{\partial x_i}$$

*satisfies*

$$\partial_i X_t^x = e_i - \int_0^t \text{Hess } V(X_s^x) \partial_i X_s^x ds,$$

*where  $e_i$  is the  $i^{\text{th}}$  canonical unit vector, i.e.*

$$\partial_i X_t^x = \left( \exp \left( - \int_0^t \text{Hess } V(X_s^x) ds \right) \right) e_i.$$

If  $f$  is a smooth function, we thus have

$$\frac{\partial}{\partial x_i} (f(X_t^x)) = \left\langle (\nabla f)(X_t^x), \left( \exp \left( - \int_0^t \text{Hess } V(X_s^x) ds \right) \right) e_i \right\rangle$$

so that

$$|\nabla(f(X_t^x))|^2 = \left\langle (\nabla f)(X_t^x), \left( \exp \left( - 2 \int_0^t \text{Hess } V(X_s^x) ds \right) \right) (\nabla f)(X_t^x) \right\rangle.$$

Since  $P_t f(x) = \mathbb{E}(f(X_t^x))$  we may calculate  $\nabla P_t f(x)$  by differentiating under the expectation, if such a differentiation is allowed. But from what precedes

$$\begin{aligned} |\nabla(f(X_t^x))| &\leq |(\nabla f)(X_t^x)| \sup_{u \in \mathbb{S}^{n-1}} \left\langle u, \left( \exp \left( - 2 \int_0^t \text{Hess } V(X_s^x) ds \right) \right) u \right\rangle^{\frac{1}{2}} \quad (2.1) \\ &\leq e^{-\left(\int_0^t \rho_m(X_s^x) ds\right)} |(\nabla f)(X_t^x)|, \end{aligned}$$

where  $\rho_m$  is defined in (1.8).

Indeed, if  $\lambda(t, x)$  denotes the smallest eigenvalue of the real and symmetric matrix  $\int_0^t \text{Hess } V(X_s^x) ds$  the supremum in (2.1) is attained for a corresponding normalized eigenvector  $v(t, x)$  and is equal to  $e^{-\lambda(t, x)}$ . But

$$\begin{aligned} \lambda(t, x) &= \left\langle v(t, x), \left( \int_0^t \text{Hess } V(X_s^x) ds \right) v(t, x) \right\rangle \\ &= \int_0^t \langle v(t, x), \text{Hess } V(X_s^x) v(t, x) \rangle ds \\ &\geq \left( \int_0^t \rho_m(X_s^x) ds \right) |v(t, x)|^2 = \int_0^t \rho_m(X_s^x) ds. \end{aligned}$$

Hence using Lebesgue's theorem for differentiating under the expectation we have obtained the following precise version of Proposition 1.3

**Theorem 2.2.** *Assume that  $V$  is  $C^3$ , that the process is conservative and that  $x \mapsto \mathbb{E} \left( e^{-\int_0^t \rho(X_s^x) ds} \right)$  is locally bounded,  $\rho$  being defined in (1.8). This is the case for instance when the curvature is bounded from below. Then for all smooth  $f$ ,*

$$|\nabla P_t f|(x) \leq \mathbb{E}_x \left( e^{-\int_0^t \rho(X_s) ds} |\nabla f|(X_t) \right).$$

Note that if  $\rho$  is bounded from below by some constant (which may be negative) and  $V$  is smooth enough, this theorem does apply.

## 2.2. Proof via coupling.

We shall now follow the method used by two of us in [18], namely synchronous coupling. The argument below is very close to the one of [44, Theorem 6], that derived a strongly related Wasserstein distance estimate for diffusions on manifolds.

Since (1.6) admits a strong solution, one can build with the *same* Brownian motion a pair of solutions starting from  $(x, y)$ , denoted again by  $(X^x, X^y)$ .



First recall that

$$\begin{aligned} \langle \nabla V(z) - \nabla V(z'), z - z' \rangle &= \int_0^1 \langle z - z', \text{Hess } V(\lambda z + (1 - \lambda)z') (z - z') \rangle d\lambda \\ &\geq \int_0^1 \rho(\lambda z + (1 - \lambda)z') |z - z'|^2 d\lambda. \end{aligned}$$

Assuming again that the curvature is bounded from below, we thus have

$$e^{\int_0^t 2 \int_0^1 \rho(\lambda X_s^x + (1 - \lambda)X_s^y) d\lambda ds} |X_t^x - X_t^y|^2 = |x - y|^2 - 2A(t) \quad (2.2)$$

with

$$\begin{aligned} A(t) &= \int_0^t \left( \langle \nabla V(X_s^x) - \nabla V(X_s^y), X_s^x - X_s^y \rangle - \left( \int_0^1 \rho(\lambda X_s^x + (1 - \lambda)X_s^y) d\lambda \right) |X_s^x - X_s^y|^2 \right) \\ &\quad \times e^{\int_0^s 2 \int_0^1 \rho(\lambda X_u^x + (1 - \lambda)X_u^y) d\lambda du} ds \\ &\leq 0. \end{aligned}$$

Hence

$$|X_t^x - X_t^y| \leq |x - y| e^{-\int_0^t \int_0^1 \rho(\lambda X_s^x + (1 - \lambda)X_s^y) d\lambda ds}. \quad (2.3)$$

Following [18] p.5 we thus have, using the mean value theorem

$$|P_t f(x) - P_t f(y)| \leq \mathbb{E}(|f(X_t^x) - f(X_t^y)|) \leq |x - y| \mathbb{E} \left( |\nabla f(z_t)| e^{-\int_0^t \int_0^1 \rho(\lambda X_s^x + (1 - \lambda)X_s^y) d\lambda ds} \right)$$

for some  $z_t$  sandwiched by  $X_t^x$  and  $X_t^y$ . It remains to divide by  $|x - y|$ , use that  $X_s^y$  goes to  $X_t^x$  almost surely as  $y \rightarrow x$ , that the curvature is bounded from below and Lebesgue's theorem to get a slightly different version of Theorem 2.2

**Theorem 2.3.** *Assume that  $V$  is  $C^2$ , that the curvature is bounded from below and that  $\rho$  is continuous,  $\rho$  being defined in (1.8). Then for all smooth  $f$ ,*

$$|\nabla P_t f|(x) \leq \mathbb{E}_x \left( e^{-\int_0^t \rho(X_s) ds} |\nabla f|(X_t) \right).$$

### 3. A DIRECT CONTROL OF THE $W_1$ DISTANCE.

Unlike what is done in [18], studying the evolution of the Wasserstein distance starting from (2.3) is not immediate. The difference is that in the constant curvature case there is no need to interpolate between  $X_t^x$  and  $X_t^y$ .

#### 3.1. Reinforcing the Ricci bound.

To bypass this issue, we shall reinforce inequality (1.8) by assuming that there exists a function  $\kappa : \mathbb{R}^n \rightarrow \mathbb{R}$  such that for all  $(x, y) \in \mathbb{R}^{2n}$

$$\langle \nabla V(x) - \nabla V(y), x - y \rangle \geq (\kappa(x) + \kappa(y)) |x - y|^2. \quad (3.1)$$

If this holds for some  $\kappa$ , taking limits  $y \rightarrow x$  it is easily seen that  $2\kappa(x) \leq \rho_m(x)$ . We can then follow the previous proof and get

$$|X_t^x - X_t^y| \leq |x - y| e^{-\int_0^t (\kappa(X_s^x) + \kappa(X_s^y)) ds}. \quad (3.2)$$

As a consequence,

$$W_1(P_t^* \delta_x, P_t^* \delta_y) \leq \mathbb{E}_x^{\frac{1}{2}} \left( e^{-\int_0^t 2\kappa(X_s) ds} \right) \mathbb{E}_y^{\frac{1}{2}} \left( e^{-\int_0^t 2\kappa(X_s) ds} \right) |x - y|. \quad (3.3)$$

The advantage of this bound is that we do not need a sup-norm control anymore. Another approach, using a related coarse, non-local version of curvature bounds was used in [44]. Of course, if we want to replace the Dirac masses by general measures, the situation is not as simple. Nevertheless we have for any coupling  $\pi$  of  $\nu$  and  $\mu$ ,

$$\begin{aligned} W_1(P_t^* \nu, \mu) &\leq \mathbb{E} \left( |X_0^\mu - X_0^\nu| e^{-\int_0^t \kappa(X_s^\mu) ds} e^{-\int_0^t \kappa(X_s^\nu) ds} \right) \\ &\leq \mathbb{E}_\pi^{\frac{1}{p}} (|y - x|^p) \mathbb{E}_\mu^{\frac{1}{r}} \left( e^{-\int_0^t r \kappa(X_s) ds} \right) \mathbb{E}_\nu^{\frac{1}{q}} \left( e^{-\int_0^t q \kappa(X_s) ds} \right) \end{aligned} \quad (3.4)$$

provided  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ . By choosing an appropriate coupling  $\pi$ , we then obtain the following:

**Proposition 3.1.** *Assume that  $V$  is  $C^2$  and that (3.1) is satisfied for some  $\kappa$  which is bounded from below by  $\kappa_0 \in \mathbb{R}$ . Then for all  $(p, q, r)$  with  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ ,*

$$W_1(P_t^* \nu, \mu) \leq W_p(\nu, \mu) \mathbb{E}_\mu^{\frac{1}{r}} \left( e^{-\int_0^t r \kappa(X_s) ds} \right) \mathbb{E}_\nu^{\frac{1}{q}} \left( e^{-\int_0^t q \kappa(X_s) ds} \right).$$

Notice that since  $2\kappa \leq \rho_m$ , the assumptions in the previous proposition imply that the curvature is bounded from below, hence that the process is conservative.

*Remark 3.2.* Since this proof is based on a coupling argument, we can also extend it to control  $W_p$  distances with  $p > 1$ .

*Remark 3.3.* The set of  $V$  satisfying (3.1) contains of course all uniformly convex functions (for which  $\kappa$  is a positive constant), but also functions like  $V(x) = g(|x|^2)$  for some convex function  $g : \mathbb{R}^+ \rightarrow \mathbb{R}$  with  $\kappa(x) = g'(|x|^2)$ . Indeed for such functions

$$\begin{aligned} \langle \nabla V(x) - \nabla V(y), x - y \rangle &= 2(g'(|x|^2)|x|^2 + g'(|y|^2)|y|^2 - (g'(|x|^2) + g'(|y|^2)) \langle x, y \rangle) \\ &\geq (g'(|x|^2)|x|^2 + g'(|y|^2)|y|^2) - 2(g'(|x|^2) + g'(|y|^2)) \langle x, y \rangle \\ &= (g'(|x|^2) + g'(|y|^2))|x - y|^2 \end{aligned}$$

since the difference equal to

$$(g'(|x|^2) - g'(|y|^2))(|x|^2 - |y|^2) \geq 0$$

thanks to the convexity of  $g$ . ◇

*Remark 3.4.* Actually, as explained in [18, Remark 6], using [8, Theorem 4.1 and below], one can deduce the gradient commutation property from the decay of the  $W_1$  distance. This is done in [18, Theorem 14] to derive such a commutation under the assumption

$$\langle \nabla V(x) - \nabla V(y), x - y \rangle \geq \kappa(|x - y|) |x - y|^2,$$

replacing the synchronous coupling by the mirror coupling and assuming that  $\kappa(u)$  goes to  $\kappa_\infty > 0$  as  $u$  goes to  $+\infty$  ( $\kappa(0)$  is not defined). This assumption is satisfied when  $V$  is uniformly convex at infinity. ◇

### 3.2. Some special cases.

Assumption (3.1) is nevertheless much stronger than (1.8). It is thus natural to ask whether one can obtain similar results starting from (2.3)

$$|X_t^x - X_t^y| \leq |x - y| e^{-\int_0^t \int_0^1 \rho(\lambda X_s^x + (1-\lambda)X_s^y) d\lambda ds}.$$

The difficulty is to control  $\rho(\lambda x + (1-\lambda)y)$  in terms of  $\rho(x)$  and  $\rho(y)$ . In dimension one however we may use the following well known monotonicity property of the coupling we used:

$$\text{if } x \leq y, \text{ then for all } s \in \mathbb{R}, X_s^x \leq X_s^y.$$

We may then state a first result:

**Proposition 3.5.** *Assume that  $n = 1$ ,  $V$  is  $C^2$  and that (1.8) is satisfied for some  $\rho$  which is bounded from below by  $\rho_0 \geq 0$  (log-concave case). Assume in addition that  $\rho$  is non-increasing on  $(-\infty, a)$  and non-decreasing on  $(a, +\infty)$  for some  $a \in \mathbb{R}$ .*

*Then for all  $(p, q, r)$  with  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ ,*

$$W_1(P_t^* \nu, \mu) \leq W_p(\nu, \mu) A_t$$

where

$$\begin{aligned} A_t &= \mathbb{E}_\mu^{\frac{1}{r}} \left( e^{-\int_0^t r \mathbf{1}_{X_s > a} \rho(X_s) ds} \right) \mathbb{E}_\nu^{\frac{1}{q}} \left( e^{-\int_0^t q \mathbf{1}_{X_s < a} \rho(X_s) ds} \right) + \\ &+ \mathbb{E}_\mu^{\frac{1}{r}} \left( e^{-\int_0^t r \mathbf{1}_{X_s < a} \rho(X_s) ds} \right) \mathbb{E}_\nu^{\frac{1}{q}} \left( e^{-\int_0^t q \mathbf{1}_{X_s > a} \rho(X_s) ds} \right). \end{aligned}$$

*Proof.* If  $x \leq y$ , and since  $\rho(z) \geq 0$ , it holds

$$\rho(\lambda X_s^x + (1-\lambda)X_s^y) \geq \rho(X_s^x) \mathbf{1}_{X_s^x > a} + \rho(X_s^y) \mathbf{1}_{X_s^y < a}.$$

It follows

$$\begin{aligned} W_1(P_t^* \nu, \mu) &\leq \mathbb{E} \left( |X_0^\mu - X_0^\nu| \mathbf{1}_{X_0^\mu \leq X_0^\nu} e^{-\int_0^t (\rho(X_s^\mu) \mathbf{1}_{X_s^\mu > a} + \rho(X_s^\nu) \mathbf{1}_{X_s^\nu < a}) ds} \right) + \\ &+ \mathbb{E} \left( |X_0^\mu - X_0^\nu| \mathbf{1}_{X_0^\mu \geq X_0^\nu} e^{-\int_0^t (\rho(X_s^\mu) \mathbf{1}_{X_s^\mu < a} + \rho(X_s^\nu) \mathbf{1}_{X_s^\nu > a}) ds} \right) \end{aligned}$$

yielding the result.  $\square$

The assumptions on  $\rho$  is of course satisfied in some cases, but if  $a$  is not infinite it often implies that  $V$  is uniformly convex at infinity. In order to extend the previous idea to the multi-dimensional setting we will introduce two different notions of lower bounds on the curvature, to control its behavior at infinity:

**Definition 3.6.** Let  $\rho$  satisfying (1.8). For  $x \in \mathbb{R}^n$  we define

$$\rho_{inf}(|x|) = \inf_{|y| \leq |x|} \rho(y) \quad \text{and} \quad \rho_{sup}(|x|) = \inf_{|y| \geq |x|} \rho(y).$$

We have

**Proposition 3.7.** *Assume that  $V$  is  $C^2$  and that (1.8) is satisfied for some  $\rho$  which is bounded from below by  $\rho_0 \geq 0$  (log-concave case). Then if  $\rho_{inf}$  (resp.  $\rho_{sup}$ ) is non-increasing (resp. non-decreasing) on  $(a, +\infty)$  for some  $a \geq 0$ , for all  $x$  and  $y$  in  $\mathbb{R}^n$ ,*

$$W_1(P_t^* \delta_x, P_t^* \delta_y) \leq |x - y| \mathbb{E}_x \left( e^{-\int_0^t \mathbf{1}_{|X_s| > a + |x-y|} \rho_{inf}(|X_s| + |x-y|) ds} \right)$$

respectively

$$W_1(P_t^* \delta_x, P_t^* \delta_y) \leq |x - y| \mathbb{E}_x \left( e^{-\int_0^t \mathbf{1}_{|X_s| > a + |x-y|} \rho_{sup}(|X_s| - |x-y|) ds} \right).$$

*Proof.* It is enough to remark that

$$\rho(\lambda z + (1 - \lambda)z') \geq \rho_{inf}(|z| + |z - z'|)$$

as soon as  $|\lambda z + (1 - \lambda)z'| \geq a$  and similarly for  $\rho_{sup}$  and that  $\rho \geq 0$  otherwise.  $\square$

#### 4. USING WI DEVIATION INEQUALITIES.

##### 4.1. Transport-information inequalities.

As said in the introduction we shall extensively use the results of [27] on functional inequalities and concentration estimates for additive functionals. To this end we introduce some notations.

Let  $c$  be a *cost function* defined on  $\mathbb{R}^n \otimes \mathbb{R}^n$ , i.e  $c$  is lower semicontinuous, non negative and satisfies  $c(x, x) = 0$  for all  $x$ . We shall also only consider cases where  $c(x, y) = c(y, x)$ . We may consider the *transportation cost*  $T_c$  defined by

$$T_c(\nu, \mu) = \inf_{\{\pi, \pi \circ x^{-1} = \nu, \pi \circ y^{-1} = \mu\}} \int c(x, y) \pi(dx, dy). \quad (4.1)$$

When  $c(x, y) = |x - y|^p$ , we recover the usual  $W_p$  distance, but other costs may also be of interest for our purpose here. The Kantorovich duality theorem states that

$$T_c(\nu, \mu) = \sup_{(u, v) \in \Phi_c} \nu(u) - \mu(v), \quad (4.2)$$

where

$$\Phi_c := \{(u, v) \text{ Borel and bounded s.t. } u(x) - v(y) \leq c(x, y), \forall (x, y)\}.$$

Also recall the definition of the Fischer information: for  $\nu = f\mu$  with  $f > 0$  and locally Lipschitz

$$I(\nu|\mu) := \frac{1}{4} \mu \left( \frac{|\nabla f|^2}{f} \right) = \mu \left( |\nabla(\sqrt{f})|^2 \right). \quad (4.3)$$

If  $\nu$  does not satisfy the previous assumptions then,  $I(\nu|\mu) = +\infty$ .

The following is a consequence of Theorem 2.2 in [27], connecting functional inequalities and deviation estimates of additive functionals along drift-diffusion S.D.E.

**Theorem 4.1.** *Let  $\alpha : [0, +\infty) \rightarrow [0, +\infty)$  be a continuous increasing function with  $\alpha(0) = 0$ . Then the following statements are equivalent.*

- (1) *The following transport-information  $WI(\alpha, c)$  inequality is satisfied*

$$\alpha(T_c(\nu, \mu)) \leq I(\nu|\mu), \quad \forall \nu.$$

- (2) *For any initial measure  $\beta \ll \mu$  such that  $d\beta/d\mu \in \mathbb{L}^2(\mu)$ , all  $u, v$  such that  $(u, v) \in \Phi_c$ , all  $R, t > 0$ ,*

$$\mathbb{P}_\beta \left( \frac{1}{t} \int_0^t u(X_s) ds \geq \mu(v) + R \right) \leq \|d\beta/d\mu\|_{\mathbb{L}^2(\mu)} e^{-t\alpha(R)}.$$

(3) *Defining*

$$P_t^{\lambda u} f(x) = \mathbb{E}_x \left[ f(X_t) e^{\int_0^t \lambda u(X_s) ds} \right],$$

then for all  $u, v$  such that  $(u, v) \in \Phi_c$  and all  $t > 0$  it holds

$$\| P_t^{\lambda u} f \|_{\mathbb{L}^2(\mu)} \leq e^{t(\lambda \mu(v) + \alpha^*(\lambda))} \| f \|_{\mathbb{L}^2(\mu)},$$

where  $\alpha^*$  is the Legendre transform of  $\alpha \mathbf{1}_{[0, \infty)}$ , restricted to  $[0, \infty)$ .

In particular, for any initial measure  $\beta \ll \mu$  such that  $d\beta/d\mu \in \mathbb{L}^2(\mu)$ ,

$$\mathbb{E}_\beta \left[ \exp \left( \int_0^t \lambda u(X_s) ds \right) \right] \leq \| d\beta/d\mu \|_{\mathbb{L}^2(\mu)} \exp[t(\lambda \mu(v) + \alpha^*(\lambda))].$$

When curvature is bounded from below by a positive constant, a quadratic transport-information inequality holds, with the  $L^2$  Wasserstein distance. Concentration estimates for additive functionals can also be directly derived from positive curvature [29]. However, positive curvature is not a necessary condition, and we can also use non-quadratic inequalities. There are also direct approaches starting from Poincaré inequality such as [17] with bounded curvature or [5] only requiring curvature bounded from below, but whose constants may not lead to a strict improvement of our results.

**Definition 4.2.** Let  $c$  be a cost function and  $u$  a measurable function. We define

$$\| u \|_c := \sup_{x \neq y} \frac{|u(x) - u(y)|}{c(x, y)} \in [0, \infty].$$

*Remark 4.3.* Notice that a simple monotone convergence argument allows us to extend Theorem 4.1 to non necessarily bounded functions  $u$  such that  $\| u \|_c = 1$ . This will be used in the sequel without further mention.  $\diamond$

When  $u$  is bounded from below we can derive another bound

**Corollary 4.4.** Let  $\alpha : [0, +\infty) \rightarrow [0, +\infty)$  be a continuous increasing function with  $\alpha(0) = 0$ . Let  $u$  be a measurable function bounded from below by  $u_0 \in \mathbb{R}$  and such that  $\mu(u) > 0$ . Then for all initial measure  $\beta \ll \mu$  such that  $d\beta/d\mu \in \mathbb{L}^2(\mu)$ , all  $\lambda$  and  $t$  strictly positive, all  $1 > \varepsilon > 0$

$$\mathbb{E}_\beta \left[ \exp \left( - \int_0^t \lambda u(X_s) ds \right) \right] \leq (1 + \| d\beta/d\mu \|_{\mathbb{L}^2(\mu)}) \max \left( e^{-\lambda \varepsilon \mu(u) t}, e^{-t(\lambda u_0 + \alpha((1-\varepsilon)\mu(u)/\|u\|_c))} \right).$$

*Proof.* Introduce the set

$$A = \left\{ - \int_0^t \lambda (u - \mu(u))(X_s) ds \geq Rt \right\}.$$

We thus have, using Theorem 4.1,

$$\begin{aligned} E_\beta \left( e^{-\int_0^t \lambda u(X_s) ds} \right) &\leq e^{-\lambda u_0 t} \mathbb{P}_\beta(A) + e^{-\lambda \mu(u) t} E_\beta \left( e^{-\int_0^t \lambda (u - \mu(u))(X_s) ds} \mathbf{1}_{A^c} \right) \\ &\leq \| d\beta/d\mu \|_{\mathbb{L}^2(\mu)} e^{t(-\lambda u_0 - \alpha(R/\lambda \|u\|_c))} + e^{-(\lambda \mu(u) - R)t}. \end{aligned} \quad (4.4)$$

Choosing  $R = \lambda \mu(u) (1 - \varepsilon)$  gives the result.  $\square$

*Remark 4.5.* It is also possible to obtain exponential convergence of additive functionals via Lyapunov-type conditions, see for example [31, 24], but the constants obtained there are not so explicit, and typically depend on the dimension, so they are not so suitable for our purpose here.

#### 4.2. Rates of convergence in $W_1$ distance.

We start with an immediate application of what precedes:

**Theorem 4.6.** *Assume that  $V$  is  $C^2$ , that (3.1) is satisfied with some function  $\kappa$  which satisfies  $\|\kappa\|_c < +\infty$ , bounded from below by some constant  $\kappa_0 \in \mathbb{R}$  and such that  $\mu(\kappa) > 0$ . Assume that  $\mu$  satisfies a WI( $\alpha, c$ ) inequality for some function  $\alpha$ . Then, for all  $\nu$  such that  $d\nu/d\mu \in \mathbb{L}^2(\mu)$ , all  $t > 0$ ,*

$$W_1(P_t^* \nu, \mu) \leq C(\nu) e^{-\theta t},$$

in the following cases:

(1) If  $\alpha^*$  satisfies

$$\mu(\kappa) > \frac{\alpha^*(2 \|\kappa\|_c)}{2}, \quad (4.5)$$

then for any  $r > 2$  such that  $\mu(\kappa) > \frac{\alpha^*(r \|\kappa\|_c)}{r}$ , and  $p = r/(r-2)$ , we may choose

$$C(\nu) = \|d\nu/d\mu\|_{\mathbb{L}^2(\mu)}^{\frac{1}{r}} W_p(\nu, \mu)$$

and

$$\theta = 2(\mu(\kappa) - (\alpha^*(r \|\kappa\|_c)/r)).$$

(2) If  $\alpha$  satisfies

$$2\kappa_0 + \alpha(\mu(\kappa)/\|\kappa\|_c) > 0, \quad (4.6)$$

then for any  $r > 2$ ,  $1 > \varepsilon > 0$  such that  $r\kappa_0 + \alpha((1-\varepsilon)\mu(\kappa)/\|\kappa\|_c) > 0$ , and  $p = r/(r-2)$ , we may choose

$$C(\nu) = (1 + \|d\nu/d\mu\|_{\mathbb{L}^2(\mu)})^{\frac{1}{r}} W_p(\nu, \mu),$$

and

$$\theta = 2 \min \left( \varepsilon \mu(\kappa), \kappa_0 + \frac{1}{r} \alpha((1-\varepsilon)\mu(\kappa)/\|\kappa\|_c) \right).$$

*Proof.* According to Proposition 3.1, it is enough to control  $\mathbb{E}_\nu^{\frac{1}{r}} \left( e^{-\int_0^t r \kappa(X_s) ds} \right)$  for some well chosen  $r \geq 2$ .

Part (1) is an immediate consequence of (3) in Theorem 4.1 applied with  $u = -\kappa/\|\kappa\|_c$  and  $\lambda = r \|\kappa\|_c$ .

Part (2) follows from Corollary 4.4. When  $r\kappa_0 + \alpha((1-\varepsilon)\mu(\kappa)/\|\kappa\|_c) > 0$ , both exponential terms go to 0, so that the exponential decay rate is given by

$$\min \left( \varepsilon \mu(\kappa), \kappa_0 + \frac{1}{r} \alpha((1-\varepsilon)\mu(\kappa)/\|\kappa\|_c) \right).$$

□

We thus have a balance between the rate of exponential convergence and the initial control i.e.  $r$  and  $p$  or  $\varepsilon$ . Note that since we want some regularity for  $\kappa$ , it may be that the one to use is not the one that optimizes (3.1), but some smoother minorant.

One can notice that in case (2) the rate is always less than or equal to  $\varepsilon \mu(\kappa)$  and that it can be achieved by choosing

$$r(\varepsilon) = \frac{\alpha((1 - \varepsilon)\mu(\kappa)/\|\kappa\|_c)}{\varepsilon\mu(\kappa) - \kappa_0},$$

provided the latter is larger than 2 and  $\varepsilon\mu(\kappa) - \kappa_0 > 0$ . We shall come back to this issue in Section 5.

In order to extend the result to more general initial measures we will use in addition the following regularization result

**Proposition 4.7.** *Assume that  $V$  is  $C^2$  and that the process is conservative. Then for all  $x \in \mathbb{R}^n$  and  $t > 0$ ,  $P_t^* \delta_x(dy) \ll dy$  and its density  $p(t, x, y)$  is continuous (in  $y$ ).*

*We will denote  $r(t, x, y) = p(t, x, y) e^{V(y)}$  the density of  $P_t^* \delta_x$  w.r.t.  $\mu$ . Notice that  $r(t, x, y) = r(t, y, x)$ .*

This result is certainly well known in P.D.E. theory. The assumptions we are using are those of [38, Proposition 4.2], where the proof is based on Malliavin calculus.

Then, as argued in [20], symmetry combined with the Chapman-Kolmogorov relation and the continuity of  $r$  yield for all  $t > 0$ ,

$$\int r^2(t, x, y) \mu(dy) = \int r(t, x, y) r(t, y, x) \mu(dy) = r(2t, x, x) < +\infty.$$

If  $\nu$  is any probability measure,  $P_t^* \nu$  is thus absolutely continuous w.r.t.  $\mu$  with density  $r(t, \nu, y) = \int r(t, x, y) \nu(dx)$  which belongs to  $\mathbb{L}^2(\mu)$  as soon as  $\int r(2t, x, x) \nu(dx) < +\infty$ .

It follows that for all  $\eta > 0$ ,

$$\begin{aligned} \mathbb{E}_\nu \left( e^{-r \int_0^t \kappa(X_s) ds} \right) &= \mathbb{E}_\nu \left( e^{-r \int_0^\eta \kappa(X_s) ds} \mathbb{E}_{X_\eta} \left( e^{-r \int_0^{t-\eta} \kappa(X'_s) ds} \right) \right) \\ &\leq e^{-r \kappa_0 \eta} \mathbb{E}_{P_\eta^* \nu} \left( e^{-r \int_0^{t-\eta} \kappa(X_s) ds} \right) \end{aligned}$$

Hence as before we easily get

$$\begin{aligned} \mathbb{E}_\nu \left( e^{-\int_0^t \kappa(X_s) ds} \right) &\leq \mathbb{E}_\nu^{\frac{1}{r}} \left( e^{-r \int_0^t \kappa(X_s) ds} \right) \\ &\leq e^{-\kappa_0 \eta} \mathbb{E}_{P_\eta^* \nu}^{\frac{1}{r}} \left( e^{-r \int_0^{t-\eta} \kappa(X_s) ds} \right) \end{aligned}$$

and we may argue as in the previous proof (using that the  $\mathbb{L}^2$  norm of a probability density is larger than 1) to obtain

**Corollary 4.8.** *The conclusions of Theorem 4.6, are true for any probability measure  $\nu$ , any  $t > \eta > 0$ , with the same  $\theta$  but*

$$C(\nu) = 2^{\frac{1}{r}} \left( \int r(2\eta, x, x) \nu(dx) \right)^{\frac{1}{2r}} e^{(\theta - \kappa_0) \eta} W_p(\nu, \mu),$$

where  $r(\eta, \dots)$  is defined in proposition 4.7.

In order to consider the more general curvature assumption (1.8) we would have to get uniform bounds in  $x$  when the initial measure is  $\delta_x$  (see (1.9)). We can follow the same path as before provided  $\rho$  satisfies the same assumptions as  $\kappa$ . Notice that this time we only need  $r \geq 1$ .

**Theorem 4.9.** *Assume that  $V$  is  $C^2$ , that (1.8) is satisfied with some function  $\rho$  which satisfies  $\|\rho\|_c < +\infty$ , is bounded from below by some constant  $\rho_0 \in \mathbb{R}$  and such that  $\mu(\rho) > 0$ . Assume moreover that  $\rho$  is continuous or  $V$  is  $C^3$  and  $\rho$  is locally bounded. Finally, assume that  $\mu$  satisfies a  $WI(\alpha, c)$  inequality for some function  $\alpha$ . Then, for all  $\nu$ , all  $\eta > 0$  and all  $t > \eta$ ,*

$$W_1(P_t^* \nu, \mu) \leq C(\nu) e^{-\theta t} W_1(\nu, \mu),$$

with

$$C(\nu) = 2^{\frac{1}{r}} \left( \sup_x r(2\eta, x, x) \right)^{\frac{1}{2r}} e^{(\theta - \rho_0)\eta},$$

in the following cases:

(1) If  $\alpha^*$  satisfies

$$\mu(\rho) > \alpha^*(\|\rho\|_c), \quad (4.7)$$

for all  $r \geq 1$  such that  $\mu(\rho) > \frac{\alpha^*(r\|\rho\|_c)}{r}$ , we may choose

$$\theta = (\mu(\rho) - (\alpha^*(r\|\rho\|_c)/r)).$$

(2) If  $\alpha$  satisfies

$$\rho_0 + \alpha(\mu(\rho)/\|\rho\|_c) > 0, \quad (4.8)$$

for all  $r \geq 1$ ,  $1 > \varepsilon > 0$  such that  $r\rho_0 + \alpha((1 - \varepsilon)\mu(\rho)/\|\rho\|_c) > 0$ , we may choose

$$\theta = \min \left( \varepsilon \mu(\rho), \rho_0 + \frac{1}{r} \alpha((1 - \varepsilon)\mu(\rho)/\|\rho\|_c) \right).$$

*Remark 4.10.* One can think that the assumption used here on the density  $r(2\eta, \cdot, \cdot)$  is too strong. Consider for example the Ornstein-Uhlenbeck corresponding to the standard gaussian measure. Then

$$r(t, x, x) = c(t) e^{-\frac{1}{2}|x|^2 \left( \frac{(1-e^{-t})^2}{1-e^{-2t}} - 1 \right)} = c(t) e^{|x|^2 e^{-t}(1-e^{-t})},$$

so that it is not bounded. However if the semi-group is *ultra-bounded*, i.e. maps continuously  $\mathbb{L}^1(\mu)$  in  $\mathbb{L}^\infty(\mu)$  for all  $t > 0$ , then, since  $r(\eta, x, \cdot)$  is a density of probability w.r.t.  $\mu$ ,  $r(2\eta, x, \cdot)$  is bounded, with a sup norm only depending on  $\eta$ , and not of  $x$ . Actually we only need the  $(\mathbb{L}^1, \mathbb{L}^\infty)$  continuity for a single time  $\eta$ .  $\diamond$

### 4.3. Entropic pre-factor.

When curvature is nonnegative, we can get a variant of Theorem 4.6 with an entropic pre-factor instead of an  $L^2$  norm, but at the cost of weakening the rate of convergence. We recall the definition of the relative entropy functional

$$H(\nu|\mu) := \int g \log g d\mu; \quad \text{if } d\nu = g d\mu$$

and takes value  $+\infty$  if  $\nu$  is not absolutely continuous with respect to  $\mu$ .



**Theorem 4.11.** *Assume that the transport-information inequality*

$$\alpha(T_c(\nu, \mu)) \leq I(\nu|\mu)$$

*holds, and that  $\|\kappa\|_c < \infty$ . Assume moreover that  $\kappa \geq \kappa_0$  for some  $\kappa_0 \in \mathbb{R}$  and that there exists  $q > 1$  such that*

$$(q\|\kappa\|_c)^{-1}\alpha^*(q\|\kappa\|_c) \leq \mu(\kappa) + \kappa_0.$$

*Then for any  $r, p$  with  $\frac{1}{p} + \frac{1}{q} + \frac{1}{q} = 1$  and  $\varepsilon \in (0, 1)$  such that  $(1-\varepsilon)\mu(\kappa) > (q\|\kappa\|_c)^{-1}\alpha^*(q\|\kappa\|_c)$ , we have*

$$W_1(P_t^* \nu, \mu) \leq W_p(\nu, \mu) (H_1(t) + H_2(t))$$

*with*

$$H_1(t) = \left( \frac{\log 2 + H(\nu|\mu)}{t\alpha((1-\varepsilon)/\|\kappa\|_c)} \right)^{1/r} e^{-(\mu(\kappa) + \kappa_0 - (q\|\kappa\|_c)^{-1}\alpha^*(q\|\kappa\|_c))t},$$

*and*

$$H_2(t) = e^{-((1+\varepsilon)\mu(\kappa) - (q\|\kappa\|_c)^{-1}\alpha^*(q\|\kappa\|_c))t}.$$

Since  $H(\nu|\mu) \leq \log(1 + \|\rho\|_{L^2(\mu)})$ , this entropic pre-factor is often much smaller than the  $L^2$  pre-factor of Theorem 4.6.

*Proof.* We shall use the classical entropy inequality

$$P(A) \leq \frac{\log 2 + H(P|Q)}{\log(1 + 1/Q(A))} \quad (4.9)$$

for any two probability measures  $P$  and  $Q$ , and event  $A$ . It can be derived as a consequence of the dual formulation

$$H(P|Q) = \sup_f P(f) - \log Q(e^f)$$

by taking  $f = \lambda \mathbf{1}_A$  and then optimizing in  $\lambda$ .

We wish to apply (4.9) with  $P$  the distribution of a trajectory of the diffusion process with initial data distributed according to  $\nu$ , and  $Q$  the law of a trajectory of the same process, but started from the equilibrium distribution  $\mu$ . It is known, for example as a consequence of the Girsanov formula [35], that in this case

$$H(P|Q) = H(\nu|\mu).$$

If we take  $A$  the event  $\int_0^t \kappa(X_s) ds \geq \mathbb{E}_\mu(\kappa) + r$ , we get

$$\mathbb{P}(A) \leq (\log 2 + H(\nu|\mu)) / (t\alpha(r)).$$

The proof of the first part then proceeds as in the proof of Theorem 4.6. Combining the transport-information inequality and the above bound, the counterpart to (4.4) for  $r, R$  and  $A$  as in the proof of Theorem 4.6, is

$$\begin{aligned} \mathbb{E}_\nu \left( e^{-\int_0^t r \kappa(X_s) ds} \right) &\leq e^{-r \kappa_0 t} \mathbb{P}_\nu(A) + e^{-r \mu(\kappa) t} \mathbb{E}_\nu \left( e^{-\int_0^t r (\kappa - \mu(\kappa))(X_s) ds} \mathbf{1}_{A^c} \right) \\ &\leq \frac{\log 2 + H(\nu|\mu)}{t\alpha(R/(r\|\kappa\|_c))} e^{-r \kappa_0 t} + e^{-(r \mu(\kappa) - R) t}. \end{aligned} \quad (4.10)$$

We then take as before,  $R = (1-\varepsilon)r\mu(\kappa)$  and get

$$\mathbb{E}_\nu \left( e^{-\int_0^t r \kappa(X_s) ds} \right)^{1/r} \leq \left( \frac{\log 2 + H(\nu|\mu)}{t\alpha((1-\varepsilon)/\|\kappa\|_c)} \right)^{1/r} e^{-\kappa_0 t} + e^{-\varepsilon \mu(\kappa) t}.$$

Moreover,

$$\mathbb{E}_\mu \left[ \exp \left( \int_0^t -q\kappa(X_s) ds \right) \right]^{1/q} \leq \exp(-t\mu(\kappa) + t(q\|\kappa\|_c)^{-1}\alpha^*(q\|\kappa\|_c)).$$

Using Proposition 3.1 then concludes the proof.  $\square$

We could also straightforwardly derive results similar to Corollary 4.8 and Theorem 4.9, but for the sake of brevity we do not do so here.

#### 4.4. Bounds on the Poincaré constant.

Recall that according to the discussion in the introduction, thanks to [20, Lemma 2.12],  $C_P(\mu) \leq \frac{2}{c}$  as soon as

$$\mathbb{E}_\mu \left( e^{-2 \int_0^t \rho(X_s) ds} \right) \leq C e^{-ct}.$$

For  $r > 2$  we may use

$$\mathbb{E}_\mu \left( e^{-\int_0^t 2\rho(X_s) ds} \right) \leq \mathbb{E}_\mu^{\frac{2}{r}} \left( e^{-\int_0^t r\rho(X_s) ds} \right),$$

together with Theorem 4.1 or Corollary 4.4. This yields

**Theorem 4.12.** *Assume that the assumptions of Theorem 4.9 are fulfilled.*

(1) *if  $\alpha^*$  satisfies*

$$\mu(\rho) > \alpha^*(2\|\rho\|_c)/2, \tag{4.11}$$

*then*

$$C_P(\mu) \leq \frac{1}{\mu(\rho) - \alpha^*(2\|\rho\|_c)/2}.$$

(2) *If  $\alpha$  satisfies*

$$2\rho_0 + \alpha(\mu(\rho)/\|\rho\|_c) > 0, \tag{4.12}$$

*then for all  $r > 2$ ,  $1 > \varepsilon > 0$  such that  $r\rho_0 + \alpha((1-\varepsilon)\mu(\rho)/\|\rho\|_c) > 0$ ,*

$$C_P(\mu) \leq \frac{1}{\min(\varepsilon\mu(\rho), \rho_0 + (\alpha((1-\varepsilon)\mu(\rho)/\|\rho\|_c)/r))}.$$

*If  $c$  is continuous, the optimal  $\varepsilon$  satisfies*

$$\alpha((1-\varepsilon)\mu(\rho)/\|\rho\|_c) = 2(\varepsilon\mu(\rho) - \rho_0)$$

*provided the right hand side is non-negative and in this case  $C_P(\mu) \leq 1/(\varepsilon\mu(\rho))$ .*

For the last assertion, first remark that for a given  $\varepsilon$  the optimal  $r$  is such that both terms in the *min* are equal, i.e.

$$r(\varepsilon) = \frac{\alpha((1-\varepsilon)\mu(\rho)/\|\rho\|_c)}{\varepsilon\mu(\rho) - \rho_0}$$

provided  $r(\varepsilon) \geq 2$ . Then remark that  $\varepsilon \mapsto h(\varepsilon) = \alpha((1-\varepsilon)\mu(\rho)/\|\rho\|_c)$  is non increasing,  $\varepsilon \mapsto g(\varepsilon) = 2(\varepsilon\mu(\rho) - \rho_0)$  is non decreasing, they are both continuous and  $h(0) > g(0)$  while  $h(1) < g(1)$  yielding the optimal  $\varepsilon$ .

*Remark 4.13.* If  $V$  is not smooth enough, the previous result is still true in the following situation: there exists a family  $V_\eta$  of smooth potentials (for instance  $V * \gamma_\eta$  where  $\gamma_\eta$  is a centered gaussian kernel with covariance matrix  $\eta Id$ ) and a bounded  $\rho$  such that  $\rho_\eta \geq \rho$  for all  $\eta$  and the measure  $e^{-V_\eta} dx$  converges weakly to  $\mu$  as  $\eta \rightarrow 0$ .  $\diamond$

*Remark 4.14.* We will now discuss some possible choices for the cost  $c$ . Notice the transport-information  $WI(\alpha, c)$  is only important for the single function  $\rho$  (or  $\kappa$ ). A natural choice will thus be

$$c(x, y) = |\rho(x) - \rho(y)|, \quad (4.13)$$

so that  $\|\rho\|_c = 1$ . The difficulty will be to find tractable conditions for  $WI(\alpha, c)$ , in particular because  $c$  is not necessarily a distance.

The most classical choice is  $c(x, y) = |x - y|$  so that  $\|u\|_c = \|u\|_{Lip} = \|\nabla u\|_\infty$ .  $WI(\alpha, c)$  then becomes the following  $W_1 I(\alpha)$  inequality

$$\alpha(W_1(\nu, \mu)) \leq I(\nu|\mu), \quad \forall \nu. \quad (4.14)$$

Of course, in many situations one may replace the *natural*  $\rho$  or  $\kappa$  by some smaller function. For instance, if we consider  $V(x) = |x|^4$  as in remark 3.3, we may replace the natural  $\kappa(x) = 2|x|^2$  by  $\kappa(x) = 2 \min(|x|^2, |x|)$  which is (globally) Lipschitz, and thus apply the previous results provided  $W_1 I(\alpha)$  is satisfied.

If we choose the Hamming distance  $c(x, y) = \mathbf{1}_{x \neq y}$ ,  $T_c(\nu, \mu) = \frac{1}{2} \|\nu - \mu\|_{TV}$  and  $\|u\|_c = \text{Osc}(u) = \sup_{x, y} |u(x) - u(y)|$ .  $\diamond$

**Remark 4.15. About homogeneity.**

It is well known that the Poincaré inequality is 2-homogeneous w.r.t. dilations. This means the following: for  $\lambda > 0$ , define  $\mu_\lambda(f) = \int f(\lambda x) \mu(dx)$  so that

$$\begin{aligned} \mu_\lambda(dx) &= \frac{1}{\lambda} e^{-V(x/\lambda)} dx := e^{-V_\lambda(x)} dx, \\ V_\lambda(x) &= \ln(\lambda) + V(x/\lambda), \quad \nabla V_\lambda(x) = \frac{1}{\lambda} \nabla V(x/\lambda), \quad \rho_\lambda(x) = \frac{1}{\lambda^2} \rho(x/\lambda), \\ \mu_\lambda(\rho_\lambda) &= \int \rho_\lambda(\lambda x) \mu(dx) = \frac{1}{\lambda^2} \mu(\rho), \quad \rho_{\lambda,0} = \frac{1}{\lambda^2} \rho_0 \\ \|\rho_\lambda\|_{Lip} &= \frac{1}{\lambda^3} \|\rho\|_{Lip}, \quad \text{Osc}(\rho_\lambda) = \frac{1}{\lambda^2} \text{Osc}(\rho), \end{aligned}$$

and finally

$$C_P(\mu_\lambda) = \lambda^2 C_P(\mu).$$

Of course the key point here is that the dilation only concerns the invariant measure and not the dynamics. In particular there is no immediate correspondence between the natural dynamics corresponding to  $\mu$  and its dilation, so that we cannot expect to get a direct comparison of the rate of convergence in  $W_1$  distance for both.

Now consider  $T_c(\mu_\lambda, \nu)$ . If we assume that  $c$  is homogeneous of degree  $\eta$ , i.e.

$$c(\lambda x, \lambda y) = \lambda^\eta c(x, y)$$

we have

$$T_c(\mu_\lambda, \nu) = \lambda^\eta T_c(\mu, \nu_{1/\lambda}).$$

It is easy to check that

$$I(\nu_{1/\lambda}|\mu) = \lambda^2 I(\nu|\mu_\lambda),$$

so that if  $\mu$  satisfies a  $WI(\alpha, c)$  inequality,

$$\begin{aligned} T_c(\nu, \mu_\lambda) &= \lambda^\eta T_c(\nu_{1/\lambda}, \mu) \leq \lambda^\eta \alpha^{-1}(I(\nu_{1/\lambda}|\mu)) \\ &\leq \lambda^\eta \alpha^{-1}(\lambda^2 I(\nu|\mu_\lambda)), \end{aligned}$$

and  $\mu_\lambda$  satisfies a  $WI(\alpha_\lambda, c)$  inequality with

$$\alpha_\lambda^{-1}(s) = \lambda^\eta \alpha^{-1}(\lambda^2 s).$$

For simplicity we will only consider the case  $\alpha(s) = cs^m$  for some  $m > 0$ , i.e.

$$\alpha_\lambda(s) = c \lambda^{-(m\eta+2)} s^m.$$

If we consider both cases,  $c(x, y) = |x - y|$ , yielding  $\eta = 1$ , or  $c(x, y) = \mathbf{1}_{x \neq y}$  yielding  $\eta = 0$ , it is easily seen that our Theorem 4.12 is homogeneous w.r.t. to dilations.  $\diamond$

*Remark 4.16.* The criterion  $\mu(\rho) > 0$  can be called positive *average curvature*. Recall that in the symmetric situation we are looking at, a Poincaré inequality is equivalent to an integral  $\Gamma_2$  criterion (see e.g [2] proposition 5.5.4), namely

**Proposition 4.17.** *The following statements are equivalent*

- For all smooth  $f$ ,

$$\mu(\|\text{Hess } f\|_2^2) + \mu(\langle \nabla f, \text{Hess } V \nabla f \rangle) \geq C \mu(|\nabla f|^2),$$

- The Poincaré constant satisfies

$$C_P(\mu) \leq 1/C,$$

where  $\|\text{Hess } f(x)\|_2$  denotes the Hilbert-Schmidt norm of  $\text{Hess } f(x)$ .

In particular if (1.8) is satisfied,  $C_P(\mu) \leq 1/C$  as soon as  $\mu(\rho|\nabla f|^2) \geq C \mu(|\nabla f|^2)$  for all smooth  $f$ . If we may choose a positive constant for  $\rho$  we recover the Bakry-Emery criterion, but if  $\rho(x) < -\varepsilon$  on a small ball, one can build some  $f$  such that the previous inequality is not satisfied. Hence our Theorem 4.12 is of a different nature.  $\diamond$

## 5. EXAMPLES AND APPLICATIONS.

### 5.1. About $W_1I$ inequalities.

Since the results of the previous section are based on  $W_1I$  inequalities, it is important to have some sufficient conditions for these inequalities to be satisfied. Actually not so much is known and we recall below the main examples:

**Proposition 5.1.** *We have the following properties*

- (1) (see [27, Theorem 3.1]) Assume that  $\mu$  satisfies a Poincaré inequality with optimal constant  $C_P(\mu)$ , then  $\mu$  satisfies a  $W_1I$  inequality

$$\| \nu - \mu \|_{TV}^2 \leq 4C_P(\mu) I(\nu|\mu) \quad \forall \nu;$$

i.e. a  $WI(\alpha, c)$  inequality with  $\alpha(s) = (1/C_P(\mu)) s^2$  and  $c(x, y) = \mathbf{1}_{x \neq y}$ .

- (2) (see [42]) Assume that  $\mu$  satisfies a logarithmic Sobolev inequality with optimal constant  $C_{LS}(\mu)$ . Then  $\mu$  satisfies a  $W_2I$  inequality

$$W_2^2(\nu, \mu) \leq C_{LS}^2(\mu) I(\nu|\mu) \quad \forall \nu.$$

Consequently, since  $W_1 \leq W_2$ , it satisfies a  $WI(\alpha, c)$  inequality with  $\alpha(s) = (1/C_{LS}^2(\mu)) s^2$  and  $c(x, y) = |x - y|$ .

Of course similar inequalities are satisfied for any  $C \geq C_P(\mu)$  (resp.  $C_{LS}(\mu)$ ).

It may seem strange to derive bounds for the Poincaré inequality starting from such a bound, but we may hope that starting from an a priori bad bound, the method developed in the previous section will help to improve upon the constant.

We will thus assume that  $\mu$  satisfies some  $W_1I(2)$  inequality in the terminology of [27], i.e. we have

$$\alpha(s) = \frac{s^2}{C} \quad \text{so that} \quad \alpha^*(s) = \frac{C}{4} s^2. \quad (5.1)$$

This yields, using  $\alpha^*$ , the following values in our previous results:

- (1) in Theorem 4.6, for  $r > 2$ ,

$$\theta = 2 \left( \mu(\kappa) - \frac{C}{4} r \|\kappa\|_c^2 \right), \quad (5.2)$$

- (2) in Theorem 4.9,

$$\theta = \mu(\rho) - \frac{C}{4} \|\rho\|_c^2, \quad (5.3)$$

- (3) in Proposition 4.12,

$$C_P(\mu) \leq \frac{1}{\mu(\rho) - \frac{C}{2} \|\rho\|_c^2}, \quad (5.4)$$

provided  $\theta > 0$  or  $\mu(\rho) - \frac{C}{2} \|\rho\|_c^2 > 0$ .

Similarly, using  $\alpha$  instead of  $\alpha^*$ , we get:

- (1) in Theorem 4.6,  $\theta = 2\varepsilon \mu(\kappa)$  with  $r > 2$  satisfying

$$r = \frac{1}{C} \frac{1}{\varepsilon \mu(\kappa) - \kappa_0} \frac{(1 - \varepsilon)^2 \mu^2(\kappa)}{\|\kappa\|_c^2}, \quad (5.5)$$

- (2) in Theorem 4.9,  $\theta = \varepsilon \mu(\rho)$  with

$$1 = \frac{1}{C} \frac{1}{\varepsilon \mu(\rho) - \rho_0} \frac{(1 - \varepsilon)^2 \mu^2(\rho)}{\|\rho\|_c^2}, \quad (5.6)$$

(3) in Proposition 4.12,  $C_P(\mu) \leq \frac{1}{\varepsilon \mu(\rho)}$  with

$$2 = \frac{1}{C} \frac{1}{\varepsilon \mu(\rho) - \rho_0} \frac{(1 - \varepsilon)^2 \mu^2(\rho)}{\|\rho\|_c^2}, \quad (5.7)$$

provided  $0 < \varepsilon < 1$ .

In the latter cases, denoting by  $g$  either  $\kappa$  or  $\rho$  and by  $a$  either 2 or 1 we have

$$aC(\varepsilon\mu(g) - g_0) \|g\|_c^2 = (1 - \varepsilon)^2 \mu^2(g), \quad (5.8)$$

so that denoting by  $\eta = \varepsilon\mu(g)$ ,

$$h(\eta) = \eta^2 - \eta(2\mu(g) + aC \|g\|_c^2) + (\mu^2(g) + aCg_0 \|g\|_c^2) = 0.$$

Since  $h(0) = \mu^2(g) + aCg_0 \|g\|_c^2$  and  $h(\mu(g)) = aC \|g\|_c^2 (g_0 - \mu(g)) < 0$ , we always have a solution  $0 < \varepsilon < 1$ , provided  $h(0) > 0$ , solution given by

$$\varepsilon\mu(g) = \left( \mu(g) + \frac{a}{2} C \|g\|_c^2 \right) \left( 1 - \sqrt{1 - \frac{\mu^2(g) + aCg_0 \|g\|_c^2}{(\mu(g) + \frac{a}{2} C \|g\|_c^2)^2}} \right). \quad (5.9)$$

One can of course compare the bounds obtained above with  $\alpha^*$  and  $\alpha$ . For simplicity denote by  $\beta^* = \mu(g) - \frac{aC}{4} \|g\|_c^2$  the bound obtained with  $\alpha^*$  and by  $\beta$  the bound obtained with  $\alpha$  in (5.9). We have:

- (1) if  $\beta^* \leq 0$  and  $\mu^2(g) + aCg_0 \|g\|_c^2 \leq 0$ , then neither of the bounds are available,
- (2) if  $\beta^* \leq 0$  and  $\mu^2(g) + aCg_0 \|g\|_c^2 > 0$ , then we may use  $\beta$ ,
- (3) if  $\beta^* > 0$  and  $\mu^2(g) + aCg_0 \|g\|_c^2 \leq 0$ , then we may use  $\beta^*$ ,
- (4) if  $\beta^* > 0$  and  $\mu^2(g) + aCg_0 \|g\|_c^2 > 0$ , then  $\beta^* \leq \beta$  provided  $\mu(g) \leq \frac{5a}{16} C \|g\|_c^2 + g_0$ , and the converse inequality holds otherwise.

The proof of (4) is elementary. These results clearly indicate that both approaches may be of interest.

## 5.2. The positive curvature case.

Let us illustrate first the previous results when  $\rho_0 > 0$ .

In this case we already know that a lot of inequalities are satisfied, in particular

$$W_1(P_t^* \nu, \mu) \leq e^{-\rho_0 t} W_1(\nu, \mu) \quad (5.10)$$

according to [45] and

$$C_P(\mu) \leq \frac{1}{\rho_0} \quad \text{and} \quad C_{LS}(\mu) \leq \frac{2}{\rho_0} \quad (5.11)$$

according to the Bakry-Emery criterion. As we have previously mentioned, this statement is rigid, and under a non-trivial variable curvature bound the constants should get strictly better. We shall see what improvements can be obtained using the previous results.

Since  $\rho_0 > 0$ , (5.8) always admits a solution for some  $\varepsilon \in (0, 1)$ . In order to compare with the previous bounds one can solve (5.8) using  $\eta = \varepsilon\mu(\rho) - \rho_0$ , yielding another expression

equivalent to (5.9)

$$\varepsilon \mu(\rho) = \rho_0 + \left( (\mu(\rho) - \rho_0) + \frac{a}{2} C \|\rho\|_c^2 \right) \left( 1 - \sqrt{1 - \frac{(\mu(\rho) - \rho_0)^2}{((\mu(\rho) - \rho_0) + \frac{a}{2} C \|\rho\|_c^2)^2}} \right). \quad (5.12)$$

We have thus obtained the following results (using  $\alpha$ ):

**Theorem 5.2.** *Assume that the assumptions of Theorem 4.9 are satisfied. Assume in addition that  $\rho$  is bounded from below by some constant  $\rho_0 > 0$ .*

*Then, for all  $\nu$ , all  $t > \eta > 0$ ,*

$$W_1(P_t^* \nu, \mu) \leq 2 \left( \sup_x r(2\eta, x, x) \right)^{\frac{1}{2}} e^{(\theta - \rho_0)\eta} e^{-\theta t} W_1(\nu, \mu),$$

where  $\theta$  is given by one of the following expressions:

(1) either

$$\theta = \rho_0 + \left( (\mu(\rho) - \rho_0) + \frac{1}{2\rho_0} \text{Osc}^2(\rho) \right) \left( 1 - \sqrt{1 - \frac{(\mu(\rho) - \rho_0)^2}{((\mu(\rho) - \rho_0) + \frac{1}{2\rho_0} \text{Osc}^2(\rho))^2}} \right),$$

(2) or

$$\theta = \rho_0 + \left( (\mu(\rho) - \rho_0) + \frac{2}{\rho_0^2} \|\rho\|_{Lip}^2 \right) \left( 1 - \sqrt{1 - \frac{(\mu(\rho) - \rho_0)^2}{((\mu(\rho) - \rho_0) + \frac{2}{\rho_0^2} \|\rho\|_{Lip}^2)^2}} \right).$$

**Theorem 5.3.** *Similarly, under the assumptions of Theorem 4.12 and if  $\rho_0 > 0$ ,*

$$C_P(\mu) \leq \frac{1}{\rho_0 + \varepsilon'}$$

with

(1) either

$$\varepsilon' = \left( (\mu(\rho) - \rho_0) + \frac{1}{\rho_0} \text{Osc}^2(\rho) \right) \left( 1 - \sqrt{1 - \frac{(\mu(\rho) - \rho_0)^2}{((\mu(\rho) - \rho_0) + \frac{1}{\rho_0} \text{Osc}^2(\rho))^2}} \right),$$

(2) or

$$\varepsilon' = \left( (\mu(\rho) - \rho_0) + \frac{4}{\rho_0^2} \|\rho\|_{Lip}^2 \right) \left( 1 - \sqrt{1 - \frac{(\mu(\rho) - \rho_0)^2}{((\mu(\rho) - \rho_0) + \frac{4}{\rho_0^2} \|\rho\|_{Lip}^2)^2}} \right).$$

What is interesting here is that one (almost) always obtains improvements of the Bakry-Emery constants, in agreement with the rigidity phenomenon [21]. Of course, for the convergence in  $W_1$  distance some constant has to appear in front of the exponential rate. Also notice that when  $\rho_0 \rightarrow 0$  or when  $\text{Osc}(\rho)$  or  $\|\rho\|_{Lip} \rightarrow +\infty$ ,  $\varepsilon'$  and  $\theta$  go to 0. Also notice that if  $\rho$  is constant, its oscillation and its Lipschitz norm vanish so that  $\theta$  and  $\varepsilon'$  are equal to 0.

As explained before, in some cases it is possible to obtain better constants by using  $\alpha^*$  instead of  $\alpha$ . Explicit computations are left to the interested reader.

*Remark 5.4.* Recall that if the semi-group is *ultra-bounded*, then  $\sup_x r(2\eta, x, x)$  is finite. Examples of ultra-bounded semi groups are well known. For example if  $V$  is bounded below, goes to infinity at infinity and satisfies

$$\frac{1}{2} |\nabla V|^2 - \Delta V \geq c|V|^\gamma - d \mathbf{1}_{|x| \leq R}$$

for some  $R > 0$ ,  $c, d > 0$  and  $\gamma > 1$ , the semi-group is ultra-bounded (see [16] Remark 5.5). With  $\gamma = 1$  one gets a log-Sobolev inequality, and with  $\gamma = 0$  a Poincaré inequality. Other conditions, such as Foster-Lyapunov conditions, may also be used.  $\diamond$

### 5.3. Using the log-Sobolev constant.

In this subsection we will assume that  $\mu$  satisfies a log-Sobolev inequality with constant  $C_{LS}(\mu)$ . In this situation it is well known that it also satisfies a Talagrand inequality  $T_2(C_{LS}(\mu))$ , so that we have

$$W_1(P_t^* \nu, \mu) \leq W_2(P_t^* \nu, \mu) \leq \sqrt{C_{LS}(\mu) H(P_t^* \nu | \mu)} \leq \sqrt{C_{LS}(\mu)} e^{-(1/C_{LS}(\mu))t} \sqrt{H(\nu, \mu)}. \quad (5.13)$$

According to [11] remark 2.8, one can improve on the previous inequalities and show that

$$W_1(P_t^* \nu, \mu) \leq W_2(P_t^* \nu, \mu) \leq K(\mu) e^{-(1/C_{LS}(\mu))t} W_2(\nu, \mu). \quad (5.14)$$

However we cannot replace  $W_2$  by  $W_1$  in the right hand side. According to Theorem 4.6, we can also deduce a similar result, using our method.

But according to what precedes, Theorem 5.2 extends to the case  $\rho_0 \leq 0$ , yielding

$$W_1(P_t^* \nu, \mu) \leq C(\mu) e^{-\theta t} W_1(\nu, \mu)$$

in the following two situations,

- (1) if the log-Sobolev constant satisfies

$$C_{LS}^2(\mu) < \frac{4\mu(\rho)}{\|\rho\|_{Lip}^2}, \quad (5.15)$$

we may choose  $\theta = \mu(\rho) - \frac{C_{LS}^2(\mu)}{4} \|\rho\|_{Lip}^2$ ,

- (2) if the log-Sobolev constant satisfies

$$C_{LS}^2(\mu) < \frac{\mu(\rho)}{|\rho_0| \|\rho\|_{Lip}^2}, \quad (5.16)$$

we may choose

$$\theta = \left( \mu(\rho) + \frac{1}{2} C_{LS}^2(\mu) \|\rho\|_{Lip}^2 \right) \left( 1 - \sqrt{1 - \frac{\mu^2(\rho) + C_{LS}^2(\mu) \rho_0 \|\rho\|_{Lip}^2}{(\mu(\rho) + \frac{1}{2} C_{LS}^2(\mu) \|\rho\|_{Lip}^2)^2}} \right). \quad (5.17)$$



#### 5.4. The log-concave case.

The log-concave case corresponds to  $\rho_0 = 0$ . It was shown by Bobkov ([10]) that such measures satisfy a Poincaré inequality. A simple proof was given in [6].

In addition, general quantitative bounds are known for  $C_P(\mu)$  in this situation. The easiest one, mentioned in [1, p.11] but credited to [30] reads as follows

**Proposition 5.5.** *Let  $\mu$  be a log-concave distribution of probability. Define  $\text{Var}_\mu(x) = \mu(|x - \mu(x)|^2)$  and  $\lambda^*$  as the largest eigenvalue of the covariance matrix of  $\mu$ . Then*

$$C_P(\mu) \leq 4 \text{Var}_\mu(x) \leq 4n\lambda^* \quad (5.18)$$

It is conjectured (Kannan-Lovasz-Simonovits or K-L-S conjecture) that

$$C_P(\mu) \leq C\lambda^*$$

for some universal constant  $C$  which is thus dimension free. One can look at the original paper [30] and at the monograph [1] for an almost up to date state of the art. The converse inequality  $\lambda^* \leq C_P(\mu)$  is always true. The best known result in this direction is due to the work of Lee and Vempala [34]: for some universal constant  $C$ ,

$$C_P(\mu) \leq Cn^{\frac{1}{2}}\lambda^*. \quad (5.19)$$

Except in the gaussian case where all quantities coincide, the relationship between the curvature and  $\lambda^*$  is quite unclear, except that, according to the Bakry-Emery criterion  $\lambda^* \leq 1/\rho_0$ .

Notice that we may assume that  $0 = \rho_0 = \inf \rho$ . Indeed if  $\inf \rho > 0$  we may always find a smaller modification on a set of small measure with an infimum equal to 0 and then pass to the limit w.r.t. the measure of this set. Hence we may assume that  $\text{Osc}(\rho) = \sup \rho$ .

Similarly we may replace  $\rho$  by  $\rho_K = \min(\rho, K)$ , satisfying again  $\rho_0 = 0$  and for a  $K$  to be chosen later such that  $\mu(\rho_K) > 0$ . Of course  $\mu(\rho_K) \leq \mu(\rho)$ . In order to get a lower bound for  $\mu(\rho_K)$ , it is enough to write

$$\mu(\rho_K) \geq K \mu(\rho \geq K) \geq \frac{1}{2} \text{med}_\mu(\rho)$$

if we choose  $K = \text{med}_\mu(\rho)$ . Of course the choice of a median is artificial and we could replace it by any quantile, up to a modification of the constant pre-factor.

The other specific feature is that, still in the log-concave case, we can replace  $\rho$  by  $\lambda\rho$  for any  $0 < \lambda \leq 1$ . In particular we may choose  $\lambda$  in (5.3) in such a way that

$$\mu(\rho) - \frac{C\lambda}{4} \|\rho\|_c^2 \geq \frac{\mu(\rho)}{2}.$$

It follows

**Theorem 5.6.** *Assume that the assumptions of Theorem 4.9 are satisfied. Assume in addition that  $\rho_0 = 0$ .*

*Then, for all  $\nu$ , all  $t > \eta > 0$ ,*

$$W_1(P_t^* \nu, \mu) \leq 2 \left( \sup_x r(2\eta, x, x) \right)^{\frac{1}{2}} e^{\theta\eta} e^{-\theta t} W_1(\nu, \mu),$$

*where  $\theta$  is given by one of the following expressions:*

$$(1) \quad \theta = \min \left( \frac{\mu(\rho)}{2}; \frac{\mu^2(\rho)}{\|\rho\|_{Lip}^2 C_{LS}^2(\mu)} \right),$$

$$(2) \quad \theta = \min \left( \frac{\mu(\rho)}{2}; \frac{\mu^2(\rho)}{\text{Osc}^2(\rho) C_P(\mu)} \right).$$

In particular we may always take  $\theta = \frac{1}{4} \min \left( \text{med}_\mu(\rho); \frac{1}{C_P(\mu)} \right)$ .

*Remark 5.7.* If we assume that the semi-group is ultra-bounded, then it is hyper-bounded so that  $\mu$  satisfies a defective log-Sobolev inequality  $\text{Ent}_\mu(f^2) \leq C \mu(|\nabla f|^2) + A \mu(f^2)$ . According to the Rothaus lemma, since a log-concave  $\mu$  always satisfies a Poincaré inequality, the previous defective log-Sobolev inequality can be tightened into a true log-Sobolev inequality with a log-Sobolev constant  $C_{LS}(\mu) \leq (C + (A + 2)C_P(\mu))$ . The difference with the previous subsection is that we do not have to assume that this constant is small enough. In addition the value of the constant  $C$  in the defective LSI approach depends on the  $L^\infty$  norm of the density of the measure itself.

Notice that the Theorem applies for instance with  $V(x) = |x|^\beta$  with  $\beta > 2$ .

It is tempting to use the same normalization trick in (5.4) with  $C = C_P$  in order to get a self improving inequality. Unfortunately this yields a trivial  $\text{Osc}^2(\rho)/\mu^2(\rho) \geq 1$ , i.e. (5.4) is not self improving. The same disappointing fact, somewhat more tricky to check, holds with (5.9).  $\diamond$

*Remark 5.8.* In the log-concave situation another approach is possible for the Poincaré constant, using the famous Brascamp-Lieb inequality

$$\text{Var}_\mu(f) \leq \int \langle \nabla f, (\text{Hess}V)^{-1} \nabla f \rangle d\mu. \quad (5.20)$$

In particular it holds

$$\text{Var}_\mu(f) \leq \left( \int \sup_{u \in \mathbb{S}^{n-1}} \langle u, (\text{Hess}V)^{-1} u \rangle d\mu \right) \|\nabla f\|_\infty^2. \quad (5.21)$$

Using E. Milman's result (see [39, Theorem 2.4] or [19, Theorem 2.7 and Theorem 2.14] for extensions) we deduce that there exists an universal constant  $C$  (see [19] for explicit bounds) such that

$$C_P(\mu) \leq C \left( \int \sup_{u \in \mathbb{S}^{n-1}} \langle u, (\text{Hess}V)^{-1} u \rangle d\mu \right) \leq C \mu(1/\rho). \quad (5.22)$$

We thus have a control of the Poincaré constant using some *mean* of the “inverse” curvature, which recovers, up to the value of the constant prefactor, Veysseire's result [44].  $\diamond$

## 6. COMPARING WITH A DIRECT APPROACH IN THE ULTRA-BOUNDED CASE.

In many of the previous results the quantity

$$\sup_x r(2\eta, x, x)$$

appears in the pre-factor for the  $W_1$  convergence. As we said, finiteness of the above quantity is ensured when the semi-group is ultra bounded, i.e. for all  $s > 0$ ,

$$\sup_{x,y} r(s, x, y) \leq K(s) < +\infty.$$

Actually, we do not know examples where the semi-group is not ultra bounded, but  $r(2\eta, x, x)$  is bounded in  $x$ . In the ultra-bounded situation, one may directly control the  $W_1$  decay and see that the rate of the exponential decay obtained in the previous two subsections is not optimal.

First recall that any Boltzmann probability measure  $e^{-V(x)} dx$  with  $V$  locally bounded satisfies a weak Poincaré inequality (see [43, 15]). If the semi-group is ultra-bounded it is hyper-bounded and together with the weak Poincaré inequality we get that a (true) Poincaré inequality is satisfied (see [15, Proposition 5.13]), hence according to the Rothaus lemma  $\mu$  satisfies a logarithmic Sobolev inequality. In our approach here, we shall only use the Poincaré inequality.

Recall that

$$W_1(P_t^* \nu, \mu) = \sup_{\|f\|_{Lip}=1} (P_t^* \nu(f) - \mu(f)) .$$

Replacing  $f$  by  $f - f(0)$  does not change the previous expression. Therefore, since we know that

$$P_\eta^* \nu(dy) = \left( \int r(\eta, x, y) \nu(dx) \right) \mu(dy) = r(\eta, \nu, y) \mu(dy)$$

and since  $P_s$  is symmetric, using the Lipschitz bound on  $f$  we get

$$W_1(P_t^* \nu, \mu) \leq \int |x| |P_{t-\eta} r(\eta, \nu, \cdot) - 1| \mu(dx) . \quad (6.1)$$

If  $r(\eta, \nu, \cdot) \in \mathbb{L}^2(\mu)$ , using first Cauchy-Schwarz inequality and then the Poincaré inequality we get

$$\begin{aligned} W_1(P_t^* \nu, \mu) &\leq \mu^{\frac{1}{2}}(|x|^2) e^{-\frac{t-\eta}{C_P(\mu)}} \mu^{\frac{1}{2}}(|r(\eta, \nu, \cdot) - 1|^2) \\ &\leq \mu^{\frac{1}{2}}(|x|^2) e^{-\frac{t-\eta}{C_P(\mu)}} \left( \int r(2\eta, x, x) \nu(dx) - 1 \right)^{\frac{1}{2}} . \end{aligned} \quad (6.2)$$

It follows that the rate of exponential convergence is  $\theta = 1/C_P(\mu)$ , whatever the curvature is. We are thus gaining a factor 4 compared with Theorem 5.6, but we are losing the pre-factor  $W_1(\nu, \mu)$  in this direct approach.

If in addition  $r(s, \cdot, \cdot)$  is bounded,

$$\left( \int r(2\eta, x, x) \nu(dx) - 1 \right)^{\frac{1}{2}} = \mu^{\frac{1}{2}}(|r(\eta, \nu, \cdot) - 1|^2) \leq (K(\eta) - 1)^{\frac{1}{2}} \|\nu - \mu\|_{TV}^{\frac{1}{2}} .$$

The following can thus be favorably compared with our previous results in some cases

**Proposition 6.1.** *If the semi-group is ultra-bounded, then for all  $\eta > 0$ , denoting  $K(\eta) = \sup_{x,y} r(\eta, x, y)$  we have*

$$W_1(P_t^* \nu, \mu) \leq (K(\eta) - 1)^{\frac{1}{2}} \mu^{\frac{1}{2}}(|x|^2) e^{\eta/C_P(\mu)} \|\nu - \mu\|_{TV}^{\frac{1}{2}} e^{-t/C_P(\mu)} .$$

**Acknowledgments:** This work was supported by Project EFI (ANR-17-CE40-0030) of the French National Research Agency (ANR). PC and MF were supported by ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02. MF was also supported by ANR Project MESA (ANR-18-CE40-006) and by the Partenariat Hubert Curien Procopé. We thank Karl-Theodor Sturm for his lectures in Toulouse in the Spring 2019, during the semester

program *Calculus of Variations and Probability* of the Labex CIMI. We also thank Matthias Erbar, Benjamin Jourdain, Georg Menz, Julien Reygner and Gabriel Stoltz for valuable discussions.

## REFERENCES

- [1] D. Alonso-Gutierrez and J. Bastero. *Approaching the Kannan-Lovasz-Simonovits and variance conjectures*, volume 2131 of *LNLM*. Springer, 2015.
- [2] 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 Mathématique de France, Paris, 2000.
- [3] M. Arnaudon, M. Bonnefont, and A. Joulin. Intertwinings and generalized Brascamp-Lieb inequalities. *Rev. Mat. Iberoam.*, 34(3):1021–1054, 2018.
- [4] M. Arnaudon, A. Thalmaier, and F. Y. Wang. Equivalent Harnack and gradient inequalities for pointwise curvature lower bound. *Bull. Sci. Math.*, 138:643–655, 2014.
- [5] R. Assaraf, B. Jourdain, T. Lelièvre, and R. Roux. Computation of sensitivities for the invariant measure of a parameter dependent diffusion. *Stoch. Partial Differ. Equ. Anal. Comput.*, 6(2):125–183, 2018.
- [6] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures. *Elec. Comm. in Prob.*, 13:60–66, 2008.
- [7] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov diffusion operators.*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2014.
- [8] D. Bakry, I. Gentil, and M. Ledoux. On Harnack inequalities and optimal transportation. *Ann. Scuola Norm.-Sci.*, 14:705–727, 2015.
- [9] F. Barthe and E. Milman. Transference principles for log-sobolev and spectral-gap with applications to conservative spin systems. *Comm. Math. Physics*, 323:575–625, 2013.
- [10] S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4):1903–1921, 1999.
- [11] 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.
- [12] M. Bonnefont and A. Joulin. Intertwining relations for one-dimensional diffusions and application to functional inequalities. *Potential Analysis*, 41:1005–1031, 2014.
- [13] M. Bonnefont, A. Joulin, and Y. Ma. Spectral gap for spherically symmetric log-concave probability measures, and beyond. *J. Funct. Anal.*, 270(7):2456–2482, 2016.
- [14] M. Braun, K. Habermann, and K. T. Sturm. Optimal transport, gradient estimates, and pathwise Brownian coupling on spaces with variable Ricci bounds. Available on Math. ArXiv. 1906.09186 [math.FA], 2019.
- [15] P. Cattiaux. A pathwise approach of some classical inequalities. *Potential. Anal.*, 20:361–394, 2004.
- [16] P. Cattiaux. Hypercontractivity for perturbed diffusion semigroups. *Ann. Fac. Sciences Toulouse Math. (6)*, 14(4):609–628, 2005.
- [17] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of Markov processes. *ESAIM Probab. Stat.*, 12:12–29, 2008.
- [18] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 231–292. Springer, Cham, 2014.
- [19] P. Cattiaux and A. Guillin. On the Poincaré constant of log-concave measures. To appear in *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA)*, 2019.
- [20] P. Cattiaux, A. Guillin, and P. A. Zitt. Poincaré inequalities and hitting times. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 49(1):95–118, 2013.
- [21] X. Cheng and D. Zhou. Eigenvalues of the drifted Laplacian on complete metric measure spaces. *Comm. Contemp. Math.*, 17, 2017.
- [22] T.A. Courtade and M. Fathi. Stability in the Bakry-Emery theorem on  $\mathbb{R}^n$ . *Arxiv preprint*, 2018.
- [23] G. De Philippis and A. Figalli. Rigidity and stability of Caffarelli’s log-concave perturbation theorem. *Nonlinear Anal.*, 154:59–70, 2017.
- [24] G. Ferré, M. Rousset, and G. Stoltz. More on the long time stability of Feynman-Kac semigroups. *arxiv preprint*, 2018.

- [25] N. Gigli, C. Ketterer, K. Kuwada, and S. Ohta. Rigidity for the spectral gap on  $RCD(K, \infty)$ -spaces. *to appear in Amer. J. Math.*, 2018.
- [26] N. Gozlan, C. Roberto, and P.-M. Samson. From concentration to logarithmic Sobolev and Poincaré inequalities. *J. Funct. Anal.*, 260:1491–1522, 2013.
- [27] A. Guillin, C. Léonard, L. Wu, and N. Yao. Transportation-information inequalities for Markov processes. *Probab. Theory Related Fields*, 144(3-4):669–695, 2009.
- [28] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981.
- [29] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38(6):2418–2442, 2010.
- [30] R. Kannan, L. Lovasz, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13(3-4):541–559, 1995.
- [31] I. Kontoyiannis and S. P. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10:no. 3, 61–123, 2005.
- [32] H. Kunita. Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d’été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 143–303. Springer, Berlin, 1984.
- [33] Kazumasa Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11):3758–3774, 2010.
- [34] Y. T. Lee and S. S. Vempala. Eldan’s stochastic localization and the KLS hyperplane conjecture: An improved lower bound for expansion. In Proc. of IEEE FOCS, 2017., 2017.
- [35] C. Léonard. Girsanov theory under a finite entropy condition. *Séminaire de probabilités de Strasbourg, XLIV*(Lecture Notes in Mathematics 2046):429–465, 2012.
- [36] P. Lezaud. Chernoff and Berry-Essen inequalities for Markov processes. *ESAIM Probability and Statistics*, 5:183–201, 2001.
- [37] F. Malrieu and D. Talay. Concentration inequalities for Euler schemes. In *Monte Carlo and quasi-Monte Carlo methods 2004*, pages 355–371. Springer, Berlin, 2006.
- [38] A. Millet, D. Nualart, and M. Sanz. Integration by parts and time reversal for diffusion processes. *Ann. Probab.*, 17(1):208–238, 1989.
- [39] E. Milman. On the role of convexity in isoperimetry, spectral-gap and concentration. *Invent. math.*, 177:1–43, 2009.
- [40] E. Milman. Isoperimetric and concentration inequalities: Equivalence under curvature lower bound. *Duke Math. J.*, 154(2):207–239, 2010.
- [41] S. Ohta and A. Takatsu. Equality in the logarithmic Sobolev inequality. *to appear in Manuscripta Math.*, 2019.
- [42] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173:361–400, 2000.
- [43] M. Röckner and F. Y. Wang. Weak Poincaré inequalities and  $\mathbb{L}^2$ -convergence rates of Markov semigroups. *J. Funct. Anal.*, 185(2):564–603, 2001.
- [44] L. Veysseire. Improved spectral gap bounds on positively curved manifolds. Available on Math. ArXiv. 1105.6080 [math.PR], 2011.
- [45] M. K. von Renesse and K. T. Sturm. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Comm. Pure Appl. Math.*, 68:923–940, 2005.
- [46] F.-Y. Wang. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields*, 109(3):417–424, 1997.
- [47] F. Y. Wang. *Functional inequalities, Markov processes and Spectral theory*. Science Press, Beijing, 2005.
- [48] L. Wu. A deviation inequality for non-reversible Markov process. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 36(4):435–445, 2000.

**Patrick CATTIAUX**, INSTITUT DE MATHÉMATIQUES DE TOULOUSE. CNRS UMR 5219., UNIVERSITÉ PAUL SABATIER,, 118 ROUTE DE NARBONNE, F-31062 TOULOUSE CEDEX 09.

*E-mail address:* `patrick.cattiaux@math.univ-toulouse.fr`

**Max FATHI**, CNRS AND INSTITUT DE MATHÉMATIQUES DE TOULOUSE. CNRS UMR 5219., UNIVERSITÉ PAUL SABATIER,, 118 ROUTE DE NARBONNE, F-31062 TOULOUSE CEDEX 09.

*E-mail address:* `max.fathi@math.univ-toulouse.fr`

**Arnaud GUILLIN**, LABORATOIRE DE MATHÉMATIQUES, CNRS UMR 6620, UNIVERSITÉ CLERMONT-AUVERGNE, AVENUE DES LANDAIS, F-63177 AUBIÈRE.

*E-mail address:* `arnaud.guillin@uca.fr`