Semi Log-Concave Markov Diffusions.

P. Cattiaux and A. Guillin

Abstract

In this paper we intend to give a comprehensive approach of functional inequalities for diffusion processes under various "curvature" assumptions. One of them coincides with the usual Γ_2 curvature of Bakry and Emery in the case of a (reversible) drifted Brownian motion, but differs for more general diffusion processes. Our approach using simple coupling arguments together with classical stochastic tools, allows us to obtain new results, to recover and to extend already known results, giving in many situations explicit (though non optimal) bounds. In particular, we show new results for gradient/semigroup commutation in the log concave case. Some new convergence to equilibrium in the granular media equation is also exhibited.

Key words : Functional inequalities, transport inequalities, diffusion processes, coupling, convergence to equilibrium.

MSC 2000 : 26D10, 35K55, 39B62, 47D07, 60J60.

1 Introduction and main results.

In this paper we shall investigate some properties of time marginals (at time T finite or infinite) of Markov diffusion processes satisfying some logarithmic

P. Cattiaux
Institut de Mathématiques de Toulouse. Université de Toulouse. CNRS UMR 5219, 118 route de Narbonne, F-31062 Toulouse cedex 09 e-mail: cattiaux@math.univ-toulouse.fr

A. Guillin
Laboratoire de Mathématiques, CNRS UMR 6620, Université Blaise Pascal, avenue des Landais, F-63177 Aubière. e-mail: guillin@math.univ-bpclermont.fr

semi-convexity like property. The properties we are interested in are functional inequalities (Poincaré, log-Sobolev) or transportation inequalities. We shall also give some consequences for the long time behavior of such processes.

Our main tools are on one hand coupling techniques and on the other hand stochastic calculus. We shall mainly use the so called "synchronous" coupling, i.e. using the same Brownian motion, but we also give some new results by using the "mirror" coupling (or coupling by reflection) introduced by Lindvall and Rogers in [35]. The main stochastic tool is (a very simple form of) Girsanov theory and h-processes.

The use of coupling techniques for obtaining analytic estimates is far to be new. It is impossible (and dangerous) to give here, even an account of the existing literature (see however [45] and references therein). The use of Girsanov theory for this goal is not new too. We shall recall later some references. The conjunction of both techniques is not usual.

The meaning of logarithmic semi-convexity will generalize the "usual" one we recall now.

Let U be a smooth (C^∞) potential defined on $\mathbb{R}^n$ and satisfying for some $K \in \mathbb{R}$,

$$\textbf{(H.C.K)} \quad \text{for all } (x,y), \quad \langle \nabla U(x) - \nabla U(y), x-y \rangle \; \geq \; K \, |x-y|^2 \, .$$

This property is called K-semi-convexity of U. It is clearly equivalent to the convexity of $U(x) - K|x|^2$. We denote $\Upsilon(dx) = e^{-U(x)} dx$ the Boltzmann measure associated to the potential U. If e^{-U} is dx integrable, we also introduce the normalized $\mu(dx) = \frac{1}{Z_U} \, e^{-U(x)} \, dx$ which is a probability measure. If U is semi-convex, μ is said to be semi log-concave.

Consider first the diffusion process, given by the solution of the Ito stochastic differential system

$$\begin{aligned} dX_t &= dB_t - \frac{1}{2} \, \nabla U(X_t) \, dt \, ; \\ \mathcal{L}(X_0) &= \mu_0 \, . \end{aligned} \tag{1}$$

$B_.$ being a standard brownian motion. It is known that (1) has an unique non explosive strong solution, in particular on can build a solution on any probability space equipped with some brownian motion. This is an easy consequence of Hasminski's explosion test using the Lyapunov function $x \mapsto |x|^2$. Usual notations are in force: for a nice enough f, $P_t f(x) = \mathbb{E}(f(X_t^x))$ where $X_.^x$ denotes a solution such that $\mu_0 = \delta_x$; L denotes the infinitesimal generator i.e.

$$L = \frac{1}{2} \, \Delta \, - \, \frac{1}{2} \, \langle \nabla U, \nabla \rangle \, ,$$

and Γ denotes the carré du champ, namely here

$$\Gamma(f,g) = \frac{1}{2}\,\langle \nabla f\,,\ \nabla g\rangle \text{ and for simplicity } \Gamma(f) = \Gamma(f,f)\,.$$

$\mathbb{P}_{\mu_0}$ will denote the law of the solution of (1), abridged in $\mathbb{P}_x$ when $\mu_0 = \delta_x$, i.e. $\mathbb{P}$ is defined on the usual space Ω of continuous paths; μ_t will denote the law of X_t for $t \geq 0$ and $P(t,x,.)$ denotes the law of X_t^x.

It is known that Υ is a symmetric (reversible) measure for the diffusion process, and is actually the unique invariant (stationary) measure for the process. If Υ is bounded, μ is ergodic.

In the latter case, P_t is thus a symmetric semi-group on $L^2(\mu)$. The domain $D(L)$ of its generator contains the algebra $\mathcal{A}$ generated by the constant functions and C_c^∞. In particular, if $f \in \mathcal{A}$, $\partial_t P_t f = L P_t f = P_t L f$ in $L^2(\mu)$, so that since $\partial_t - L$ is hypo-elliptic $(t,x) \mapsto P_t f(x) \in C^\infty$.

L is the basic example of generator satisfying the celebrated $C(K/2,+\infty)$ Bakry-Emery curvature condition (see [1]). Indeed if we define

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

(H.C.K) is equivalent to $\Gamma_2(f) \geq (K/2)\,\Gamma(f)$.

This curvature condition is known to imply (and is in fact equivalent to) a lot of nice inequalities for the semi-group, in particular for all $T>0$ and all x, a commutation between Γ and the semi group P_t holds, namely

$$\Gamma(P_T f) \leq e^{-KT} P_T \left(\sqrt{\Gamma(f)}\right)^2\,, \tag{2}$$

which in turn implies powerful functional inequalities such as

$$P(T,x,.) \text{ satisfies a log-Sobolev inequality with constant } \frac{4}{K}\,(1-e^{-KT}). \tag{3}$$

Recall that ν satisfies a (usual) log-Sobolev inequality with constant C_{LS} if

$$\mathrm{Ent}_\nu(f) := \int f^2\,\log(f^2)\,d\nu - \left(\int f^2 d\nu\right)\,\log\left(\int f^2 d\nu\right)\,\leq\,C_{LS}\,\int \Gamma(f)\,d\nu\,. \tag{4}$$

(3) is exactly what is (a little bit improperly) called a “local” log-Sobolev inequality in [1] (theorem 5.4.7). For further informations and more, see the forthcoming book [5].

It is well known that a log-Sobolev inequality implies a (usual) Poincaré inequality

$$\mathrm{Var}_\nu(f) := \int f^2\,d\nu\,-\,\left(\int f d\nu\right)^2\,\leq\,C_P\,\int \Gamma(f)\,d\nu\,, \tag{5}$$

with $C_P = \frac{1}{2}\,C_{LS}$, as well as a T_2 transportation inequality

$$W_2^2(\eta,\nu) \, \leq \, C_W \, H(\eta|\nu) \, , \tag{6}$$

with $C_W = C_{LS}$. Here W_2 denotes the Wasserstein distance between the probability measures η and ν, i.e.

$$W_2^2(\eta,\nu) = \frac{1}{2} \, \inf_{\pi} \, \int \, |x-y|^2 \, \pi(dx,dy) \, ,$$

where π is a coupling of η and ν (i.e. has respective marginals equal to η and ν) and

$$H(\eta|\nu) \, = \, \int \left(\frac{d\eta}{d\nu}\right) \, \log \, \left(\frac{d\eta}{d\nu}\right) d\nu \, ,$$

denotes the Kullback-Leibler information or relative entropy of η w.r.t. ν. The latter property is due to Otto-Villani [37]. Another approach and related properties were developed by Bobkov, Gentil and Ledoux [9]. For a nice survey on transportation inequalities we refer to [26]. One can find in all these references another remarkable consequence of semi log-concavity, namely that a log-Sobolev inequality derives from a transportation inequality. This is a consequence of the following (H.W.I) inequality that holds for any nice μ density of probability h,

(H.W.I) If (H.C.K) holds then

$$H(h\mu|\mu) \leq \left(2 \int \, \frac{|\nabla h|^2}{h} \, d\mu\right)^{\frac{1}{2}} \, W_2(h\mu,\mu) - K \, W_2^2(h\mu,\mu) \, .$$

As a consequence, if (H.C.K) holds for some $K \, \leq \, 0$, a T_2 transportation inequality for μ implies a log-Sobolev inequality with constant $C_{LS} \leq (4/C_W) \, (1+(K/C_W))^{-2}$ provided $1+(K/C_W)>0$, in particular if $K=0$. Let us finally remark that the starting point of this approach is the Γ_2 commutation property (2) which fails however to give a direct proof of the T_2 inequality.

Our first goal is to show that functional and transportation inequalities can be derived, in the previous situation, by using coupling techniques and simple tools of stochastic calculus.

Proving (2). As a warming up, let us see how (2) can be easily derived, just using synchronous coupling, i.e. the processes $X_.^x$ and $X_.^y$ built with *the same* Brownian motion.

Applying Ito formula yields (almost surely)

$$\begin{aligned} & e^{Kt} \, |X_t^x - X_t^y|^2 \\ & = |x-y|^2 + \int_0^t \, \left(K|X_s^x - X_s^y|^2 - \langle \nabla U(X_s^x) - \nabla U(X_s^y), X_s^x - X_s^y \rangle\right) \, e^{Ks} \, ds \\ & \leq |x-y|^2 \, , \end{aligned}$$

thanks to (H.C.K). Hence,

$$|X_t^x - X_t^y| \leq e^{-Kt/2}\,|x-y|\,, \quad \mathbb{P}\ a.s. \tag{7}$$

so that, using the mean value theorem,

$$|P_t f(x) - P_t f(y)| \leq \mathbb{E}(|f(X_t^x) - f(X_t^y)|) \leq e^{-Kt/2}\,\mathbb{E}(|\nabla f(z_t)|\,|x-y|)$$

for some z_t sandwiched by X_t^x and X_t^y. It remains to use the continuity (and boundedness) of ∇f and the fact that X_t^y goes almost surely to X_t^x as $y \to x$ to conclude. ♢

Now consider a classical diffusion process, given by the solution of an Ito stochastic differential system

$$\begin{aligned} dX_t &= \sigma(X_t)\,dB_t + b(X_t)\,dt\,; \\ \mathcal{L}(X_0) &= \mu_0, \end{aligned} \tag{8}$$

$B_.$ being a standard brownian motion. For simplicity we assume that σ is a squared matrix. Assumptions will be made ensuring again strong uniqueness and non explosion. The notations used previously are still in force.

Generalizing (2). Let us mimic what we did previously, using again the synchronous coupling and applying Ito formula. We get

$$\begin{aligned} e^{Kt}\,&|X_t^x - X_t^y|^2 \\ &= |x-y|^2 + \int_0^t 2e^{Ks}\langle \sigma(X_s^x) - \sigma(X_s^y), X_s^x - X_s^y\rangle\,dB_s \\ &\quad + \int_0^t (K|X_s^x - X_s^y|^2 + |\sigma(X_s^x) - \sigma(X_s^y)|_{HS}^2 \\ &\qquad + 2\,\langle b(X_s^x) - b(X_s^y), X_s^x - X_s^y\rangle)\,e^{Ks}\,ds\,. \end{aligned}$$

This suggests to extend (H.C.K), say for Higher Convexity of order K, to this new situation

$$\textbf{(H.C.K)} \quad \forall (x,y),\ |\sigma(x) - \sigma(y)|_{HS}^2 + 2\,\langle b(x) - b(y), x-y\rangle \ \leq\ -K\,|x-y|^2\,. \tag{9}$$

Indeed, if (H.C.K) holds

$$e^{Kt}\,|X_t^x - X_t^y|^2 \leq |x-y|^2 + \int_0^t 2e^{Ks}\langle \sigma(X_s^x) - \sigma(X_s^y), X_s^x - X_s^y\rangle\,dB_s\,. \tag{10}$$

If the right hand side of (10) is a (true) martingale, we obtain

$$\mathbb{E}(|X_t^x - X_t^y|^2) \leq e^{-Kt}\,|x-y|^2\,. \tag{11}$$

Let $f \in \mathcal{A}$, then

$$f(X_t^x) - f(X_t^y) \leq \langle \nabla f(X_t^y), X_t^x - X_t^y \rangle + C|X_t^x - X_t^y|^2$$

for some constant C, so that

$$\begin{aligned} |P_t f(x) - P_t f(y)| &= |\mathbb{E}(f(X_t^x) - f(X_t^y))| \\ &\leq |\mathbb{E}(\langle \nabla f(X_t^y), X_t^x - X_t^y \rangle)| + C\mathbb{E}(|X_t^x - X_t^y|^2) \\ &\leq \left(\mathbb{E}(|\nabla f(X_t^y)|^2)\right)^{\frac{1}{2}} \left(\mathbb{E}(|X_t^x - X_t^y|^2)\right)^{\frac{1}{2}} + C\mathbb{E}(|X_t^x - X_t^y|^2) \\ &\leq e^{-Kt/2} \left(P_t(|\nabla f|^2)(y)\right)^{\frac{1}{2}} |x-y| + C\, e^{-Kt}\, |x-y|^2 \,, \end{aligned}$$

so that, provided we know that $\nabla P_t f$ exists, we have obtained a weaker form of (2),

$$|\nabla P_t f|^2 \leq e^{-Kt}\, P_t(|\nabla f|^2)\,. \tag{12}$$

We shall now explain that (12) is in general very different from (2). $\diamondsuit$

Notice that if σ and b are C-Lipschitz, (H.C.K) is satisfied for $K = -(C^2n^2 + C)$, but if σ is C-Lipschitz, (H.C.K) can be satisfied for a non-negative K provided b is sufficiently repealing. Contrary to the case of a constant diffusion coefficient, (H.C.K) is not related to the Bakry-Emery curvature condition which involves in this situation controls on derivatives of higher order of the coefficients.

For simplicity in the sequel we shall assume that $\sigma \in C_b^3$ hence is C-Lipschitz and that b is C^3, but not necessarily bounded nor with bounded derivatives. With these assumptions, once again if we assume that (H.C.K) is in force, (8) admits a unique non explosive strong solution using $x \mapsto x^2$ as a Lyapunov function for non explosion. We shall show this and other properties of the process in subsection 3.1.

We still use the notations introduced before, but now

$$L = \frac{1}{2}\, \sigma\, \sigma^*\, \nabla^2 \,+\, b\nabla \,=\, \frac{1}{2}\, a\, \nabla^2 \,+\, b\nabla\,,$$

and Γ the carré du champ is now

$$\Gamma(f,g) = \frac{1}{2}\, \langle \sigma\, \nabla f \,,\, \sigma\, \nabla g \rangle\,.$$

Notice that, contrary to the Bakry-Emery bounded curvature case, the previous commutation property (12) holds with the *usual gradient* and not with the *natural* one i.e. $\Gamma^{\frac{1}{2}}$.

If the proof we gave of (2) is more or less part of the folklore, the previous proof of (12) already appeared in a slightly different form in Lemma 2.2 of [41].

Starting from (2), the Γ_2 calculus of Bakry-Emery uses this commutation property and the control of the derivative of $\psi(s) = P_s(P_{t-s}f \log(P_{t-s}f))$

to get a "local" logarithmic Sobolev inequality, i.e. a logarithmic Sobolev inequality satisfied by $P(T, x, .)$ for all finite T. Note that considering rather $\psi(s) = P_s((P_{t-s}f)^2)$ leads to a local Poincaré inequality. A similar calculation can be done, at least in the uniformly elliptic situation, in the general case starting from (2), to again prove a Poincaré inequality (see e.g. [46]).

In the next two sections we propose an alternate method based on the use of h-processes. This idea is close to the one used in [14] in the stationary case, and in the paper by Djellout, Guillin and Wu [22], Theorem 5.6 (condition 4.5 therein being exactly (H.C.K) for $K > 0$) where these authors are looking at transportation inequalities on the path space. What we shall show is that the same scheme of proof also furnishes functional inequalities. This unified treatment of functional inequalities and transportation inequalities using an ad-hoc coupling is the novelty here.

Another interest is that the method does not require uniform ellipticity of a so that some hypo-coercive examples enter this framework as explained in subsection 3.5. In this situation, Bakry-Emery curvature is actually equal to $-\infty$.

It also extends to the non (time-)homogeneous situation, see section 5. In addition in this section we show how to directly obtain convergence to equilibrium and properties of the invariant measure for non linear diffusions of Mc Kean-Vlasov type, simplifying arguments in [36].

The case of positive curvature ($K > 0$) in (2) is important since it implies exponential decay to the equilibrium μ_∞ (or of $P_t f$ to $\int f \, d\mu_\infty$ in variance or entropy). It also provides the exponential *contraction* in W_2 distance, i.e for all initial μ_0 and ν_0,

$$W_2^2(\mu_t, \nu_t) \leq e^{-Kt} \, W_2^2(\mu_0, \nu_0)$$

as it clearly derives from (7). Actually as shown first by Sturm and Von Renesse [44], this contraction property is equivalent to positive curvature K. A similar statement holds under the general (H.C.K) assumption in the elliptic case (see Theorem 1).

In full generality however, (H.C.K) for a positive K, only implies the following Poincaré inequality for the invariant measure :

$$\mathrm{Var}_{\mu_\infty}(f) \leq \frac{M}{K} \int |\nabla f|^2 \, d\mu_\infty \, , \tag{13}$$

where M denotes the uniform norm of a. This inequality is in general strictly weaker than (5) (recall that $\Gamma(f) = |\sigma \, \nabla(f)|^2$), except in the uniformly elliptic case. In particular it is not sufficient to ensure the exponential decay of $\mathrm{Var}_{\mu_\infty}(P_t f)$ to 0.

If μ_∞ is not only invariant but *symmetric*, it was remarked in [17] Remark 4.9, that an exponential decay of the Wasserstein distance

$$W_2(\mu_t, \mu_\infty) \leq C\, e^{-Kt}\, W_2(\mu_0, \mu_\infty)$$

for some $C \geq 1$, implies that μ_∞ satisfies a T_2 inequality (6), and consequently a Poincaré inequality (5). This property thus holds in positive curvature.

But still in the symmetric case, one can reinforce the previous result:

Proposition 1. *Assume that for all bounded (resp. Lipschitz) density of probability h we have $W_0(P_t h\mu, \mu) \leq c_h(t)$ (reps. W_1). Then for all bounded (resp. Lipschitz and bounded) f, there exist c_f and h such that $Var_\mu(P_t f) \leq c_f\, c_h(2t)$.*
In particular if $c_h(t) = c_h\, e^{-\beta t}$, μ satisfies a Poincaré inequality (5).

The latter statement is a consequence of the following lemma one can find for instance in [19] lemma 2.12:

Lemma 1. *Assume that P_t is μ-symmetric. Then, if there exists $\beta > 0$ such that for all f in a dense subset of $\mathbb{L}^2(\mu)$ there exists c_f with $Var_\mu(P_t f) \leq c_f\, e^{-\beta t}$ then $Var_\mu(P_t f) \leq e^{-\beta t}\, Var_\mu(f)$ for all $f \in \mathbb{L}^2(\mu)$.*
Hence μ satisfies a Poincaré inequality with constant $C_P \leq 1/\beta$.

It is thus interesting to look at other Wasserstein distances, in particular W_1.

$\mathbf{W_1}$ and synchronous coupling. Using again the synchronous coupling and applying Ito formula, we have

$$\begin{aligned} e^{Kt}\, |X_t^x - X_t^y| &= |x-y| + \int_0^t e^{Ks} \left\langle \sigma(X_s^x) - \sigma(X_s^y), \frac{X_s^x - X_s^y}{|X_s^x - X_s^y|} \right\rangle dB_s \\ &+ \int_0^t K|X_s^x - X_s^y| e^{Ks} ds \\ &+ \int_0^t \frac{e^{Ks}}{2|X_s^x - X_s^y|} \left(|\sigma(X_s^x) - \sigma(X_s^y)|_{HS}^2 + 2\, \langle b(X_s^x) - b(X_s^y), X_s^x - X_s^y \rangle \right) ds \\ &- \int_0^t \frac{1}{2|X_s^x - X_s^y|} \left| (\sigma(X_s^x) - \sigma(X_s^y)) \left(\frac{X_s^x - X_s^y}{|X_s^x - X_s^y|} \right) \right|^2 e^{Ks}\, ds\,, \end{aligned}$$

almost surely for $t < T_C$, where T_C denotes the coupling time, i.e. the first time when $X_t^x = X_t^y$. After this time both processes coincide thanks to pathwise uniqueness.

The gain with respect to (H.C.K) replacing $|\sigma(x) - \sigma(y)|_{HS}^2$ by

$$|\sigma(x) - \sigma(y)|_{HS}^2 - \left| (\sigma(x) - \sigma(y)) \left(\frac{x-y}{|x-y|} \right) \right|^2$$

is mainly irrelevant except in the one dimensional case where the latter quantity is equal to 0.

But even in this case an exponential decay of W_1 will furnish some weak commutation property for the gradient and the semi-group namely

$$|\nabla P_t f| \leq C\, e^{-Kt} \ \parallel \nabla f \parallel_\infty \,, \tag{14}$$

which, nevertheless, allows us to derive weak functional inequalities. ♢

Another approach for studying the exponential decay of W_1 was proposed by Eberle [23, 24]. Instead of synchronous coupling it uses the *mirror* (or *reflection*) coupling introduced by Lindvall and Rogers [35], and then extended by Cranston. Eberle's method allows him to look at drifted Brownian motions when the drift satisfies some "uniform convexity at infinity" property, i.e. when (H.C.K) is satisfied for some $K > 0$ but for $|x - y|$ large enough.

This situation cannot be treated by using synchronous coupling.

We recall Eberle's method and obtain some new consequences of his result in section 7. In addition, up to an extra condition, we show that his result (and all the consequences we derived) can be extended to general elliptic diffusion processes. We will also use this mirror coupling to show that we may get a weak version of the commutation property in the log concave case with the "convexity at infinity" property at least in dimension one, which is the first result we know of in this direction. Still in dimension one, we will also consider using mirror coupling for non linear diffusions.

Let us come back to (1). Among potentials U satisfying Eberle's condition, one find those written as $U = V + W$ with V K-uniformly convex and W Lipschitz continuous. That $\mu(dx) = e^{-U(x)}dx$ satisfies a log-Sobolev (and a Poincaré) inequality is already known in this situation, and also when W is bounded, with a constant $C_{LS} = (4/K)\, \exp(\text{Osc}W)$ where Osc denotes the oscillation of W. But, both approaches are "dimensional", i.e. furnish constants which are dimension dependent as for the celebrated double well case $|x|^4 - |x|^2$, but even for convex potentials which are not uniformly convex like $|x|^4$.

In section 6, we introduce the following extension of (H.C.K).

Let α be a non decreasing function defined on $\mathbb{R}^+$. We shall say that **(H.α.K)** is satisfied for some $K > 0$ if for all (x, y) and all $\varepsilon > 0$,

$$\langle \nabla U(x) - \nabla U(y), x - y \rangle \geq K\, \alpha(\varepsilon)\, (|x - y|^2 - \varepsilon)\,.$$

When $\alpha(a) = 1$, we may take $\varepsilon = 0$ and we recognize (H.C.K). In Proposition 15 we show that $U(x) = |x|^{2\beta}$ (with $\beta \geq 1$) satisfies (H.α.K_β) for $\alpha(a) = a^{\beta-1}$ and an explicit $K_\beta > 0$.

The main result of this section is then that, for suitable functions α,

if (H.α.K) holds (for $K > 0$), then μ satisfies a log-Sobolev inequality.

See theorems 9 and 10. These theorems thus (partly) extend the Bakry-Emery criterion (3) to some non uniformly convex potentials. However, they are dealing with the invariant measure only and not with the law at time T (only incomplete results are proved in this section for these distributions). Finally they provide some explicit bound, but presumably not optimal.

Section 8 is peculiar. Using the results we have described for the Ornstein-Uhlenbeck process we show how to recover known results on the stability of functional inequalities under convolution (provided one of the terms is gaussian).

The use of stochastic calculus in deriving such inequalities is not new but only a small number of papers dealt with. One can trace back to the paper of Borell [13], who used Girsanov theory to study the propagation of log-concavity along the Schrödinger dynamics (not the Fokker-Planck one we are looking at here). In addition to [22] for transportation inequalities, and [41, 46], one can also mention [14, 15] where similar ideas are used to study hyper-boundedness. More recently, using similar arguments, Lehec [34] has studied gaussian functional inequalities and Fontbona and Jourdain [25] obtained a pathwise version of the Γ_2 theory.

Since we will meet several types of inequalities, from now on, in the whole paper, we shall say that ν satisfies a log-Sobolev inequality with constant C_{LS} if

$$\mathrm{Ent}_\nu(f) := \int f^2 \, \log(f^2)\, d\nu - \left(\int f^2 d\nu\right) \, \log\left(\int f^2 d\nu\right) \, \leq \, C_{LS} \, \int |\nabla f|^2 \, d\nu \, . \tag{15}$$

Similarly for the Poincaré inequality, replacing the entropy by the variance.

2 Semi log-concave drifted Brownian motion.

In this first warming up section we shall look at the simplest situation given by (1) and derive the classical inequalities. Though the methods are the same in the general case, we prefer to detail the proofs first in this simpler situation. *We emphasize that the results of this section are very well known. What is new is the method of proof.*

2.1 Commutation property

Recall the result we proved in the introduction using synchronous coupling:

Proposition 2. *In the situation of (1), assume (H.C.K). Then for all $f \in \mathcal{A}$,*

$$\begin{aligned} W_2(P_t(x,\cdot), P_t(y,\cdot)) &\leq e^{-Kt/2} |x-y|, \\ |\nabla P_t f| &\leq e^{-Kt/2} \, P_t |\nabla f| \, . \end{aligned} \tag{16}$$

Remark 1. If instead of (x, y) the processes start with initial distribution π_0 the "optimal coupling" between μ_0 and ν_0 for the W_2 distance, the previous shows that $W_2^2(\mu_T, \nu_T) \leq e^{-KT}\, W_2^2(\mu_0, \nu_0)$. ♢

2.2 h-processes and functional inequalities.

We now introduce the standard notion of h-process. Let $T > 0$ and h be a non-negative function such that $\int h\, d\mu_T = 1$. For simplicity, we assume for the moment that there exist c and C such that $C \geq h \geq c > 0$. We thus may define on the path-space up to time T a new probability measure

$$\frac{d\mathbb{Q}}{d\mathbb{P}_{\mu_0}}|_{\mathcal{F}_T} \, = \, h(\omega_T)\,.$$

It is immediately seen that

$$\mathbb{Q} \circ \omega_s^{-1} = P_{T-s} h\, \mu_s \quad \text{ for all } \; 0 \leq s \leq T\,.$$

In this situation, it is well known (Girsanov transform theory) that one can find a progressively measurable process u_s such that

$$\frac{d\mathbb{Q}}{d\mathbb{P}_{\mu_0}}|_{\mathcal{F}_T} \, = \, P_T h(\omega_0) \, \exp\left(\int_0^T \, \langle u_s, dM_s \rangle \, - \, \frac{1}{2} \int_0^T \, |u_s|^2 \, ds\right) \, ,$$

where ω denotes the canonical element of the path-space and M denotes the martingale part of ω under $\mathbb{P}_{\mu_0}$. In addition, it is easily seen (see e.g. [20]) that

$$H(\mathbb{Q}|\mathbb{P}_{\mu_0}) \, = \, H(h\mu_T|\mu_T) \, = \, H(P_T h \mu_0|\mu_0) + \frac{1}{2}\, \mathbb{E}^{\mathbb{Q}} \left(\int_0^T \, |u_s|^2 \, ds\right) \, . \tag{17}$$

Actually, if $h \in \mathcal{A}$, it is immediate to check (applying Ito formula) that

$$u_s = \nabla \, \log P_{T-s} h(\omega_s)$$

both $\mathbb{P}_{\mu_0}$ and $\mathbb{Q}$ almost surely.

(17) thus becomes

$$H(h\mu_T|\mu_T) \, = \, H(P_T h\mu_0|\mu_0) + \frac{1}{2} \int_0^T \, \left(\int \, \frac{|\nabla P_s h|^2}{P_s h} \, d\mu_{T-s}\right) \, ds\,. \tag{18}$$

If h is smooth we may apply Proposition 2 in order to get

$$\begin{aligned}
H(h\mu_T|\mu_T) &\leq H(P_T h\mu_0|\mu_0) + \frac{1}{2}\int_0^T \left(\int e^{-Ks}\,\frac{P_s^2(|\nabla h|)}{P_s h}\,d\mu_{T-s}\right) ds \\
&\leq H(P_T h\mu_0|\mu_0) + \frac{1}{2}\int_0^T e^{-Ks}\left(\int P_s\left(\frac{|\nabla h|^2}{h}\right) d\mu_{T-s}\right) ds \\
&\leq H(P_T h\mu_0|\mu_0) + \frac{1}{2}\int_0^T e^{-Ks}\left(\int \frac{|\nabla h|^2}{h}\,d\mu_T\right) ds \\
&\leq H(P_T h\mu_0|\mu_0) + \frac{1-e^{-KT}}{2K}\int \frac{|\nabla h|^2}{h}\,d\mu_T\,, \qquad (19)
\end{aligned}$$

where we have used Cauchy-Schwarz inequality for the second inequality and the Markov property for the third one. The previous inequality then extends to any h in C^1 for which the right hand side makes sense, by density.

We have thus obtained the following

Proposition 3. *In the situation of (1), assume (H.C.K). If μ_0 satisfies a log-Sobolev inequality with constant $C_{LS}(0)$, μ_T satisfies a log-Sobolev inequality with constant*

$$C_{LS}(T) = e^{-KT}\,C_{LS}(0) \,+\, \frac{2(1-e^{-KT})}{K}\,.$$

When $K=0$ one has to replace $\frac{(1-e^{-KT})}{K}$ by T. This applies in particular to $\mu_T = P(T,x,.)$ since δ_x satisfies a log-Sobolev inequality with constant equal to 0.

Proof. Apply the log-Sobolev inequality to μ_0. It furnishes (since $\int P_T h d\mu_0 = 1$),

$$H(P_T h\mu_0|\mu_0) \leq \frac{C_{LS}(0)}{4}\int \frac{|\nabla P_T h|^2}{P_T h}\,d\mu_0 \leq e^{-KT}\,\frac{C_{LS}(0)}{4}\int \frac{|\nabla h|^2}{h}\,d\mu_T\,,$$

similarly as what we did in (19). Hence the result applying (19).

As we recalled in the introduction a log-Sobolev inequality implies a T_2 transportation inequality. It is interesting to see that one can directly obtain such an inequality for semi log-concave measures, by using the previous construction. But before to do this, just remark that the above proof using $h = 1+\varepsilon g$ with $\int g d\mu_T = 0$ allows us to obtain a similar result replacing the log-Sobolev inequality by a Poincaré inequality i.e.

Proposition 4. *In the situation of (1), assume (H.C.K). If μ_0 satisfies a Poincaré inequality with constant $C_P(0)$, μ_T satisfies a Poincaré inequality with constant*

$$C_P(T) = e^{-KT}\,C_P(0) \,+\, \frac{1-e^{-KT}}{K}\,.$$

When $K=0$ one has to replace $\frac{(1-e^{-KT})}{K}$ by T. This applies in particular to $\mu_T = P(T,x,.)$ since δ_x satisfies a Poincaré inequality with constant equal to 0.

2.3 Transportation inequalities.

The existence of u_s and (17) are ensured as soon as $H(h\mu_T|\mu_T) < +\infty$ (see [20]). For our goal we do not need the explicit expression of u_s.

Indeed, Girsanov theory and Paul Lévy characterization of Brownian motion tell us that on $(\Omega, \mathbb{Q})$, there exists some standard brownian motion w (independent of ω_0) such that, up to time T,

$$\omega_t = \omega_0 + w_t - \frac{1}{2}\int_0^t \nabla U(\omega_s)\, ds + \int_0^t u_s\, ds\,.$$

Since (1) has an unique *strong* solution, one can build (on $(\Omega, \mathbb{Q})$) a solution of

$$z_t = z_0 + w_t - \frac{1}{2}\int_0^t \nabla U(z_s)\, ds\,,$$

the law of which being given by

$$\mathbb{P}_{\nu_0} \quad \text{with} \quad \nu_0 = \mathcal{L}(z_0)\,.$$

For instance we may choose $\nu_0 = \mu_0$ or $z_0 = \omega_0$ in which case $\nu_0 = P_T h\mu_0$. But in all situations we choose the distribution of (ω_0, z_0) in such a way that $\mathbb{E}^{\mathbb{Q}}(|\omega_0 - z_0|^2) = 2W_2^2(\nu_0, P_T h\mu_0)$ (or we take approximating sequences).

In particular

$$z_t - \omega_t = (z_0 - \omega_0) + \frac{1}{2}\int_0^t (\nabla U(\omega_s) - \nabla U(z_s))\, ds - \int_0^t u_s\, ds\,,$$

$\mathbb{Q}$ almost surely. Applying Ito's formula and (H.C.K) we obtain

$$\begin{aligned} \eta_t &:= \mathbb{E}^{\mathbb{Q}}(|z_t - \omega_t|^2) && (20)\\ &\le \mathbb{E}^{\mathbb{Q}}(|z_0 - \omega_0|^2) - K\int_0^t \eta_s\, ds + 2\int_0^t \mathbb{E}^{\mathbb{Q}}|\langle (z_s - \omega_s), u_s\rangle|\, ds \\ &\le \eta_0 - K\int_0^t \eta_s\, ds + 2\left(\int_0^t \eta_s\, ds\right)^{\frac{1}{2}} \left(\mathbb{E}^{\mathbb{Q}}\left(\int_0^t |u_s|^2\, ds\right)\right)^{\frac{1}{2}} && (21)\\ &\le \eta_0 - K\int_0^t \eta_s\, ds + 2\sqrt{2}\, H^{\frac{1}{2}}(h\mu_T|\mu_T)\left(\int_0^t \eta_s\, ds\right)^{\frac{1}{2}}. \end{aligned}$$

We have then different situations depending on the sign of K.

1) First in the case where $K > 0$, one has using that $2ab \leq Ka^2 + b^2/K$

$$\begin{aligned} \eta_t &\leq \eta_0 \, - \, K \int_0^t \eta_s \, ds + 2\sqrt{2} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \, \left(\int_0^t \eta_s \, ds \right)^{\frac{1}{2}} \\ &\leq \eta_0 + \frac{2}{K} H^{\frac{1}{2}}(h\mu_T|\mu_T) \end{aligned}$$

so that we recover an uniform transportation inequality when $\eta_0 = 0$, which is moreover optimal for the invariant measure, considering logarithmic Sobolev inequality and Poincaré inequality. If μ_0 satisfies some transportation inequality then one obtains that μ_T satisfies a transportation inequality with constant the sum of the initial constant plus $\frac{2}{K}$.

2) The previous simple argument has however a serious drawback in the sense that in positive curvature, μ_T does not forget the "initial measure". Let us see how to deal with this problem. Start once again from the first estimation, but using Itô's formula between t and $t+\varepsilon$ and (H.C.K)

$$\eta_{t+\varepsilon} \leq \eta_t - \, K \int_t^{t+\varepsilon} \eta_s \, ds + \, 2 \int_t^{t+\varepsilon} \mathbb{E}^{\mathbb{Q}} |\langle (z_s - \omega_s), u_s \rangle| \, ds$$

so that we may differentiate in time to get for all positive λ

$$\begin{aligned} \eta_t' &\leq -K\eta_t + 2\mathbb{E}^{\mathbb{Q}} |\langle (z_t - \omega_t), u_t \rangle| \, ds, \\ &\leq -(K+\lambda)\eta_t + \frac{1}{\lambda} \mathbb{E}^{\mathbb{Q}} |u_t|^2. \end{aligned}$$

Using Gronwall's lemma, we get that

$$\eta_T \leq e^{(-K+\lambda)T} \, \eta_0 + \frac{1}{\lambda} \int_0^T e^{(K-\lambda)(s-T)} \mathbb{E}^{\mathbb{Q}} |u_t|^2 \, dt.$$

so that if $K > 0$ we get, for $\lambda < K$

$$\eta_T \leq e^{(-K+\lambda)T} \, \eta_0 + \frac{1}{\lambda} \int_0^T \mathbb{E}^{\mathbb{Q}} |u_t|^2 \, dt \leq e^{(-K+\lambda)T} \, \eta_0 + \frac{2}{\lambda} \, H(h\mu_T|\mu_T).$$

Note that this is once again optimal for the limiting measure, and captures the fact that it forgets the initial condition. When $K < 0$, we then have

$$\eta_T \leq e^{(-K+\lambda)T} \, \eta_0 + \frac{2}{\lambda} e^{(-K+\lambda)T} \, H(h\mu_T|\mu_T).$$

Note however the presence of the additional parameter λ.

3) Let us see how a direct approach may get rid of this additional parameter, which is particularly important in negative curvature. Define

$$a_t = e^{Kt} \int_0^t \eta_s \, ds \, - \, \frac{e^{Kt}}{K} \, \eta_0 \, .$$

We have

$$a'_t \, \leq \, 2\sqrt{2} \, e^{Kt/2} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \, \left(a_t + \frac{e^{Kt}}{K} \, \eta_0 \right)^{\frac{1}{2}} \, .$$

Since $\frac{e^{Kt}}{K} \leq \frac{e^{KT}}{K}$ we obtain

$$\left(a_t + \frac{e^{KT}}{K} \, \eta_0 \right)^{\frac{1}{2}} \leq \left(a_0 + \frac{e^{KT}}{K} \, \eta_0 \right)^{\frac{1}{2}} + 2\sqrt{2} \, \frac{e^{Kt/2} - 1}{K} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \, .$$

It follows

$$\left(\int_0^T \eta_s \, ds \right)^{\frac{1}{2}} \leq \left(\frac{1 - e^{-KT}}{K} \, \eta_0 \right)^{\frac{1}{2}} + 2\sqrt{2} \, \left(\frac{1 - e^{-KT/2}}{K} \right) \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \, .$$

For $K \geq 0$, this yields, since $W_2^2 \, (h\mu_T, \nu_T) \, \leq \, \frac{1}{2} \, \mathbb{E}^{\mathbb{Q}}(|z_t - \omega_t|^2)$, and using $\sqrt{a} \, \sqrt{b} \, \leq \frac{1}{2} \, (a + b)$,

$$\begin{aligned} W_2^2 \, (h\mu_T, \nu_T) \leq & \left(1 + \sqrt{2} \, \frac{1 - e^{-KT}}{K} \right) \, W_2^2 \, (P_T h\mu_0, \nu_0) \\ & + \left(\frac{\sqrt{2}}{2} + \frac{4(1 - e^{-KT/2})}{K} \right) \, H(h\mu_T|\mu_T) \, . \end{aligned} \tag{22}$$

If $\eta_0 = 0$, (22) can be improved in

$$W_2^2 \, (h\mu_T, \nu_T) \, \leq \, \frac{4(1 - e^{-KT/2})}{K} \, H(h\mu_T|\mu_T) \, . \tag{23}$$

When $K \leq 0$, we obtain

$$\begin{aligned} W_2^2 \, (h\mu_T, \nu_T) \leq & \left(1 + \sqrt{2} \, \frac{1 - e^{-KT}}{K} + 2(e^{-KT} - 1) \right) \, W_2^2 \, (P_T h\mu_0, \nu_0) \\ & + \left(\frac{\sqrt{2}}{2} + 4 \, \frac{(1 - e^{-KT/2})}{K} - \, 4 \, \frac{(1 - e^{-KT/2})^2}{K} \right) \, H(h\mu_T|\mu_T) \, . \end{aligned} \tag{24}$$

Again if $\eta_0 = 0$, (24) can be improved in

$$W_2^2 \, (h\mu_T, \nu_T) \, \leq \, 4 \, \left(\frac{(1 - e^{-KT/2})}{K} - \, \frac{(1 - e^{-KT/2})^2}{K} \right) \, H(h\mu_T|\mu_T) \, . \tag{25}$$

The previous inequalities then extend to any non-negative h (not necessarily bounded below nor above).

If we choose $\mu_0 = \delta_x$, we have $\mu_T = P(T, x, .)$, $1 = \int h d\mu_T = P_T h(x)$ and so $\nu_0 = \delta_x$ and $\nu_T = \mu_T$. Hence

Proposition 5. *In the situation of (1), assume (H.C.K). Then $P(T, x, .)$ satisfies a T_2 transportation inequality*

$$W_2^2\left(hP(T,x,.), P(T,x,.)\right) \leq C_T\, H(hP(T,x,.)|P(T,x,.))\,,$$

with

$$C_T = \min\left(\frac{2}{K}, \frac{4(1-e^{-KT/2})}{K}\right)$$

when $K > 0$, $2T$ when $K = 0$ and

$$C_T = \frac{4(1-e^{-KT/2})}{K} - 4\,\frac{(1-e^{-KT/2})^2}{K}$$

when $K \leq 0$.

If we choose $\nu_0 = \mu_0$, we may use the convexity of $t \mapsto t \log t$, i.e

$$H(P_T h\mu_0|\mu_0) = \int P_T h\, \log P_T h\, d\mu_0 \leq \int P_T(h\,\log h)\, d\mu_0 = H(h\mu_T|\mu_T)\,,$$

in order to get

Proposition 6. *In the situation of (1), assume (H.C.K). If μ_0 satisfies T_2 with constant $C(0)$, then μ_T satisfies T_2 with a constant $C(T)$ given,*

1. *when $K > 0$, for $0 < \lambda < K$,*

$$C(T) = e^{-(K-\lambda)T} C(0) + \frac{2}{\lambda}\,,$$

2. *and when $K \leq 0$, $C(T) = C_T + \frac{\sqrt{2}}{2} + B_T\, C(0)$ with*

$$B_T = 1 + \sqrt{2}\,\frac{1-e^{-KT}}{K} + 2(e^{-KT}-1)\,.$$

Remark 2. All what precedes holds even if Υ is not bounded (i.e. if the process is not positive recurrent), in which case of course, $K < 0$. ♢

Remark 3. If we choose $\mu_0 = \mu$ (assuming that Υ is bounded), we have to choose $\nu_0 = P_T h\mu$ hence $\nu_T = P_{2T} h\, \mu_0$. After noticing that we can slightly refine the previous bound replacing $H(h\mu_T|\mu_T)$ by $H(h\mu_T|\mu_T) - H(P_T h\mu_0|\mu_0)$ according to (17), we obtain

$$W_2(P_{2T} h\mu, h\mu) \leq \sqrt{C_T\,(H(h\mu|\mu) - H(P_T h\mu|\mu))}$$

and finally

$$W_2(h\mu,\mu) \le \sqrt{C_T\left(H(h\mu|\mu) - H(P_T h\mu|\mu)\right)} + W_2(P_{2T}h\mu,\mu)\,. \qquad (26)$$

The latter has to be compared with remark 4.9 in [17] which shows that the inequality

$$W_2(h\mu,\mu) \le \sqrt{T\left(H(h\mu|\mu) - H(P_T h\mu|\mu)\right)} + W_2(P_T h\mu,\mu)$$

always holds. ♢

Remark 4. If $K > 0$ we may let T go to $+\infty$ in Proposition 3 and recover that μ satisfies a log-Sobolev inequality with constant $2/K$, hence a T_2 transportation inequality with constant $1/K$ (in particular we are loosing a factor 4 in Proposition 5).
Similarly, when $T \to +\infty$, (26) shows that if $K > 0$, μ satisfies a T_2 inequality , and since μ is log-concave, satisfies a log-Sobolev inequality. This scheme of proof does not require Proposition 2, but the (H.W.I) inequality. Unfortunately it does not furnish the optimal constant. ♢

2.4 Transportation-Fisher Inequalities.

Let us look now at another type of Transportation Information inequality recently introduced in [27], which is weaker but quite close to logarithmic Sobolev inequality (in fact equivalent under bounded curvature). For simplicity we assume here that $K \ge 0$. We are obliged to come back to the initial inequality in (20) which becomes in our new situation

$$\eta_t \le \eta_0 - K\int_0^t \eta_s\,ds + 2\int_0^t \mathbb{E}^{\mathbb{Q}}\left(|z_s - \omega_s|\,|\nabla \log P_{T-s}h(\omega_s)|\right) ds\,. \qquad (27)$$

Replacing the pair $(0,t)$ by $(t,t+\varepsilon)$ we thus have

$$\begin{aligned}\eta_{t+\varepsilon} &\le \eta_t - K\int_t^{t+\varepsilon} \eta_s\,ds + 2\int_t^{t+\varepsilon} \eta_s^{\frac12}\,\left(\mathbb{E}^{\mathbb{Q}}\left(|\nabla \log P_{T-s}h(\omega_s)|^2\right)\right)^{\frac12}\,ds\\ &\le \eta_t - K\int_t^{t+\varepsilon} \eta_s\,ds + 2\int_t^{t+\varepsilon} \eta_s^{\frac12}\,\left(\int \frac{|\nabla P_{T-s}h|^2}{P_{T-s}h}\,d\mu_s\right)^{\frac12}\,ds\,.\end{aligned}$$

It follows that $t \mapsto \eta_t$ is differentiable and satisfies,

$$\begin{aligned}
\eta_t' &\le -K\,\eta_t + 2\,\eta_t^{\frac{1}{2}}\,\left(\int \frac{|\nabla P_{T-t}h|^2}{P_{T-t}h}\,d\mu_t\right)^{\frac{1}{2}} \\
&\le -K\,\eta_t + 2\,\eta_t^{\frac{1}{2}}\,\left(\int \frac{P_{T-t}^2|\nabla h|}{P_{T-t}h}\,d\mu_t\right)^{\frac{1}{2}} \\
&\le -K\,\eta_t + 2\,\eta_t^{\frac{1}{2}}\,\left(\int P_{T-t}\left(\frac{|\nabla h|^2}{h}\right)\,d\mu_t\right)^{\frac{1}{2}} \\
&\le -K\,\eta_t + 2\,\eta_t^{\frac{1}{2}}\,\left(\int \frac{|\nabla h|^2}{h}\,d\mu_T\right)^{\frac{1}{2}}\,. \qquad (28)
\end{aligned}$$

(for the second inequality, recall that (H.C.0) is satisfied so that, for short, $|\nabla P_s| \le P_s|\nabla|$.) To explore (28) we shall use the usual trick $ab \le \lambda a^2 + \frac{1}{\lambda}\,b^2$ for a, b, λ positive. Hence

$$\eta_t' \le \,(-K + 2\lambda)\;\eta_t + \frac{2}{\lambda}\,\left(\int \frac{|\nabla h|^2}{h}\,d\mu_T\right)\,. \qquad (29)$$

We deduce, denoting $A = K\, - 2\lambda$,

$$W_2^2(h\mu_T, \mu_T) \le \eta_T \le \eta_0\,e^{-AT} + \frac{2(1 - e^{-AT})}{A\lambda}\,\int \frac{|\nabla h|^2}{h}\,d\mu_T\,.$$

This inequality is close to what is called a $W_2 I$ inequality (see [26] definition 10.4 or [27] for examples and details on properties of WI inequality). Here we obtain a defective $W_2 I$ inequality. However, as $T \to +\infty$, we recover the true $W_2 I$ inequality for the invariant distribution, which together with the (H.W.I) inequality allows us to recover the log-Sobolev inequality. Nevertheless, we get

Proposition 7. *Assume (H.C.K) for some $K \ge 0$. Then, for all $\lambda < K/2$, $P(T, x, \cdot)$ satisfies a WI inequality of constant $\frac{2(1-e^{-(K-2\lambda)T})}{(K-2\lambda)\lambda}$. If we suppose moreover that μ_0 satisfies a WI inequality with constant $D(0)$ then μ_T satisfies a WI inequality with constant $D(T) = e^{-(K-2\lambda)T}D(0) + \frac{2(1-e^{-(K-2\lambda)T})}{(K-2\lambda)\lambda}$.*

Remark 5. As remarked, under (H.C.K), the inequalities verified by the law μ_T depend on the inequalities verified by the initial measure, in the range between Poincaré and logarithmic Sobolev inequality. Indeed, a logarithmic Sobolev inequality implies a WI inequality, but to get the the WI inequality for P_T we need only a WI inequality for the initial measure. As seen by the example of the Gaussian measure, which satisfies (H.C.K), no stronger inequalities can be obtained. ♢

3 General diffusion processes.

We shall now extend the results of the previous section to the general situation of (8),

$$\begin{aligned} dX_t &= \sigma(X_t)\, dB_t + b(X_t)\, dt\,. \\ \mathcal{L}(X_0) &= \mu_0\,. \end{aligned} \tag{30}$$

First of all we have to discuss some properties of the process and the associated quantities.

3.1 Some properties of the process.

3.1.1 Non explosion.

Since we assume that $\sigma \in C_b^2$, when (H.C.K) is fulfilled, b satisfies

$$2\,\langle b(x) - b(y), x - y\rangle \;\leq\; -D\,|x-y|^2\,,$$

for some $D \in \mathbb{R}$. In particular,

$$2\,\langle b(x), x\rangle \;\leq\; -D\,|x|^2 + 2\,|b(0)||x|\,.$$

Thus, if S_k denotes the exit time from the ball $B(x,k)$, and $t_k = t \wedge S_k$ it holds

$$\begin{aligned} \mathbb{E}(|X^x_{t_k}|^2) &= |x|^2 + \mathbb{E}\left(\int_0^{t_k} Trace(a(X^x_s)) + 2\,\langle b(X^x_s), X^x_s\rangle\, ds\right) \\ &\leq |x|^2 + Nt + |D| \int_0^t \mathbb{E}(|X^x_{s_k}|^2)\, ds + 2\,|b(0)| \int_0^t \mathbb{E}(|X^x_{s_k}|)\, ds \\ &\leq |x|^2 + (N + 2|b(0)|)t + (|D| + 2|b(0)|) \int_0^t \mathbb{E}(|X^x_{s_k}|^2)\, ds \end{aligned}$$

where, since σ is bounded, we have defined

$$N = \|\, Trace(a(.))\, \|_\infty\,,$$

and where we used $|y| \leq 1 + |y|^2$. Applying Gronwall lemma we obtain that $\mathbb{E}(|X^x_{t_k}|^2)$ is bounded independently on k, so that we may pass to the limit in k. This proves non explosion up to time t (since the explosion time is the increasing limit of the sequence S_k) for all t.

It is then easily seen that one can perform similar calculations with $g(t,x) = \exp(e^{-Ct}|x|^2)$ for a large enough C in order to kill the integrated term, i.e

Lemma 2. *There exists a large enough $C_e > 0$, such that $\mathbb{E}(\exp(e^{-C_e t}|X_t^x|^2)) \leq e^{|x|^2}$.*

It is interesting to notice that one can similarly obtain some "deviation" bound from the starting point. Indeed arguing as before, one can show the existence of constants $\alpha(T,D)$ and $\beta(T,D)$ such that for $0 \leq t \leq T$,

$$\mathbb{E}(|X_t^x - x|^2) \leq (\alpha(T,D)N + \beta(T,D)|b(x)|^2)\, t\,. \tag{31}$$

3.1.2 Properties of the semi-group.

Recall that, if (H.C.K) holds

$$e^{Kt}\,|X_t^x - X_t^y|^2 \leq |x-y|^2 + \int_0^t 2e^{Ks}\langle\sigma(X_s^x) - \sigma(X_s^y), X_s^x - X_s^y\rangle\, dB_s\,. \tag{32}$$

Notice that with our assumptions, the right hand side of (32) is a (true) martingale, so that

$$\mathbb{E}(|X_t^x - X_t^y|^2) \leq e^{-Kt}\,|x-y|^2\,. \tag{33}$$

Hence

Theorem 1. *If (H.C.K.) holds true, then*

$$W_2(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-Kt/2}\,|x-y|. \tag{34}$$

Moreover, if $\sigma\sigma^*$ is positive, (34) implies back (H.C.K.).
If we suppose moreover for some $m \geq 2$, the condition (H.C.K.m): $\forall(x,y)$

$$\frac{m}{2}|\sigma(x)-\sigma(y)|_{HS}^2 + m\langle b(x)-b(y), x-y\rangle + m\left(\frac{m}{2}-1\right)\frac{\|(x-y)(\sigma(x)-\sigma(y))^t\|^2}{\|x-y\|^2} \leq -K\,|x-y|^2$$

then

$$W_m(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-Kt/(m)}\,|x-y|. \tag{35}$$

Proof. The contraction in W_2 distance inherited from (H.C.K.) has already been proved. The contraction in W_m distance is done exactly in the same way using once again synchronous coupling. The necessary part comes from [10] (or more precisely section 4. in the Arxiv version 1110.3606). Let us explain the ideas of the proof. In fact, one may compute the time derivative of the Wasserstein distance: note $M : N = \sum_{i,j} M_{ij}N_{ij}$ when M and N are two matrices, then denoting ν_t and μ_t two solutions starting respectively from ν_0 and μ_0

$$\frac{d}{dt}W_2(\nu_t, \mu_t = 2\,J(\nu_t|\mu_t)$$

where if $\nu_t = \nabla\phi_t \# \mu_t$,

$$J(\nu_t, |\mu_t) = \int \Big[\frac{1}{2}\sigma\sigma^*(x) : (\nabla^2\phi_t(x) - I) \\ + \frac{1}{2}\sigma\sigma * (\nabla\phi_t(x)x) : (\nabla^2\phi_t(x)^{-1} - I) - \langle b(\nabla\phi_t(x)) \\ - b(x), \nabla\phi_t(x) - x\rangle\Big] d\mu_t.$$

Then the contraction property implies that at time 0 for $\nu_0 = \delta_y$ and $\mu_0 = \delta_x$

$$\frac{K}{2}|x-y|^2 \leq J(\nu_0, \mu_0).$$

A clever choice of ϕ then enables to prove the result.

Let f be C-Lipschitz continuous. It holds $|f(X_t^x) - f(X_t^y)| \leq C\,|X_t^x - X_t^y|$ so that, using (11), $P_t f$ is Lipschitz continuous with Lipschitz constant less than $C\,e^{-Kt/2}$.

As we said in the introduction, when $K > 0$ one deduces the existence and uniqueness of an invariant probability measure μ_∞, to which μ_T converges weakly.

In order to mimic what we have done in the previous section, we need the following: if $f \in \mathcal{A}$ (see the introduction), then $(t,x) \mapsto P_t f(x)$ is regular and satisfies (for $t > 0$)

$$\partial_t P_t f = P_t L f = L P_t f\,.$$

First if $f \in \mathcal{A}$, Lf is C_c^0 and we have $P_t f(x) - f(x) = \int_0^t P_s(Lf)(x)\,ds$. It follows that $\lim_{s\to 0} \frac{1}{s}\,(P_{t+s}f(x) - P_t f(x)) = P_t(Lf)(x)$ for all x, since $v \mapsto P_v Lf(x)$ is continuous. So $\partial_t P_t f = P_t L f$.

Now we need to know about the smoothness of $(t,x) \mapsto P_t f(x)$. But according to Theorems 40 and 49 in chapter V of Protter's book [39], $x \mapsto X_t^x$ is N times differentiable up to its explosion time, provided σ and b are $N+1$ times differentiable. So if we assume that σ and b are C^3, and that (H.C.K) holds true, we easily get the required regularity of P_t.

The commutation of L and P_t can then be proved as in [30] (where it was assumed that the coefficients belong to C_b^∞).

3.2 Commutation property with the gradient

Since we know that $\nabla P_t f$ exists, the calculations made in the introduction furnished a weaker form of Proposition 2. If σ is constant, we may use the arguments of the previous section.

Proposition 8. *Assume (H.C.K) or the weaker contraction property (34). Let $f \in \mathcal{A}$. It holds*
$$|\nabla P_t f|^2 \leq e^{-Kt}\, P_t(|\nabla f|^2)\,.$$
If σ is a constant matrix, we have the stronger
$$|\nabla P_t f| \leq e^{-Kt/2}\, P_t(|\nabla f|)\,.$$

Notice that, contrary to the Bakry-Emery bounded curvature case, the previous commutation property holds with the *usual gradient* and not with the *natural* one i.e. $\Gamma^{\frac{1}{2}}$.

If Proposition 2 allowed us to obtain logarithmic Sobolev inequalities, the weaker Proposition 8 will allow us to obtain a weaker inequality, namely a Poincaré inequality.

Remark 6. It is worth mentioning here the following alternate proof of the commutation property, starting from Wasserstein contraction, as derived in the recent paper [4] following our suggestion, i.e. using Kantorovitch-Rubinstein duality we have for all bounded Lipschitz ϕ denoting the inf convolution operator $Q_t\phi(x) = \inf_y\{\phi(y) + \frac{|x-y|^2}{2t}\}$ and initial measure μ_0 and ν_0

$$\begin{aligned}\int Q_1\phi d\mu_t - \int \phi d\nu_t &= \int P_t Q_1 \phi d\mu_0 - \int P_t\phi d\nu_0\\ &\leq e^{-Kt} W_2^2(\mu_0,\nu_0).\end{aligned}$$

Choose now $\mu_0 = \delta_x$, $\nu_0 = \delta_y$ to get for all y

$$P_t(Q_1\phi)(x) \leq P_t\phi(y) + \frac{|x-y|^2}{2e^{Kt}}$$

which by homogeneity of the inf-convolution operator gives

$$P_t(Q_1\phi) \leq Q_{e^{Kt}}(P_t\phi).$$

This assertion is in fact stronger than the gradient commutation property which can be deduced by using the fact that the inf-convolution operator is the Hopf-Lax solution of the Hamilton-Jacobi equation. ♢

3.3 h-processes and functional inequalities.

We now introduce the corresponding h-process. Let $T>0$ and $h>0$ be such that

$$\int P_T h \, d\mu_0 = 1$$

We thus may define on the path-space up to time T a new probability measure

$$\frac{d\mathbb{Q}}{d\mathbb{P}_{\mu_0}}|_{\mathcal{F}_T} \;=\; h(\omega_T)\,.$$

Again

$$\mathbb{Q}\circ\omega_s^{-1} = P_{T-s}h\,\mu_s \quad \text{ for all } \quad 0\le s\le T\,.$$

For simplicity, we assume in what follows that there exist c and C such that $C\ge h\ge c>0$. In this situation, using again Girsanov transform theory, we know that we can find a progressively measurable process u_s such that

$$\frac{d\mathbb{Q}}{d\mathbb{P}_{\mu_0}}|_{\mathcal{F}_T} \;=\; P_T h(\omega_0)\,\exp\left(\int_0^T \langle u_s, dM_s\rangle \;-\; \frac{1}{2}\int_0^T |\sigma(\omega_s)\,u_s|^2\,ds\right)\,,$$

where ω denotes the canonical element of the path-space and M denotes the martingale part of ω under $\mathbb{P}_{\mu_0}$. In addition, it can be shown [20] that

$$H(\mathbb{Q}|\mathbb{P}_{\mu_0}) \;=\; H(h\mu_T|\mu_T) \;=\; H(P_T h\mu_0|\mu_0) + \frac{1}{2}\,\mathbb{E}^{\mathbb{Q}}\left(\int_0^T |\sigma(\omega_s)\,u_s|^2\,ds\right)\,, \tag{36}$$

and

$$u_s = \nabla\,\log P_{T-s}h(\omega_s)$$

both $\mathbb{P}_{\mu_0}$ and $\mathbb{Q}$ almost surely.

We thus have

$$H(h\mu_T|\mu_T) \;=\; H(P_T h\mu_0|\mu_0) + \frac{1}{2}\int_0^T \left(\int \frac{|\sigma\,\nabla P_s h|^2}{P_s h}\,d\mu_{T-s}\right)ds\,. \tag{37}$$

Now define

$$M = \parallel |\sigma|^2 \parallel_\infty = \sup_y \sup_{|u|=1} |\sigma(y)u|^2\,.$$

If $h\in\mathcal{A}$ we may apply Proposition 8 in order to get (recall that $h\ge c$)

$$\begin{aligned} H(h\mu_T|\mu_T) &\leq H(P_T h\mu_0|\mu_0) + \frac{M}{2}\int_0^T \left(\int e^{-Ks}\,\frac{P_s(|\nabla h|^2)}{P_s h}\,d\mu_{T-s}\right) ds \\ &\leq H(P_T h\mu_0|\mu_0) + \frac{M}{2c}\int_0^T e^{-Ks}\left(\int |\nabla h|^2\,d\mu_T\right) ds \\ &\leq H(P_T h\mu_0|\mu_0) + \frac{M(1-e^{-KT})}{2cK}\int |\nabla h|^2\,d\mu_T\,, \end{aligned} \tag{38}$$

where we have used the Markov property for the second inequality.

Now let $g \in C_c^\infty$ be such that $\int g d\mu_T = \int P_T g\, d\mu_0 = 0$ and choose $h = 1 + \eta g \in \mathcal{A}$ so that $\int P_T h d\mu_0 = 1$ and $h > c > 0$ for η small enough. Actually we will let η go to 0 so that in the limit $c = 1$. Standard manipulations thus yield

$$\int g^2\,d\mu_T \leq \int (P_T g)^2\,d\mu_0 + \frac{M(1-e^{-KT})}{K}\int |\nabla g|^2\,d\mu_T\,. \tag{39}$$

We can eventually use first the density of C_c^∞ so that, arguing as for Proposition 3 we have obtained

Proposition 9. *Assume that (H.C.K) is satisfied. Let* $M = \| \, |\sigma|^2 \, \|_\infty$.
If μ_0 *satisfies a Poincaré inequality with constant* $C_P(0)$ *then* μ_T *satisfies a Poincaré inequality with constant*

$$C_P(T) = e^{-KT}\,C_P(0) + \frac{M(1-e^{-KT})}{K}\,.$$

This applies in particular to $P(T, x, .)$ *with* $C_P(0) = 0$.
If σ *is a constant matrix, a similar statement holds with the log-Sobolev constant instead of the Poincaré constant, replacing* M *by* $2M$.

Contrary to the log-Sobolev inequality, the Poincaré inequality does not furnish a transportation inequality, so we shall try to adapt what we did in subsection 2.3.

3.4 Transportation inequalities.

The situation is a little bit less simple than in the previous section. Indeed the martingale term is no more a Brownian motion and we can no more use characterization tricks on $(\Omega, \mathbb{Q})$. Hence we have to consider the solution of

$$dY_t = \sigma(Y_t)\,dB_t + b(Y_t)dt + a(Y_t)\,\nabla \log P_{T-t}h(Y_t)\,dt\,. \tag{40}$$

As before we assume first that $h \in \mathcal{A}$, $C \geq h \geq c > 0$ so that (40) is well defined and admits a unique strong solution. We can thus build a solution with the same Brownian motion B we used in (8). Strong uniqueness follows from the local Lipschitz property of all the coefficients and non explosion (up to time T) which is ensured by construction ($\mathbb{Q}$ is a probability measure). Again we may choose in an appropriate way the distribution of the pair of initial variables.

If (H.C.K) is satisfied, it holds

$$\begin{aligned}
\eta_t &= \mathbb{E}(|Y_t - X_t|^2) \quad (41)\\
&\leq \eta_0 \, - \, K \int_0^t \eta_s \, ds \quad (42)\\
&\qquad +2 \left(\int_0^t \mathbb{E}(\langle Y_s - X_s, a(Y_s) \, \nabla \log P_{T-s} h(Y_s) \rangle \right) ds \\
&\leq \eta_0 \, - \, K \int_0^t \eta_s \, ds + \\
&\quad + 2M^{\frac{1}{2}} \left(\int_0^t \eta_s \, ds \right)^{\frac{1}{2}} \left(\mathbb{E} \left(\int_0^t |\sigma(Y_s) \nabla \log P_{T-s} h(Y_s)|^2 \, ds \right) \right)^{\frac{1}{2}} \\
&\leq \eta_0 \, - \, K \int_0^t \eta_s \, ds + 2(2M)^{\frac{1}{2}} \left(\int_0^t \eta_s \, ds \right)^{\frac{1}{2}} H^{\frac{1}{2}}(h\mu_T|\mu_T) \, . \quad (43)
\end{aligned}$$

We may thus conclude as in the previous section

Proposition 10. *Assume that (H.C.K) is satisfied. Let $M = \parallel |\sigma|^2 \parallel_\infty$. The conclusions of Proposition 5 and Proposition 6 are still true, replacing C_T by MC_T*

Of course a T_2 inequality implies a Poincaré inequality, but the constant in Proposition 9 is better (in addition we only require that μ_0 satisfies a Poincaré inequality).

Remark 7. One of the renowned consequence of such inequalities is the concentration of measure phenomenon for μ_T. In particular, under the assumptions of Proposition 10, μ_T satisfies a gaussian type concentration property. In particular $|X_T^x|^2$ has some exponential moment, fact we have already shown in lemma 2. But this integrability does not reflect all the strength of the T_2 inequality whose tensorization property is particularly useful for statistical purposes.

When L is uniformly elliptic, this concentration property follows from gaussian estimates for the transition kernel. Here we obtain much more explicit constants (even if they are certainly far from optimality) which do not depend on the ellipticity constant. ♢

Remark 8. Assume that L is uniformly elliptic, i.e.

$$e = \inf_y \inf_{|u|=1} |\sigma(y)u|^2 \, > \, 0 \, .$$

Then we deduce from Proposition 9

$$P_T g^2(x) - (P_T g(x))^2 \leq \frac{2M}{e} \, \frac{(1-e^{-KT})}{K} \, P_T(\Gamma g)(x) \, .$$

According to [1] proposition 5.4.1, this is equivalent to the $CD(K/2, \infty)$ condition provided $M = e$ hence when σ is constant times the identity. In the non constant diffusion case, our condition (H.C.K) seems to be really different from the Bakry-Emery curvature condition. ♢

3.5 An hypoelliptic example : kinetic Fokker-Planck equation

We present in this section an application of the techniques developed here in an hypoelliptic example where the Bakry-Emery curvature is $-\infty$ and where (H.C.K.) may not be satisfied also.
Let (x_t, v_t) be the solution of the following SDE

$$\begin{aligned} dx_t &= v_t dt \\ dv_t &= dB_t - \nabla V(x_t) dt - v_t dt. \end{aligned}$$

also called stochastic Hamiltonian system. The long time behavior study of such a system has been considered for a long time and have been tackled by different techniques, see for example: hypocoercivity by Villani [42] or Lyapunov function technique by Bakry&al [3]. However, due to its hight degeneracy, the Bakry-Emery curvature is $-\infty$ so that we may not apply the Γ_2 technique. Remark also that the (H.C.K.) condition reads for all (x, v) and (y, w)

$$-\langle \nabla V(x) - \nabla V(y), v - w \rangle - |v-w|^2 \leq -K(|x-y|^2 + |v-w|^2)$$

so that it is hopeless to get $K > 0$.

Let us first remark that if ∇V is Lipshitz continuous, (H.C.K) is verified for some negative K and σ a constant diagonal matrix, so that we get that for some negative K the gradient commutation property holds

$$|\nabla P_t f| \leq e^{-Kt} \, P_t |\nabla f|$$

and thus the logarithmic Sobolev inequality holds for $P_t((x,v), \cdot)$. Let us remark once again that those properties are written with the usual gradient and not the Carré-du-Champ operator $\Gamma(f) = |\nabla_v f|^2$.

One may then wonder if it is possible to get the gradient commutation property with $K > 0$. In fact, using synchronous coupling and Itô's formula applied to the function $N((x,v),(y,w)) = a|x-y|^2 + b\langle x-y, v-w\rangle + |v-w|^2$, following [12], we get that if $V(x) = |x|^2 + W(x)$ where ∇W is δ-Lipschitz with δ sufficiently small there exists a, b and $K > 0$ such that N is equivalent to the euclidean norm and

$$N((x_t^x, v_t^v), (x_t^y, v_t^w)) \leq e^{-Kt} N((x,v),(y,w))$$

so that we get as in Section 2 the commutation property for some $K > 0$ and $A > 1$

$$|\nabla P_t f| \leq A\, e^{-Kt}\, P_t |\nabla f|$$

and thus a Logarithmic Sobolev inequality holds uniformly in time.
The previous method lies on the fact that one can replace the usual euclidean distance by another one, which is equivalent but more appropriate for the calculations. The same idea will be used with the mirror coupling.

It is not hard to extend the result of this simplified setting to the case where the Brownian motion in the velocity has a diffusion coefficient which is bounded and L-Lipschitz. We may then obtain a weaker gradient commutation property

$$|\nabla P_t f|^2 \leq A\, e^{-Kt}\, P_t |\nabla f|^2$$

and local Poincaré type inequality or Transportation information inequality like in Propositions 9 or 10, and if L is sufficiently small uniform in time version of these inequalities (using functional N).

3.6 Interpolation of the gradient commutation property and local Beckner inequality

We have seen here that we cannot recover a logarithmic Sobolev inequality by our technique when (H.C.K.) is in force, except when σ is a constant matrix. Remember however that we have introduced the stronger (H.C.K.m) condition which implies a contraction in Wasserstein distance W_m. It is then not hard to deduce some interpolation of the gradient commutation property

Proposition 11. *Assume (H.C.K.m) or the weaker contraction property (35). Let $f \in \mathcal{A}$, then*

$$|\nabla P_t f|^{\frac{m}{m-1}} \leq e^{-Kt/(m-1)}\, P_t \left(|\nabla f|^{\frac{m}{m-1}}\right). \tag{44}$$

Remark once again that this property does hold even if the diffusion coefficient is degenerate, so that variations of the hypoelliptic example of the previous subsection with a diffusion coefficient in the velocity enters into this framework. This contraction property may thus lead to a reinforcement of the Poincaré inequality to a Beckner inequality.

Proposition 12. *Assume (H.C.K.m) or the weaker (44). Let* $M = \| |\sigma|^2 \|_\infty$*. Then for all nice* f*, we have the following Beckner inequality*

$$P_t f^2 - P_t \left(|f|^{\frac{2m}{m+2}} \right)^{\frac{m+2}{m}} \leq M \frac{m+2}{m} \frac{1 - e^{2Kt/m}}{K} \, P_t |\nabla f|^2 .$$

The proof is similar to the one of Wang [45] Proposition 6.3.9.

4 Convergence to equilibrium in positive curvature.

When $K > 0$ we already mentioned that μ_T weakly converges to the unique invariant probability measure μ_∞ (which exists).
In particular , for all smooth g (say C_b^2), $\text{Var}_{\mu_T}(g) \to \text{Var}_{\mu_\infty}(g)$ as well as $\int |\nabla g|^2 \, d\mu_T \to \int |\nabla g|^2 \, d\mu_\infty$. We deduce that if σ is uniformly elliptic,

$$\text{Var}_{\mu_\infty}(g) \leq \frac{M}{K} \int |\nabla g|^2 \, d\mu_\infty \leq \frac{2M}{eK} \int \Gamma(g) \, d\mu_\infty \, .$$

Summarizing all this we have obtained

Theorem 2. *Assume that* σ *is bounded and uniformly elliptic. Then if (H.C.K) holds for some* $K > 0$*, defining* M *and* e *as before, there exists an unique invariant probability measure* μ_∞ *and* μ_∞ *satisfies a Poincaré inequality with constant* M/K*. In addition for all* $f \in \mathbb{L}^2(\mu_\infty)$ *it holds*

$$Var_{\mu_\infty}(P_t f) \leq e^{-KeT/M} \; Var_{\mu_\infty}(f) \, .$$

As we said, this result is not captured by the Γ_2 theory.

But we can obtain general convergence results, even in the non uniformly elliptic case. Indeed recall that in full generality

$$\text{Var}_{\mu_\infty}(P_t g) = \frac{1}{2} \int_t^{+\infty} \int |\sigma \, \nabla P_s g|^2 \, d\mu_\infty \, ds \, .$$

Using proposition 8 we thus have

$$\begin{aligned} \text{Var}_{\mu_\infty}(P_t g) &\leq \frac{M}{2} \int_t^{+\infty} \int |\nabla P_s g|^2 \, d\mu_\infty \, ds \\ &\leq \frac{M}{2} \int_t^{+\infty} e^{-Ks} \int P_s(|\nabla g|^2) \, d\mu_\infty \, ds \\ &\leq \frac{M}{2K} \, e^{-Kt} \int |\nabla g|^2 \, d\mu_\infty \, . \end{aligned}$$

Hence

Theorem 3. *Assume that σ is bounded. Then if (H.C.K) holds for some $K > 0$, defining M as before, there exists an unique invariant probability measure μ_∞ and for all nice enough function g ,*

$$Var_{\mu_\infty}(P_t g) \leq \frac{M}{2K}\, e^{-Kt} \int |\nabla g|^2 \, d\mu_\infty \,.$$

In addition if μ_∞ is symmetric (i.e. $\int fLg d\mu_\infty = \int gLf d\mu_\infty$), it holds

$$Var_{\mu_\infty}(P_t g) \leq e^{-Kt}\, Var_{\mu_\infty}(g)\,.$$

Remark once again that what is used here is the weak gradient commutation property which is a consequence of (H.C.K.). The last part of the theorem follows from lemma 1. Of course, unless we explicitly know the invariant measure, it is not easy to see wether μ_∞ is symmetric or not.

We may further extends the previous argument to the entropic convergence to equilibrium. Let us suppose that there exists an unique invariant measure μ_∞.

Theorem 4. *Assume that σ is bounded and that the gradient commutation property*

$$|\nabla P_t f| \leq c\, e^{-Kt}\, P_t |\nabla f| \tag{45}$$

holds for some positive K. Then for all nice positive function f (defining M as before)

$$Ent_{\mu_\infty}(P_t f) \leq \frac{cM}{K} e^{-Kt} \int \frac{|\nabla g|^2}{g} d\mu_\infty.$$

Proof. The proof is as for the L_2 decay quite standard. Indeed,

$$\begin{aligned}
\mathrm{Ent}_{\mu_\infty}(P_t g) &\leq M \int_t^\infty \int \frac{|\nabla P_s g|^2}{P_s g} d\mu_\infty ds, \\
&\leq cM \int_t^\infty e^{-Ks} \int P_s \frac{|\nabla g|^2}{g} d\mu_\infty ds, \\
&\leq \frac{cM}{K} e^{-Kt} \int \frac{|\nabla g|^2}{g} d\mu_\infty.
\end{aligned}$$

Remark 9. One of the important point here is that we do not suppose any non-degeneracy on the diffusion coefficient, so that the result applies to the kinetic Fokker-Planck equation. It then provides an alternative to the approach by Villani [42], where he obtained such kind of convergence by completely different techniques with assumptions quite similar to the ones described in Section 3.5. One may then complete the approach by regularization of the Fisher Information in small time to obtain an entropic decay controlled by the initial entropy, see [42] or [28]. ♢

Remark 10. Let us point out that even in the symmetric case, such a control is not sufficient to recover a logarithmic Sobolev inequality as the analog of lemma 1 is no more valid for the entropy. Remark however that we have shown in section 2 how to recover a logarithmic Sobolev inequality for P_t using the strong commutation gradient property (45). If $K > 0$, we may then let t goes to infinity to recover a logarithmic Sobolev inequality for the invariant measure. It may be, for example, used in the context of kinetic Fokker-Planck equation with non gradient coefficient, for which the invariant measure is unknown. ♢

Remark 11. Let us consider, as in the Poincaré case via (H.C.K.) condition, a particular class of test function g such that $g \geq \varepsilon > 0$, so that $P_t g \geq \varepsilon$. We then see adapting the preceding proof that a weak commutation of gradient property

$$|\nabla P_t f|^2 \leq c\, e^{-Kt}\, P_t |\nabla f|^2 \tag{46}$$

obtained for example under (H.C.K.) condition implies that

$$\mathrm{Ent}_{\mu_\infty}(P_t g) \leq \frac{cM}{\varepsilon K}\, e^{-Kt} \int |\nabla g|^2 d\mu_\infty .$$

As we said in the introduction, an exponential decay of Wasserstein distances furnishes some Poincaré inequality for μ. In what follows W_0 denotes the total variation distance and W_1 is the usual 1-Wasserstein distance. We restate Proposition 1

Proposition 13. *Assume that μ is reversible. Assume that for all bounded (resp. Lipschitz) density of probability h we have $W_0(P_t h\mu, \mu) \leq c_h(t)$ (reps. W_1). Then for all bounded (resp. Lipschitz and bounded) f, there exist c_f and h such that $Var_\mu(P_t f) \leq c_f\, c_h(2t)$. In particular if $c_h(t) = c_h\, e^{-\beta t}$, μ satisfies a Poincaré inequality.*

Proof. Let f be bounded and centered, and

$$h = (f + \| f \|_\infty) / \int (f + \| f \|_\infty)\, d\mu = 1 + (f / \| f \|_\infty) .$$

h is thus a density of probability with $\| h \|_\infty \leq 2$. We have

$$\begin{aligned} \mathrm{Var}_\mu(P_t f) &= \| f \|_\infty^2 \ \mathrm{Var}_\mu(P_t h) \\ &\leq \| f \|_\infty^2 \int P_t h (P_t h - 1)\, d\mu = \| f \|_\infty^2 \int h (P_{2t} h - 1)\, d\mu \\ &\leq \| f \|_\infty^2 \| h \|_\infty \ W_0(P_{2t} h\mu, \mu) \ \leq \ 2 \ \| f \|_\infty^2 \ c_h(2t) . \end{aligned}$$

One can replace W_0 by W_1, just replacing $\| h \|_\infty$ by $\| \nabla h \|_\infty$ in which case

$$\mathrm{Var}_\mu(P_t f) \leq \| f \|_\infty \| \nabla f \|_\infty \ c_h(2t) .$$

Remark 12. The previous result partly extends to the non symmetric situation. Indeed if we do not use the symmetry of P_t, but only the fact that $\| P_t \| \leq 1$ in $\mathbb{L}^\infty$ we obtain that provided $W_0(P_t h\mu, \mu) \leq c_h(t)$,

$$\mathrm{Var}_\mu(P_t f) \leq 2 \ \| f \|_\infty^2 \ c_h(t) \, . \qquad \diamondsuit$$

Even when the decay is not exponential, one gets a weak form of the Poincaré inequality (called a weak Poincaré inequality).

Corollary 1. *In the situation of proposition 13, assume that $c_h(t) = c_h \, c(t)$ with $c(t) \to 0$ as $t \to +\infty$. Then*

$$Var_\mu(f) \, \leq \, \alpha(s) \int |\nabla f|^2 \, d\mu + \, s \, \Psi(f) \, ,$$

for all $s > 0$ where $\Psi(f) = c_h \ \| f - \int f d\mu \|_\infty^2$ for W_0 and $\Psi(f) = c_h \ \| f - \int f d\mu \|_\infty \| \nabla f \|_\infty$ for W_1 and $\alpha(s) = s \, \inf_{u>0} \, \frac{1}{u} \, c^{-1}(u \exp(1 - (u/s)))$; h being defined in the proof of proposition 13 .

Proof. Once we notice that the transformation $f \mapsto \lambda f$ does not change h the result follows from [40] Theorem 2.3.

Here is an application in dimension 1, suggested by B. Jourdain [31] who has others very interesting results in one dimension for the control of Wasserstein distances.
If $X_.$ is a solution of (8) in dimension 1, assume that

$$\langle b(x) - b(y) \, , \, x - y \rangle \leq -K \, |x - y|^2 \, . \tag{47}$$

As we showed in the introduction, up to the synchronous coupling time

$$e^{Kt} |X_t^x - X_t^y| \leq |x - y| + M_t$$

where M_t is a martingale term. It follows that $W_1(\mu_t, \nu_t) \leq e^{-Kt} \, W_1(\mu_0, \nu_0)$, so that there exists an unique invariant probability measure $\mu_\infty = e^{-V} dx$. For μ_∞ to be reversible, we must have $2b = a' - aV'$ where $a = \sigma^2$.

5 Non homogeneous diffusion processes.

5.1 General non homogeneous diffusion

In [21] the authors extended the Γ_2 theory to time dependent coefficients (non homogeneous diffusions). Considering the Ito system

$$\begin{aligned} dX_t &= \sigma(v_t, X_t)\, dB_t + b(v_t, X_t)\, dt\,. \\ dv_t &= dt\,, \\ \mathcal{L}(v_0, X_0) &= \delta_{t_0} \otimes \mu_0\,, \end{aligned} \tag{48}$$

we see that all what we have done can be applied to this system. Actually one can modify the "curvature" assumptions introducing for some function $K(t)$ and its derivative $K'(t)$: **(H.C.K(t))** for all (x, y), all $t \in \mathbb{R}$

$$|\sigma(t,x) - \sigma(t,y)|^2_{HS} + 2\,\langle b(t,x) - b(t,y), x - y\rangle \;\leq\; -K'(t)\,|x-y|^2\,.$$

We then have

Theorem 5. *Assume that σ and b satisfy the hypotheses of the previous section, considered as functions on $\mathbb{R} \times \mathbb{R}^n$. If (H.C.K(t)) is satisfied, then we may extend (1) (Poincaré) and (2)(log-Sobolev) replacing e^{-KT} by $e^{-K(T)}$ and $\frac{1-e^{-KT}}{K}$ by $\int_0^T e^{-K(s)}\, ds$.*

Proof. If f only depends on x, the proof of proposition 2 (resp. 8) is unchanged using the process starting from $(0, x)$ and $(0, y)$ and replacing Kt by $K(t)$. To obtain the analogue of proposition 3 and proposition 9, it suffices to remark that $\sigma\nabla$ is equal to ∇_x, and use what precedes for h depending on x only.

For the transportation inequality we have to slightly modify the method in subsection 2.3. With the notations therein, (20) has become,

$$\eta_t \leq \eta_0 \;-\; \int_0^t K'(s)\, \eta_s\, ds + 2\sqrt{2}\, H^{\frac{1}{2}}(h\mu_T|\mu_T)\, \left(\int_0^t \eta_s\, ds\right)^{\frac{1}{2}}\,,$$

so that, as in the previous section we have to come back to

$$\eta_t \leq \eta_0 \;-\; \int_0^t K'(s)\, \eta_s\, ds + 2 \int_0^t \mathbb{E}^{\mathbb{Q}}\left(|z_s - \omega_s|\, |\nabla \log P_{T-s}h(\omega_s)|\right) ds\,. \tag{49}$$

Using as usual $(ab)^{\frac{1}{2}} \leq \lambda a + \frac{1}{\lambda} b$ we obtain (see the details of the derivation in the previous section) that for all increasing function $\lambda(t)$

$$\eta'(t) \leq (-K'(t) + \lambda'(t))\, \eta_t + \frac{4}{\lambda'(t)} I_T(h)\,,$$

from which we deduce, provided we choose $K(0) = \lambda(0) = 0$,

$$\eta_T \;\leq\; e^{-K(T)+\lambda(T)}\, \eta_0 \;+\; 4\, e^{-K(T)+\lambda(T)} \left(\int_0^T \frac{e^{K(s)-\lambda(s)}}{\lambda'(s)}\, ds\right) I_T(h)\,.$$

Theorem 6. *In the situation of theorem 5. If (H.C.K(t)) is satisfied, then for any x and any increasing function λ, $P(T, x, .)$ satisfies a W_2I inequality*

$$W_2^2(hP(T,x,.),P(T,x,.)) \leq C(T) \int \frac{|\nabla h|^2}{h}\, d\mu_T\,,$$

with constant

$$C(T) \leq 4\, e^{-K(T)+\lambda(T)} \left(\int_0^T \frac{e^{K(s)-\lambda(s)}}{\lambda'(s)}\, ds \right) .$$

If μ_0 satisfies a T_2 inequality with constant $C_T(0)$, then

$$W_2^2(h\mu_T,\mu_T) \leq C_T(0)\, e^{-K(T)+\lambda(T)}\, H(h\mu_T|\mu_T) + C(T)\, I_T(h)\,.$$

The best choice of λ is not clear. If $K'(t)$ is not positive on the whole $[0,T]$, it seams that taking $\lambda(T) = \lambda T$ for some $\lambda > 0$ is enough. If $K'(t) > 0$ for all t (but not necessarily bounded from below by a positive constant), $\lambda(t) = \lambda K(t)$ seems to be natural.

Remark 13. Assume that $K(t) \to +\infty$ as $t \to +\infty$ and that $C = \int_0^{+\infty} e^{-K(s)}\, ds < +\infty$. Then, for all t, $P(t,x,.)$ (the distribution of the process starting from x and $t_0 = 0$) satisfies a Poincaré inequality (and a log-Sobolev inequality when $\sigma = Id$) with a constant bounded by MC (or $2C$). The family $(P(t,x,.))_{t>0}$ is then tight, but we do not know whether it is weakly convergent or not. Nevertheless any weak limit satisfies the same functional inequality.
When $\sigma = Id$ we know that $|X_t^x - X_t| \leq e^{-K(t)}|x - X_0|$ for any initial random variable X_0. It follows that if a sequence $P(t_k,x,.)$ is weakly convergent to some μ, the sequence μ_{t_k} weakly converges to the same limit.
In particular if we consider $\sigma = Id$, $b(t,x) = -\frac{1}{2}\,(\nabla U(x) + K'(t)\,x)$, for some convex potential U, (H.C.K(t)) is satisfied, so that any weak limit satisfies a log-Sobolev inequality. If $d\mu = e^{-U}dx$ does not satisfy a log-Sobolev inequality, it cannot be a weak limit, even if $K'(t) \to 0$. In this situation one should expect that the "perturbation" of ∇U being smaller and smaller when t growths, the convergence to μ will still hold. This is not the case. ♢

5.2 Application to some non-linear diffusions.

We shall now discuss an example that does partly enter the framework of the beginning of this section.

Following [36, 18] consider the following non-linear stochastic differential equation

$$\begin{aligned} dX_t &= dB_t - \frac{1}{2}\,\nabla V(X_t)\, dt - \frac{1}{2}\nabla W * q_t(X_t)\, dt \qquad (50)\\ \mathcal{L}(X_t) &= q_t\, dx\,. \end{aligned}$$

If a solution exists, q_t will solve

$$\partial_t q_t = \frac{1}{2}\,\nabla.\left(\nabla q_t + q_t \nabla V + q_t(\nabla W * q_t)\right)\,. \tag{51}$$

This is a non-linear diffusion of Mc Kean-Vlasov type modeling, for instance, granular media. We refer to the introduction of [18] for details and motivations. One can approximate the solution of (50) by the first coordinate of a linear large particle system with mean field interactions. This is what is done in [36, 18] to study the long time behavior of X_t.

Let us see how to apply what we have just done. First, under some conditions on V and W (we later shall give some of them) existence and weak uniqueness of (50) are ensured, provided the initial law admits some big enough polynomial moment. This will imply for all x, the existence and uniqueness of q_t^x solution of (51) with initial condition δ_x. As usual for these non-linear equations, if we consider the linear time inhomogeneous S.D.E.

$$dZ_t^{x,y} = dB_t - \frac{1}{2}\,\nabla V(Z_t^{x,y})\,dt - \frac{1}{2}\,\nabla W * q_t^x(Z_t^{x,y})\,dt \quad Z_0^{x,y} = y\,,$$

the pathwise unique solution (up to explosion) $Z_.^{x,x}$ is shown to satisfy (50) (i.e. $\mathcal{L}(Z_t^{x,x}) = q_t^x$) so that it coincides with X_t^x. So, once q_t^x and q_t^y are built, we may build our synchronous coupling (X_t^x, X_t^y) as before. We may thus state

Theorem 7. *Assume that*

H1 *V, W and their first two derivatives have at most polynomial growth of order m and $W(-x) = W(x)$,*
H2 *V satisfies (H.C.K_V) and W satisfies (H.C.K_W).*

Let $a = \max(m(m+3), 2m^2)$. If μ_0 and ν_0 have a polynomial moment of order a, there exist an unique solution of (50) and an unique solution of (51) among the set of probability flows having having a polynomial moment of order a with initial condition μ_0 or ν_0.
Furthermore

1.
$$W_2^2(\mu_T, \nu_T) = W_2^2(q_T^{\mu_0}\,dx, q_T^{\nu_0}\,dx) \le e^{-(K_V + \min(K_W,0))T}\, W_2^2(\mu_0, \nu_0)\,.$$

2. *If $V = 0$ and $\int x\mu_0(dx) = \int x\nu_0(dx)$ then*
$$W_2^2(\mu_T, \nu_T) \le e^{-K_W T}\, W_2^2(\mu_0, \nu_0)\,.$$

Introduce the conditions,

H'1 $K = K_V + \min(K_W, 0) > 0$.
H'2 $V = 0$, $\int x\mu_0(dx) = \int x\nu_0(dx)$ *and* $K_W > 0$.

If H'1 is satisfied, there exists an unique invariant distribution $\mu_\infty = q^\infty(x)dx$ of (50) and (51) satisfying the polynomial moment condition of order a, the convergence to μ_∞ in W_2 Wasserstein distance being exponential as above. If H'2 is satisfied the same result holds for each $A \in \mathbb{R}^n$ in the set of probability measures such that $\int x\mu(dx) = A$.

Proof. The moment condition ensuring existence and uniqueness is described in [18] section 2.

Recall that we may build our synchronous coupling (X_t^x, X_t^y) as before. Now introduce an independent copy $(\bar{X}_t^x, \bar{X}_t^y)$ of (X_t^x, X_t^y).

We have

$$\begin{aligned} &\mathbb{E}\left(|X_t^x - X_t^y|^2\right) \\ &= -\mathbb{E}\left(\int_0^t \langle \nabla V(X_s^x) - \nabla V(X_s^y), X_s^x - X_s^y\rangle ds\right) \qquad (52) \\ &-\mathbb{E}\int_0^t \int \langle \nabla W(X_s^x - z^x) - \nabla W(X_s^y - z^y), X_s^x - X_s^y\rangle q_s^x(z^x) q_s^y(z^y) dz^x dz^y ds. \end{aligned}$$

Remark that the last term can be written

$$\int_0^t \mathbb{E}\left(\langle \nabla W(X_s^x - \bar{X}_s^x) - \nabla W(X_s^y - \bar{X}_s^y), X_s^x - X_s^y\rangle\right) ds\,.$$

If we assume in addition (as usual) that $W(-x) = W(x)$, and remember that $\bar{X}$ is a copy of X, it is still equal to

$$-\int_0^t \mathbb{E}\left(\langle \nabla W(X_s^x - \bar{X}_s^x) - \nabla W(X_s^y - \bar{X}_s^y), \bar{X}_s^x - \bar{X}_s^y\rangle\right) ds\,.$$

Hence

$$\begin{aligned} &2\,\mathbb{E}\left(|X_t^x - X_t^y|^2\right) \\ &= \mathbb{E}\left(|X_t^x - X_t^y|^2\right) + \mathbb{E}\left(|\bar{X}_t^x - \bar{X}_t^y|^2\right) \\ &= 2|x-y|^2 - 2\,\mathbb{E}\left(\int_0^t \langle \nabla V(X_s^x) - \nabla V(X_s^y), X_s^x - X_s^y\rangle ds\right) \\ &- \int_0^t \mathbb{E}\left(\langle \nabla W(X_s^x - \bar{X}_s^x) - \nabla W(X_s^y - \bar{X}_s^y), (X_s^x - \bar{X}_s^x) - (X_s^y - \bar{X}_s^y)\rangle\right) ds\,. \end{aligned}$$

According to what precedes we have

$$\begin{aligned}\mathbb{E}\left(|X_t^x - X_t^y|^2\right) &\leq |x-y|^2 - K_V \int_0^t \mathbb{E}\left(|X_s^x - X_s^y|^2\right) ds \\ &\quad - (K_W/2) \int_0^t \mathbb{E}\left(|(X_s^x - \bar{X}_s^x) - (X_s^y - \bar{X}_s^y)|^2\right) ds \\ &\leq |x-y|^2 - K_V \int_0^t \mathbb{E}\left(|X_s^x - X_s^y|^2\right) ds \\ &\quad - K_W \left(\int_0^t \mathbb{E}\left(|X_s^x - X_s^y|^2\right) ds - \int_0^t |\mathbb{E}\left(X_s^x - X_s^y\right)|^2 ds\right) .\end{aligned}$$

Of course we may replace the initial δ_x and δ_y by probability distributions μ_0 and ν_0 satisfying the required moment conditions. This immediately furnishes the first assertion about the upper bound for the Wasserstein distance.

If $V = 0$ it is easily seen that $\int x q_t^{\mu_0}(x) dx = \int x \mu_0(dx)$ for all $t > 0$, hence

$$E\left(X_s^{\mu_0} - X_s^{\nu_0}\right) = 0$$

provided the same holds at time 0. This furnishes the second assertion for the upper bound.

Finally the convergence under strict positivity of our new "curvature" condition ensures the existence of the limiting measure μ_∞. To see that $\mu_\infty = q^\infty(x)dx$ is actually invariant, one can for instance use the following trick: first consider the solution q_t^∞ of (51) with initial condition q^∞. Similar bounds for the Markov non homogeneous process $Z_.^{q^\infty,y}$ (when we replace q_t^x by q_t^∞) are obtained applying the results of the beginning of this section. Hence the law of $Z_T^{q^\infty,q^\infty}$ (which is exactly μ_T starting with μ_∞ as we explained before) converges to some limiting measure $\mu_\infty^{\mu_\infty}$ which in turn is equal to μ_∞ and is invariant for $Z_.^{q_\infty,y}$. This achieves the proof.

Remark 14. The proof of the above result is new and direct, while the result is mainly contained in [36, 18] using particle approximation. Notice that in [36] the Γ_2 approach is developed for the non homogeneous Markov diffusion $Z_.$ and not for $X_.$. Also notice that some direct study of the decay to equilibrium in W_2 distance for granular media is done in [11]. $\diamondsuit$

As said in the previous remark the Γ_2 theory does not work directly for the process $X_.$. Actually our method to control the gradient of $x \mapsto \mathbb{E}(f(X_t^x))$ should work but we do not know whether the gradient exists or not, due to the fact that we do not have any a priori regularity in the initial condition. Fortunately, if we want to obtain some properties for the time marginal distribution μ_T we may use the fact (as done by Malrieu) that this distribution coincides with the one of the non homogeneous Markov diffusion $Z_T^{x,y}$ to which we can apply the techniques of this section. In particular, in the situation of the previous theorem, when q^∞ exists we may consider the diffusion

$$dZ_t^y = dB_t - \frac{1}{2}\nabla V(Z_t^y)\,dt - \frac{1}{2}\nabla W * q^\infty(Z_t^y)\,dt \quad Z_0^y = y\,,$$

for which $\mu_\infty(dx) = q^\infty(x)dx$ the invariant probability measure. Using the results in section 2 we thus have

Proposition 14. *In the situation of Theorem 7, if H'1 or H'2 are satisfied, μ_∞ satisfies a log-Sobolev inequality with constant $C_{LS} = 2/K$ or $C_{LS} = 2/K_W$.*

All what we have done extends to more general Mc Kean-Vlasov equations, with a diffusion coefficient σ and a drift b satisfying hypothesis (R). In particular, positive curvature (in the sense of (H.C.K)) will also imply existence of and convergence to an invariant probability measure. The only difference is that we have to replace log-Sobolev inequality by Poincaré inequality in the latter proposition. Let us explain quickly what kind of model we may consider. We do not aim to be optimal, but will provide a flavor of the results on contraction with some non constant diffusion term. We will not focus also on the existence of solution of such equation. Let X_t^x be solution of

$$dX_t^x = \sigma(X_t^x, \kappa * q_t(X_t^x))dB_t - \frac{1}{2}\nabla V(X_t^x)\,dt - \frac{1}{2}\nabla W * q_t^x(X_t^x)\,dt \qquad (53)$$

$$X_0^x = x \qquad (54)$$

$$\mathcal{L}(X_t^x) = q_t^x\,dx\,.$$

Theorem 8. *Let us suppose H1 and H2, that κ is l-Lipschitz and that*

$$|\sigma(x,y) - \sigma(x',y')|_{HS}^2 \le r(|x-x'|^2 + |y-y'|^2).$$

Then (using the notation of Th.7)

$$W_2^2(\mu_T, \nu_T) \le e^{-(K_V - r(1+4l^2) + \min(K_W,0))T}\, W_2^2(\mu_0, \nu_0).$$

Suppose moreover that $K_V - r(1+4l^2) + \min(K_W,0) > 0$, then there exists an unique invariant distribution to (53), the convergence to μ_∞ in W_2 Wasserstein distance being exponential as above.

The proof follows the same line as before except that in the Itô's formula, there is the diffusion part which comes into play for which we use the Lipschitz condition of the theorem. Note that Bolley&al [12] have considered the case of a kinetic McKean-Vlasov equation, but with a constant diffusion coefficient in speed. As before, we may obtain some functional inequality for the invariant distribution as in Prop. 14 but we have to replace log-Sobolev inequality by Poincaré inequality .

6 Extensions to some non uniformly convex potentials.

Let us come back to (1), and assume that Υ is bounded. We shall extend (H.C.K) to more general situations. The first natural extension is to replace the squared distance by some other convex functional of the distance. More precisely.

Definition 1. Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$. We say that φ belongs to $\mathcal{C}$ if it satisfies the following conditions:

- φ is increasing and convex, with $\varphi(0) = 0$ and $\varphi(1) = 1$,
- $a \mapsto \varphi(a)/a$ is non decreasing,
- there exist a positive function ψ such that for all $a > 0$ and all $\lambda > 0$, $\varphi^{-1}(\lambda a) \leq \psi(\lambda)\, \varphi^{-1}(a)$, where φ^{-1} denotes the inverse (reciprocal) function of φ.

Definition 2. Let $\varphi \in \mathcal{C}$. We shall say that **(H.φ.K)** is satisfied for some $K > 0$ if for all (x, y),

$$\langle \nabla U(x) - \nabla U(y), x - y\rangle \geq K\, \varphi(|x - y|^2)\,.$$

On one hand, since $K > 0$ and $\varphi \geq 0$, (H.φ.K) implies that U is convex. On the other hand, if (H.φ.K) is satisfied, since U is smooth, $\varphi(a)/a$ is necessarily bounded near the origin since $\limsup_{a\to 0}(\varphi(a)/a) \leq \inf |Hess(U)|$. Here of course if $\varphi \in \mathcal{C}$ the latter is automatically satisfied.

If $\varphi(a) = a$ this is nothing else but (H.C.K). If $\varphi(a)/a \to +\infty$ we shall say that U is super-convex. This terminology is justified by the example below.

Example 1. Let $U(x) = (|x|^2)^\beta$ for some $\beta > 1$. We shall see that (H.φ.K) is satisfied for $\varphi(a) = a^\beta$ and some K we shall estimate.

We start with the one dimensional case. In this case

$$(U'(x) - U'(y))(x - y) = 2\beta\, (sign(x)|x|^{2\beta-1} - sign(y)|y|^{2\beta-1})(x - y)\,.$$

If $sign(x) = sign(y)$, we may assume that $|x| \geq |y|$, write $|x| = u + |y|$ for $u \geq 0$ and remark that if $2\beta - 1 \geq 1$,

$$(u + |y|)^{2\beta-1} - |y|^{2\beta-1} \geq u^{2\beta-1}$$

so that

$$(U'(x) - U'(y))(x - y) = 2\beta\, ((u + |y|)^{2\beta-1} - |y|^{2\beta-1})u \geq 2\beta u^{2\beta} = 2\beta|x - y|^{2\beta}\,.$$

If $sign(x) = -sign(y)$, we have, using the convexity of $x \mapsto |x|^{2\beta-1}$,

$$\begin{aligned}(U'(x) - U'(y))(x - y) &= 2\beta\, (|x|^{2\beta-1} + |y|^{2\beta-1})(|x| + |y|)\\ &\geq 2\beta\, 2^{2-2\beta}\, (|x| + |y|)^{2\beta}\\ &= 2\beta\, 2^{2-2\beta}\, |x - y|^{2\beta}\,.\end{aligned}$$

Since $\beta > 1$, we may choose $K_\beta = 2\beta\, 2^{2-2\beta}$.

The general situation is a little bit more intricate.
Pick x and y in $\mathbb{R}^n$, assume that $|x| \geq |y|$ and write $x = |x|u$ and $y = |y|(\alpha u + \gamma v)$ for unit vectors u and v such that $\langle u, v\rangle = 0$ and $\alpha^2 + \gamma^2 = 1$. Then

$$\langle \nabla U(x) - \nabla U(y), x-y\rangle = 2\beta \left((|x|^{2\beta-1} - \alpha|y|^{2\beta-1})(|x| - \alpha|y|) + \gamma^2|y|^{2\beta}\right) ,$$

and

$$|x-y|^{2\beta} = \left((|x|-\alpha|y|)^2 + \gamma^2|y|^2\right)^\beta \;\leq\; 2^{\beta-1}\left((|x|-\alpha|y|)^{2\beta} + \gamma^{2\beta}|y|^{2\beta}\right).$$

If $\alpha \geq 0$, we write again $|x| = |y| + a$ with $a \geq 0$. Thus , since $0 \leq 1-\alpha \leq 1$,

$$\begin{aligned} |x|^{2\beta-1} - \alpha|y|^{2\beta-1} &= (a+|y|)^{2\beta-1} - \alpha|y|^{2\beta-1} \geq a^{2\beta-1} + (1-\alpha)|y|^{2\beta-1} \\ &\geq a^{2\beta-1} + ((1-\alpha)|y|)^{2\beta-1} \\ &\geq 2^{2-2\beta}\,(a + (1-\alpha)|y|)^{2\beta-1} = 2^{2-2\beta}\,(|x|-\alpha|y|)^{2\beta-1}. \end{aligned}$$

It follows, since $\beta \geq 1$ and $\gamma^2 \leq 1$,

$$\begin{aligned} \langle \nabla U(x) - \nabla U(y), x-y\rangle &\geq 2\beta\, \left(2^{2-2\beta}\,(|x|-\alpha|y|)^{2\beta} + \gamma^2\,|y|^{2\beta}\right) \\ &\geq 2\beta\, 2^{2-2\beta}\, \left((|x|-\alpha|y|)^{2\beta} + \gamma^{2\beta}\,|y|^{2\beta}\right) \\ &\geq 2\beta\, 2^{3-3\beta}\,|x-y|^{2\beta}. \end{aligned}$$

If $\alpha < 0$, since $|\alpha| \leq 1$, it holds

$$\begin{aligned} &\langle \nabla U(x) - \nabla U(y), x-y\rangle \\ &\quad = 2\beta\, \left((|x|^{2\beta-1} + |\alpha||y|^{2\beta-1})(|x| + |\alpha||y|) + \gamma^2\,|y|^{2\beta}\right) \\ &\quad \geq 2\beta\, \left((|x|^{2\beta-1} + (|\alpha||y|)^{2\beta-1})(|x| + |\alpha||y|) + \gamma^{2\beta}\,|y|^{2\beta}\right) \\ &\quad \geq 2\beta\, \left(2^{2-2\beta}(|x| + |\alpha||y|)^{2\beta} + \gamma^{2\beta}\,|y|^{2\beta}\right) \\ &\quad \geq 2\beta\, 2^{3-3\beta}\,|x-y|^{2\beta}. \end{aligned}$$

Proposition 15. *Let $U(x) = (|x|^2)^\beta$ for some $\beta > 1$. Then (H.φ.K) is satisfied for $\varphi(a) = a^\beta$ and $K_\beta \geq 2\beta\, 2^{3-3\beta}$. If $n = 1$ we have the better bound $K_\beta \geq 2\beta\, 2^{2-2\beta}$.* ♢

Remark 15. If $\varphi \in \mathcal{C}$, for all $a \geq 0$ and all $\varepsilon > 0$, it holds

$$\varphi(a) \geq \frac{\varphi(\varepsilon)}{\varepsilon}\, a \,-\, \varphi(\varepsilon).$$

Hence (H.φ.K) implies the following condition

$$\forall \varepsilon > 0, \forall (x,y) \quad \langle \nabla U(x) - \nabla U(y), x-y \rangle \geq K \left(\frac{\varphi(\varepsilon)}{\varepsilon} \, |x-y|^2 \, - \, \varphi(\varepsilon) \right) . \tag{55}$$

The latter appears in the study of the granular medium equation in [18] (condition(6)) for power functions φ. This formulation will be the interesting one. It can be extended in

Definition 3. Let α be a non decreasing function defined on $\mathbb{R}^+$. We shall say that **(H.α.K)** is satisfied for some $K > 0$ if for all (x, y) and all $\varepsilon > 0$,

$$\langle \nabla U(x) - \nabla U(y), x-y \rangle \geq K \, \alpha(\varepsilon) \, (|x-y|^2 - \varepsilon) \, .$$

(H.φ.K) implies (H.α.K) with the same K and $\alpha(\varepsilon) = \varphi(\varepsilon)/\varepsilon$. In this definition we do not need that $a \mapsto a\alpha(a)$ is convex.

Now we shall see how to use (H.φ.K).

6.1 Non fully convincing first results.

This subsection contains first results which are not really convincing, but have to be tested.
If we want to control the gradient $\nabla P_t f$, we may write for $t > u$,

$$\begin{aligned} |X_t^x - X_t^y|^2 &= |X_u^x - X_u^y|^2 - \int_u^t \langle \nabla U(X_s^x) - \nabla U(X_s^y), X_s^x - X_s^y \rangle \, ds \\ &\leq |X_u^x - X_u^y|^2 \, - \, K \int_u^t \varphi(|X_s^x - X_s^y|^2) \, ds \, . \end{aligned}$$

Denoting $\eta_t = |X_t^x - X_t^y|^2$, we thus have $\eta_t' \leq -K \, \varphi(\eta_t)$. If $\varphi(a) = a^\beta$, this yields

$$|X_t^x - X_t^y|^2 \leq |x-y|^2 \, \left(\frac{1}{1 + K(\beta-1)|x-y|^{2(\beta-1)} \, t} \right)^{1/(\beta-1)} . \tag{56}$$

This result (even after taking expectation) is not really satisfactory. Indeed, first we do not obtain any better control for $\nabla P_t f$ than the one for a general convex potential (in particular we do not obtain a rate of convergence to 0). In second place, the decay to 0 of the Wasserstein distance we obtain is desperately slow, while we expected an exponential decay (which we know to hold true for $U(x) = |x|^{2\beta}$ for $\beta \geq 1$). Notice however that we recover the exponential decay we obtained previously when $\beta \to 1$.

Remark 16. If instead of (H.φ.K) we use (H.α.K), it is not difficult to show that

$$\eta_t \leq \eta_0 \, e^{-K\alpha(\varepsilon)t} + \varepsilon \, .$$

If $\alpha(\varepsilon) = \varepsilon^{\beta-1}$, choosing $\varepsilon = \eta_0 \, t^{-\theta}$ for some $\theta < \beta - 1$, we get

$$W_2^2(P(t,x,.), P(t,y,.)) \leq |x-y|^2 \, (t^{-\theta} + e^{-K|x-y|t^{\beta-1-\theta}}) \, .$$

The method can be extended to the Mc Kean-Vlasov situation studied in subsection 5.2 and allows us to recover (up to the constants) Theorem 4.1 in [18] without the help of a particle approximation. However, better results in this situation are obtained in [11]. ♢

Mimicking subsection 2.3, in particular (20), do we obtain more interesting results ? Using the notation therein we have

$$\eta_t := \mathbb{E}^{\mathbb{Q}}(|z_t - \omega_t|^2) \leq \eta_0 - K \int_0^t \varphi(\eta_s) \, ds + 2\sqrt{2} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \left(\int_0^t \eta_s \, ds \right)^{\frac{1}{2}} . \tag{57}$$

Using Jensen inequality we deduce

$$\begin{aligned} \varphi\left(\frac{1}{t} \int_0^t \eta_s \, ds \right) &\leq \frac{1}{t} \int_0^t \varphi(\eta_s) \, ds \\ &\leq \frac{2\sqrt{2}}{K\,t} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \left(\int_0^t \eta_s \, ds \right)^{\frac{1}{2}} \end{aligned}$$

so that, if $v_t = \int_0^t \eta_s \, ds$,

$$\begin{aligned} v_t &\leq t \, \varphi^{-1}\left(\frac{2\sqrt{2}}{K\,t} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \, v_t^{\frac{1}{2}} \right) \\ &\leq t \, \psi\left(\frac{2\sqrt{2}}{K\,t} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \right) \, \varphi^{-1}(v_t^{1/2}) \, . \end{aligned}$$

If $\varphi(a) = a^\beta$, we thus obtain

$$\begin{aligned} \eta_T &\leq \eta_0 + 2\sqrt{2} \, H^{\frac{1}{2}}(h\mu_T|\mu_T) \left(\int_0^T \eta_s \, ds \right)^{\frac{1}{2}} \\ &\leq \eta_0 + (2\sqrt{2})^{\frac{\beta+1}{\beta}} \, K^{-1/\beta} \, H^{\frac{\beta+1}{2\beta}}(h\mu_T|\mu_T) \, T^{\frac{\beta-1}{2\beta-1}} \, . \end{aligned} \tag{58}$$

This result is certainly not fully satisfactory too. On one hand, we get a less explosive bound in time (recall that in the general convex case the bound growths like T), but on the other hand the relative entropy appears to a power less than 1. In particular such an inequality does not imply a Poincaré inequality (which is obtained for entropies going to 0), but furnishes nice concentration properties (obtained for large entropies via Marton's argument).

6.2 An improvement of Bakry-Emery criterion.

As we remarked at this end of section 2 we may come back to the initial inequality in (20) which becomes in our new situation

$$\eta_t \leq \eta_0 \, - \, K \int_0^t \varphi(\eta_s) \, ds + 2 \int_0^t \mathbb{E}^{\mathbb{Q}} \left(|z_s - \omega_s| \, |\nabla \log P_{T-s} h(\omega_s)| \right) ds \, , \quad (59)$$

and yields

$$\eta_t' \leq - \, K \, \varphi(\eta_t) + 2 \, \eta_t^{\frac{1}{2}} \, \left(\int \frac{|\nabla h|^2}{h} \, d\mu_T \right)^{\frac{1}{2}} \, . \quad (60)$$

(here again (H.C.0) is satisfied so that, for short, $|\nabla P_s| \leq P_s |\nabla|$.) To explore (60) we shall use both the remark 15 and the usual trick $ab \leq \lambda a^2 + \frac{1}{\lambda} \, b^2$ for a, b, λ positive. Hence

$$\eta_t' \leq \left(-K \, \frac{\varphi(\varepsilon)}{\varepsilon} + 2\lambda \right) \, \eta_t + \left(\frac{2}{\lambda} \, \left(\int \frac{|\nabla h|^2}{h} \, d\mu_T \right) + K \varphi(\varepsilon) \right) \, . \quad (61)$$

We deduce, denoting $A = K \, \frac{\varphi(\varepsilon)}{\varepsilon} - 2\lambda$,

$$\eta_T \leq \eta_0 \, e^{-AT} + (1 - e^{-AT}) \, \frac{\frac{2}{\lambda} \left(\int \frac{|\nabla h|^2}{h} \, d\mu_T \right) + K \varphi(\varepsilon)}{A} \, .$$

Choose $\lambda = (1/4) \, K \, (\varphi(\varepsilon)/\varepsilon)$ so that $A = (1/2) \, K \, (\varphi(\varepsilon)/\varepsilon) > 0$. η_T is thus bounded in time, but the bound is not tractable except for $T = +\infty$ (starting with $\mu_0 = \mu$) or if $\eta_0 = 0$. In both cases we have obtained

$$W_2^2(h\mu_T, \mu_T) \leq \varepsilon \, + \, \left(\frac{8\varepsilon^2}{K^2 \, \varphi^2(\varepsilon)} \right) \, \int \frac{|\nabla h|^2}{h} \, d\mu_T \, . \quad (62)$$

It remains to optimize in ε. In full generality we choose ε such that both terms in the sum of (62) are equal (we know that we are loosing a factor less than 2). Remark that we do not use the explicit form of φ, i.e. we may replace (H.φ.K) by (H.α.K) in what we did previously. We have thus obtained

Proposition 16. *Assume that U satisfies (H.α.K) for $K > 0$. Let F be the inverse (reciprocal) function of $\varepsilon \mapsto \varepsilon \, \alpha^2(\varepsilon)$. Denote $\mu_T = P(T, x, .)$ and $\mu_\infty(dx) = \mu(dx) = e^{-U(x)} \, dx$. Then for all $0 < T \leq +\infty$, μ_T satisfies for all nice h,*

$$W_2^2(h\mu_T, \mu_T) \leq 2\, F\left(\frac{8}{K^2}\, I_T(h)\right) ,$$

where $I_T(h) = \int \frac{|\nabla h|^2}{h}\, d\mu_T$ *is the Fisher information of* h.

When F is equal to identity, such an inequality is called a W_2I inequality (see [26] definition 10.4). Here we obtained a weak form of W_2I inequality (which is clear on (62)) in the spirit of the weak Poincaré or the weak log-Sobolev inequalities.

In particular, using (H.W.I), since (H.C.0) is satisfied we obtain

Corollary 2. *Under the hypotheses of proposition 16, μ satisfies the inequality*

$$H(h\mu|\mu) \leq 2\left(I(h)\, F\left(\frac{8}{K^2}\, I(h)\right)\right)^{\frac{1}{2}} .$$

Weak logarithmic Sobolev inequalities were introduced and studied in [16]. Actually, we are not exactly here in the situation of [16] because we wrote the previous inequality in terms of a density of probability. Let $h = f^2/\int f^2 d\mu$. We deduce from the previous corollary

$$\int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu \leq 4\left(\int f^2 d\mu\right)^{\frac{1}{2}} \left(\int |\nabla f|^2\, d\mu\right)^{\frac{1}{2}} F^{\frac{1}{2}}\left(\frac{32}{K^2}\, \frac{\int |\nabla f|^2\, d\mu}{\int f^2\, d\mu}\right) ,$$

so that if $F(\lambda\, a) \leq \theta(\lambda)\, F(a)$,

$$\int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu \leq 4\, \theta^{\frac{1}{2}}\left(\frac{32}{K^2}\right) G_1\left(\int f^2 d\mu\right) G_2\left(\int |\nabla f|^2\, d\mu\right) ,$$

where $G_1(a) = a^{\frac{1}{2}}\, \theta^{\frac{1}{2}}(1/a)$ and $G_2(a) = a^{\frac{1}{2}}\, F^{\frac{1}{2}}(a)$. The previous inequality looks like the Nash inequality version of a weak log-Sobolev inequality, but with the $\mathbb{L}^2$ norm of f in place of the $\mathbb{L}^\infty$ norm of $f - \int f d\mu$. So the previous inequality is not only "weak" but also "defective".

6.2.1 Super convex potentials.

In this sub(sub)section we assume that $\varphi(a) = a^\beta$ for some $\beta \geq 1$, so that $F(a) = a^{\frac{1}{2\beta-1}}$. We thus have

$$\int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu \leq 4\left(\frac{32}{K^2}\right)^{\frac{1}{2(2\beta-1)}} \left(\int f^2 d\mu\right)^{\frac{\beta-1}{2\beta-1}} \left(\int |\nabla f|^2\, d\mu\right)^{\frac{\beta}{2\beta-1}} . \tag{63}$$

Recall first that if $g \geq 0$, then $\text{Var}_\mu(g) \leq \text{Ent}_\mu(g)$ (see e.g. [17] (2.6)).
Next recall the following: defining $m_\mu(g)$ as a *median* of g, we have

$$\text{Var}_\mu(g) \;\leq\; 4 \int (g - m_\mu(g))^2 \, d\mu \;\leq\; 36 \, \text{Var}_\mu(g) \,. \tag{64}$$

We may decompose $f - m_\mu(f) = (f - m_\mu(f))_+ - (f - m_\mu(f))_- = g_+ - g_-$ so that both g_+ and g_- are non negative with median equal to 0. In addition, if f is Lipschitz, so are g_+ and g_-, $\nabla f = \nabla g_+ + \nabla g_-$, and the product of both vanishes. Hence

$$\text{Var}_\mu(f) \leq 4 \left(\int (g_+)^2 d\mu + \int (g_-)^2 d\mu \right) ,$$

while

$$\begin{aligned} \int (g_+)^2 d\mu &\leq 9 \, \text{Var}_\mu(g_+) \;\leq\; 9 \, \text{Ent}_\mu(g_+) \\ &\leq 36 \left(\frac{32}{K^2} \right)^{\frac{1}{2(2\beta-1)}} \left(\int (g_+)^2 d\mu \right)^{\frac{\beta-1}{2\beta-1}} \left(\int |\nabla g_+|^2 \, d\mu \right)^{\frac{\beta}{2\beta-1}} \,. \end{aligned}$$

It follows from (63)

$$\int (g_+)^2 d\mu \leq (36)^{\frac{2\beta-1}{\beta}} \left(\frac{32}{K^2} \right)^{\frac{1}{2\beta}} \int |\nabla g_+|^2 \, d\mu \,,$$

similarly for g_-. We have thus obtained

Theorem 9. *Assume that U satisfies $(H.\varphi.K)$ for $K > 0$ and $\varphi(a) = a^\beta$, for $\beta \geq 1$. Then, μ satisfies both a Poincaré inequality with*

$$C_P(\mu) \leq C_P(K, \beta) = 4 \, (36)^{\frac{2\beta-1}{\beta}} \left(\frac{32}{K^2} \right)^{\frac{1}{2\beta}} ,$$

and a log-Sobolev inequality with

$$C_{LS}(\mu) \leq C_{LS}(K, \beta) = \left(\frac{32}{K^2} \right)^{\frac{1}{2\beta}} \left(4^{\frac{3\beta-2}{2\beta-1}} 36^{\frac{\beta-1}{\beta}} + 8 \times 36^{\frac{2\beta-1}{\beta}} \right) .$$

Proof. The statement on the Poincaré inequality follows from the previous discussion.
Concerning the log-Sobolev inequality, let $\tilde{f} = f - \int f d\mu$. Then, Rothaus lemma (see [1] lemma 4.3.7) says that

$$\text{Ent}_\mu(f) \leq \text{Ent}_\mu(\tilde{f}) + 2 \, \text{Var}_\mu(f) \,.$$

Applying (63) to $\tilde{f}$ together with the Poincaré inequality, yield the result (after some elementary calculation).

Remark 17. This theorem applies in particular to $U(x) = |x|^{2\beta}$ for $\beta \geq 1$, according to proposition 15. The fact that μ satisfies a log-Sobolev inequality in this situation is well known, but here we obtain an explicit (though not really cute) expression for the constant that only depends on β and not on the dimension n.
Unfortunately, in this particular situation, our bounds are not optimal. Indeed, spherically symmetric log-concave probability measures are now well understood.

For the Poincaré constant, it was shown by Bobkov [8] that

$$\frac{1}{n}\,\mathrm{Var}_\mu(x) \;\leq\; C_P(\mu) \leq \frac{13}{n}\,\mathrm{Var}_\mu(x)\,.$$

It is an (easy) exercise to see that $\mathrm{Var}_\mu(x) = \Gamma((n+2)/2\beta)/\Gamma(n/2\beta)$, so that $C_P(\mu) \leq c(\beta)\, n^{\frac{1}{\beta}-1}$ which goes to 0 as $n \to +\infty$.
A famous conjecture by Kannan-Lovasz-Simonovitz is that the previous bound for spherically symmetric measures extends (up to a change of the constant 13) to any log-concave measure. If true, the KLS conjecture will presumably give a better upper bound for the Poincaré constant than ours.

Regarding the log-Sobolev constant, the work by Huet [29], furnishes a lower bound for the isoperimetric profile of μ (see Theorem 3 and the discussion p.98 therein) which indicates a similar bound for the log-Sobolev constant as above, i.e. depending on the isotropic constant ($n^{\frac{1-\beta}{2\beta}}$) of μ. ◇

6.2.2 Lack of uniform convexity.

Now choose $\alpha(a) = a^\beta$ for some $\beta \geq 1$ and $a \leq 1$, and $\alpha(a) = 1$ for $a \geq 1$. (H.α.K) is less restrictive than before since it only implies a linear behavior at infinity for the gradient of the potential.

We now have $F(a) = a^{\frac{1}{2\beta+1}}$ for $a \leq 1$ and $F(a) = a$ for $a \geq 1$. It follows

$$\mathrm{Ent}_\mu(f) \leq \; 4\,\left(\frac{32}{K^2}\right)^{\frac{1}{2(2\beta+1)}}\,\left(\int f^2 d\mu\right)^{\frac{\beta}{2\beta+1}}\,\left(\int |\nabla f|^2\, d\mu\right)^{\frac{\beta+1}{2\beta+1}} \tag{65}$$

if $\int |\nabla f|^2\, d\mu \leq \frac{K^2}{32}\,\int f^2 d\mu$ and

$$\mathrm{Ent}_\mu(f) \leq \frac{32}{K}\int |\nabla f|^2\, d\mu \quad \text{otherwise.} \tag{66}$$

Proceeding as before we obtain

Theorem 10. *Assume that U satisfies (H.α.K) for $K>0$ and $\alpha(a)=a^\beta\wedge 1$ for $\beta\geq 1$. Then, μ satisfies both a Poincaré inequality with*

$$C_P(\mu)\leq C_P(K,\beta)=\max\left(\frac{32}{K}\,,\,4\,(36)^{\frac{2\beta+1}{\beta+1}}\,\left(\frac{32}{K^2}\right)^{\frac{1}{2(\beta+1)}}\right)\,,$$

and a log-Sobolev inequality with

$$C_{LS}(\mu)\leq C_{LS}(K,\beta)=\max\left(\frac{32}{K}\,,\,\left(\frac{32}{K^2}\right)^{\frac{1}{2(\beta+1)}}\,\left(4^{\frac{3\beta+1}{2\beta+1}}36^{\frac{\beta}{\beta+1}}+8\times 36^{\frac{2\beta+1}{\beta+1}}\right)\right)\,.$$

Remark 18. Using the general form of the (H.W.I) inequality it is quite easy to adapt the previous proof in order to show the following result :
Let μ satisfying (H.C.K) for some $K>-\infty$. If μ satisfies a weak W_2I inequality,

$$W_2^2(h\mu,\mu)\leq C\,(I(h))^\beta$$

for some $0<\beta\leq 1$, then μ satisfies a log-Sobolev inequality with a constant depending on C,K,β only. In particular μ satisfies a T_2 inequality.

In particular if we know that μ_T has a bounded below curvature, the previous theorems extend to μ_T. ♢

7 Using reflection coupling.

As we have seen in the previous section, the simple coupling using the same Brownian motion is not fully well suited to deal with non uniformly convex potentials.
In a recent note [23], Eberle studied the contraction property in Wasserstein distance W_1, induced by another well known coupling method: coupling by reflection (or mirror coupling) introduced in [35]. The results of the note are extended in the recent [24] which appeared more or less at the same time than the first version of the present paper.
We shall see now how to use this coupling method in the spirit of what we have done before.

7.1 Reflection coupling for the drifted brownian motion.

In this section we consider X_t^x the solution starting from x of the Ito stochastic differential equation

$$dX_t = dB_t + b(X_t)\,dt\,, \tag{67}$$

where b is smooth enough. We introduce another formulation of the semi-convexity property, namely :

$$\kappa(r) = \inf\left\{-2\,\frac{\langle b(x)-b(y), x-y\rangle}{|x-y|^2}\;;\;|x-y|=r\right\}\,, \tag{68}$$

so that, it always holds

$$2\,\langle b(x)-b(y), x-y\rangle \;\leq\; -\kappa(|x-y|)\,|x-y|^2\,.$$

We shall say that **(H.**κ**)** is satisfied if

$$\liminf_{r\to+\infty}\,\kappa(r) = \kappa_\infty > 0\,. \tag{69}$$

This condition is typically some "uniform convexity at infinity" condition. Indeed if $b = -\frac{1}{2}\,\nabla U$ where $U = U_1 + U_2$ with U_1 satisfying (H.C.κ_∞) and U_2 compactly supported, then (H.κ) is satisfied. We shall come back later to this. Notice that if (69) is satisfied, the solution of (67) is strongly unique and non explosive, using the same tools as we used before.
Now, following [23] we introduce (with some slight change of notations)

$$\begin{aligned}
R_0 &= \inf\left\{R\geq 0\,;\,\kappa(r)\geq 0\,,\,\forall r\geq R\right\}\,,\\
R_1 &= \inf\left\{R\geq R_0\,;\,\kappa(r)\geq 8/(R(R-R_0))\,,\,\forall r\geq R\right\}\,,\\
\varphi(r) &= \exp\left(-\,\frac{1}{4}\int_0^r s\kappa^-(s)\,ds\right)\,,\quad \Phi(r) = \int_0^r \varphi(s)\,ds\,,\\
g(r) &= 1-\frac{1}{2}\left(\int_0^{r\wedge R_1}\frac{\Phi(s)}{\varphi(s)}\,ds\Big/\int_0^{R_1}\frac{\Phi(s)}{\varphi(s)}\,ds\right)\\
D(r) &= \int_0^r \varphi(s)\,g(s)\,ds\,.
\end{aligned} \tag{70}$$

Notice that

$$\frac{1}{2}\leq g\leq 1 \;\text{ and }\; \exp\left(-\,\frac{1}{4}\int_0^{R_0} s\kappa^-(s)\,ds\right) = \varphi_{min}\leq\varphi\leq 1\,.$$

If (H.κ) is satisfied, $R_0 < +\infty$ so that $\varphi_{min} > 0$ and

$$\frac{\varphi_{min}}{2}\,r\;\leq D(r)\;\leq\; r\,,$$

i.e. $D(|x-y|)$ which is actually a distance, is equivalent to the euclidean distance.
Hence a consequence of Theorem 1 in [23] is the following :

Theorem 11. *Assume that (H.κ) is satisfied. Let λ be defined by*

$$\frac{1}{\lambda} = \int_0^{R_1} \frac{\Phi(s)}{\varphi(s)}\, ds \;\leq\; \frac{R_1^2}{\varphi_{min}}\,.$$

Then for all initial distributions ν and μ, and all t, the W_1 Wasserstein distance satisfies

$$W_1(\nu_t, \mu_t) \leq \frac{2}{\varphi_{min}}\, e^{-\lambda t}\, W_1(\nu, \mu)\,.$$

In order to prove this result, Eberle adds to (67) the following Ito s.d.e.

$$dY_t = (Id - 2e_t e_t^*)\, dB_t + b(Y_t)\, dt\,,\; Y_0 = y\,, \tag{71}$$

where $e_t = (X_t - Y_t)/|X_t - Y_t|$ and e^* is the transposed of e (remark that if $n = 1$, it just changes B into $-B$). Of course one has to consider (67) and (71) together. Existence, strong uniqueness and non explosion are again easy to show. Now introduce the coupling time T_c defined by

$$T_c = \inf\, t \geq 0\,;\, X_t = Y_t\,,$$

and finally define

$$\bar{X}_t^y = Y_t^y \;\text{ if } t \leq T_c\,, \quad \bar{X}_t^y = X_t^x \;\text{ if } t \geq T_c\,.$$

It is easy to see that $X_.^y$ and $\bar{X}_.^y$ have the same law, so that the distribution of $(X_t^x, \bar{X}_t^y)$ is a coupling of $P(t, x, .)$ and $P(t, y, .)$. Of course this extends to any initial distributions (μ, ν) and furnishes a coupling of (μ_t, ν_t).

It follows that $Z_t = X_t^x - \bar{X}_t^y$ solves

$$dZ_t = (b(X_t^x) - b(\bar{X}_t^y))dt + 2\, \frac{Z_t}{|Z_t|}\, dW_t \tag{72}$$

where $W_t = \int_0^t e_s^*\, dB_s$ is a one dimensional brownian motion.

The key of the proof of Theorem 11 is then that, if $r_t = D(|X_t^x - \bar{X}_t^y|)$, $r_.$ is a semi-martingale with decomposition

$$r_t = D(|x - y|) + \int_0^{t \wedge T_c} 2\, \varphi(r_s)\, g(r_s)\, dW_s + \int_0^{t \wedge T_c} \beta_s\, ds\,, \tag{73}$$

where the drift term satisfies

$$\beta_s \leq -\, \lambda\, r_s\,. \tag{74}$$

Taking expectation, it immediately shows that the W_D Wasserstein distance decays exponentially fast. As remarked by several authors, one can then deduce as we did previously

$$|\nabla P_t f| \leq \frac{2}{\varphi_{min}} \, e^{-\lambda t} \ \parallel \nabla f \parallel_\infty \ . \tag{75}$$

As a consequence we obtain

Theorem 12. *Assume that* $b = -\frac{1}{2}\nabla U$ *satisfies(H.κ) and that* $\mu = (1/Z_U)e^{-U}$ *is a probability measure. Then* μ *satisfies a Poincaré inequality with constant* $C_P \leq (1/2\lambda)$.

Proof. Recall that

$$\mathrm{Var}_\mu(P_t f) = \frac{1}{2} \int_t^{+\infty} \int |\nabla P_s f|^2 \, d\mu \, ds \, .$$

According to (75), we thus have

$$\mathrm{Var}_\mu(P_t f) \leq \frac{1}{\lambda \, \varphi_{min}^2} \, e^{-\,2\,\lambda t} \ \parallel \nabla f \parallel_\infty^2$$

for all Lipschitz function f. According to Lemma 2.12 in [19], we deduce that $\mathrm{Var}_\mu(P_t f) \leq e^{-\,2\,\lambda t}\,\mathrm{Var}_\mu(f)$, hence the result.

Remark 19. One can see that the reflection coupling cannot furnish some information on W_2, the concavity of D near the origin being crucial. In the same negative direction, theorem 12 cannot be extended to the log-Sobolev framework, the key lemma 2.12 in [19] being restricted to the variance control. ♢

Example(1) If (H.φ.K) is satisfied with $\varphi(a) = a^\beta$, we have $\kappa(a) = a^{2(\beta-1)}$. We thus have $R_0 = 0$, $R_1 = (8/K)^{\frac{1}{2\beta}}$, $\varphi_{min} = 1$ and finally

$$C_P \leq \frac{1}{2} \left(\frac{8}{K}\right)^{\frac{1}{\beta}} \, .$$

We recover the result in Theorem 9, i.e. a bound $C_\beta \, K^{-\,1/\beta}$ but with a better constant C_β.

(2) If (H.α.K) is satisfied, one can choose $R_0 = \sqrt{2\varepsilon}$, $R_1 = \sqrt{2\varepsilon} + (4/\sqrt{K\alpha(\varepsilon)})$, $\varphi_{min} = \exp\left(-\,\frac{1}{4}\,\varepsilon^2\,K\alpha(\varepsilon)\right)$ and finally

$$C_P \leq \left(2\varepsilon + \frac{16}{K\alpha(\varepsilon)}\right) \, e^{K\varepsilon^2\,\alpha(\varepsilon)/4} \, .$$

(3) Now assume that the potential U can be written $U = U_1 + U_2$ where U_1 satisfies (H.C.K) for some $K > 0$ and U_2 satisfies $\parallel \nabla U_2 \parallel_\infty = M < +\infty$. It easily follows that (H.κ) is satisfied with $\kappa(a) = K - \frac{M}{a}$. We thus have

$$R_0 = \frac{M}{K} \, , \quad R_1 = \frac{M}{K} + \sqrt{\frac{8}{K}} \, , \quad \varphi_{min} = e^{-\,\frac{M^2}{8K}} \, .$$

We finally obtain

$$C_P \leq \left(\frac{\| \nabla U_2 \|_\infty}{K} + \sqrt{\frac{8}{K}} \right)^2 \exp \left(\frac{\| \nabla U_2 \|_\infty^2}{8K} \right) .$$

An old result by Miclo (unpublished but explained in [33]) indicates that such a result (without the square of the supremum of the gradient but without K in the exponential) can be obtained by using the usual Holley-Stroock perturbation argument. $\diamondsuit$

7.2 The log-concave situation.

Now consider the situation where b satisfies (H.C.0). In this situation we have $\lambda = 0$ ($R_1 = +\infty$, $\varphi = g = 1$) so that (73) becomes

$$dr_t = 2\, dW_t + \beta_t\, dt$$

for $t \leq T_c$ with $\beta_t \leq 0$. It follows that $r_t \leq |x-y| + 2\, W_t$ up to the first time $T_{|x-y|}$ the brownian motion $W_.$ hits $-\,|x-y|/2$.
In particular,

$$\mathbb{P}(r_t > 0) \;\leq\; \mathbb{P}(t < T_{|x-y|}) \;\leq\; \frac{|x-y|}{\sqrt{2\pi\, t}} \,,$$

since the law of $T_{|x-y|}$ is given by

$$\mathbb{P}(T_{|x-y|} \in da) = \frac{|x-y|}{2\sqrt{2\pi\, a^3}}\, e^{-|x-y|^2/8a}\, \mathbb{1}_{a>0}\, da \,.$$

As a first by-product we obtain

Proposition 17. *If b satisfies (H.C.0) then*

$$|\nabla P_t f| \leq \frac{2}{\sqrt{2\pi\, t}} \;\; \| f \|_\infty \;\; .$$

Actually if $b = -\frac{1}{2}\nabla U$ with U convex (i.e. in the zero curvature situation of the Γ_2 theory) the inequality $|\nabla P_t f| \leq \frac{1}{\sqrt{t}} \;\; \| f \|_\infty$ is well known as a consequence of what is called the reverse (local) Poincaré inequality (see [1]). The previous proposition extends this result (up to the constant) to a non-gradient drift.

Proof. Recall that $X_t^x = \bar{X}_t^y$ for $t > T_c$. It follows

$$\begin{aligned} P_t f(x) - P_t f(y) &= \mathbb{E}\left((f(X_t^x) - f(\bar{X}_t^y))\, 1\!\!1_{T_c > t} \right) \\ &\leq 2 \parallel f \parallel_\infty \, \mathbb{P}(T_c > t) \, \leq \, 2 \parallel f \parallel_\infty \, \mathbb{P}(t < T_{|x-y|}) \\ &\leq \, \frac{2 \parallel f \parallel_\infty}{\sqrt{2\pi\, t}} \, |x-y| \, , \end{aligned}$$

hence the result.

If one wants to get a contraction bound for the gradient (in the spirit of (75) or better of proposition 2) we cannot only use a comparison with the brownian motion for which $\nabla P_t f = P_t \nabla f$.

Remark 20. In the symmetric situation ($b = -\frac{1}{2}\, \nabla U$) it is known that $t \mapsto \int \, |\nabla P_t f|^2 \, d\mu$ is non increasing. It easily follows that

$$\parallel \nabla P_t f \parallel_{\mathbb{L}^2(\mu)} \leq \frac{\sqrt{2}}{\sqrt{t}} \, \parallel f \parallel_{\mathbb{L}^2(\mu)} \ .$$

If (H.C.0) is satisfied, this remark together with Proposition 17 and Riesz-Thorin interpolation theorem show that, up to an universal constant, the same holds in all $\mathbb{L}^p(\mu)$ spaces. ♢

Remark 21. If we assume that (H.κ) is satisfied, we may replace the comparison with a Brownian motion by the comparison with an Ornstein-Uhlenbeck process with parameter $\lambda/2$, according to standard comparison theorems for one dimensional Ito processes (see e.g [30] Chapter VI theorem 1.1). For the O-U process, it is known (see [38]) that

$$\mathbb{P}(T_{|x-y|} \in da) = \frac{|x-y|}{2\sqrt{2\pi}} \, \left(\frac{\lambda}{2\sinh(a\lambda/2)}\right)^{\frac{3}{2}} \, e^{\left(- \, \frac{\lambda \, |x-y|^2 e^{-a\lambda/2}}{16 \sinh(a\lambda/2)} \, + \, \frac{a\lambda}{4} \right)} \, da \, .$$

An explicit bound for $\mathbb{P}(t < T_{|x-y|})$ can be obtained by using the reflection principle in [47], namely

$$\mathbb{P}(t < T_{|x-y|}) \leq \frac{\sqrt{\lambda}\, e^{-t\lambda/2}}{\sqrt{2\pi}\, \sqrt{1 - e^{-t\lambda}}} \, |x-y| \, ,$$

yielding

$$|\nabla P_t f| \leq \, \frac{2}{\varphi_{min}} \, \frac{\sqrt{\lambda}\, e^{-t\lambda/2}}{\sqrt{2\pi}\, \sqrt{1 - e^{-t\lambda}}} \, \parallel f \parallel_\infty \ .$$

These bounds are interesting as regularization bounds (from bounded to Lipschitz functions), but notice that we have lost a factor 2 in the exponential decay. ♢

7.3 Reflection coupling for general diffusions.

The case of a general diffusion process with a non constant diffusion matrix as in section 3 is more delicate to handle, as already remarked in [35] (Theorem 1).

Assume that σ is a bounded and smooth square matrices field and that it is uniformly elliptic. The quantities we need here are (notations differ from [35])

$$M = \sup_x \sup_{|u|=1} |\sigma(x)\, u|^2 \, , \; N = \sup_x \sup_{|u|=1} |\sigma^{-1}(x)\, u|^2 \, , \; \Lambda = \sup_{x,x'} \sup_{|u|=1} |(\sigma(x) - \sigma(x'))\, u|^2 \, . \tag{76}$$

Recall the Lindvall-Rogers reflection coupling

$$\begin{aligned} dX_t &= \sigma(X_t) dB_t + b(X_t) dt \, , \\ dX'_t &= \sigma(X'_t)\, H_t\, dB_t + b(X'_t) dt \, , \end{aligned}$$

where

$$H_t \, = \, Id \, - \, 2 \left(\frac{\sigma^{-1}(X'_t)\, (X_t - X'_t)}{|\sigma^{-1}(X'_t)(X_t - X'_t)|} \right) \left(\frac{\sigma^{-1}(X'_t)\, (X_t - X'_t)}{|\sigma^{-1}(X'_t)(X_t - X'_t)|} \right)^* \, .$$

Existence and strong uniqueness can be shown as previously. Of course, as in subsection 7.1, we replace X'_t by X_t if $t > T_c$ the coupling time, but not to introduce new notation we still use $X'_.$.

As in [35] define

$$Y_t = X_t - X'_t \, , \; V_t = \frac{Y_t}{|Y_t|} \, , \; \alpha_t = \sigma(X_t) - \sigma(X'_t) H_t \, , \; \beta_t = b(X_t) - b(X'_t) \, .$$

According to (15) in [35] we have

$$d(|Y_t|) = \langle V_t, \alpha_t \, dB_t \rangle \; + \; \frac{1}{2} \, \frac{1}{|Y_t|} \, \left(2 \langle Y_t, \beta_t \rangle + Trace(\alpha_t \, \alpha_t^*) - |\alpha_t^* \, V_t|^2 \right) \, ,$$

and a simple calculation shows that

$$\begin{aligned} & Trace(\alpha_t \, \alpha_t^*) - |\alpha_t^* \, V_t|^2 = \\ &= Trace((\sigma(X_t) - \sigma(X'_t))\, (\sigma(X_t) - \sigma(X'_t))^*) \, - \, |(\sigma(X_t) - \sigma(X'_t))^* \, V_t|^2 \, , \end{aligned}$$

while

$$|\alpha_t^* \, V_t|^2 \geq \frac{2}{N} - \Lambda \, .$$

Applying Ito formula we thus have for a smooth function D

$$\mathbb{E}(D(|Y_t|)) = \frac{1}{2} \, \mathbb{E} \left(\frac{D'(|Y_t|)}{|Y_t|} \, \left(2 \langle Y_t, \beta_t \rangle + Trace(\alpha_t \, \alpha_t^*) - |\alpha_t^* \, V_t|^2 \right) + \, D''(|Y_t|) |\alpha_t^* V_t|^2 \right)$$

We introduce the natural generalization of (H.κ), namely we assume that

$$\text{for a } \kappa \text{ satisfying (69), } |\sigma(x)-\sigma(y)|^2_{HS}+2\,\langle b(x)-b(y), x-y\rangle \;\leq\; -\kappa(|x-y|)\,|x-y|^2\,. \tag{77}$$

If D is a non decreasing, concave function we thus get, provided $(2/N)-\Lambda > 0$,

$$2\,\mathbb{E}(D(|Y_t|)) \leq \mathbb{E}\left(-\,D'(|Y_t|)\,\kappa(|Y_t|)|Y_t| + D''(|Y_t|)\left(\frac{2}{N}-\Lambda\right)\right)\,.$$

Hence looking carefully at the calculations in [23], we see that, provided $(2/N)-\Lambda>0$, the only thing we have to change in (70) is the definition of φ replacing $1/4$ by the inverse of $(2/N)-\Lambda>0$, all other definitions being unchanged. We have thus obtained

Theorem 13. *Assume that (76) and (77) are satisfied. Assume in addition that $(2/N)-\Lambda>0$. Then defining*

$$\varphi_{min} = e^{-\,\frac{1}{(2/N)-\Lambda}\,\int_0^{R_0} s\kappa^-(s)\,ds}\,,$$

the conclusion of Theorem 11 is still true with $\lambda = \frac{1}{2}\,(\varphi_{min}/R_1^2)$.

All the consequences of Theorem 11 still hold (up to the modifications of the constants), in particular one can extend (H.C.K) to the situation of "convexity at infinity" as in Example 2 (3). Details are left to the reader.

As we already said, the condition $(2/N)-\Lambda>0$ already appears in [35] and ensures that the coupling by reflection is succesfull. Roughly speaking it means that the fluctuations of σ are not too big with respect to the uniform ellipticity bound.

7.4 Gradient commutation property and reflection coupling

It is of course quite disappointing at first glance that the only gradient commutation property that we get using this nice contraction results in W_1 distance, is restricted to Lipschitz function as in (75). Let us see however that we may transfer this to stronger gradient commutation properties in some cases. The main tool is the following lemma on Hölder's type inequality in Wasserstein distance.

Lemma 3. *Suppose that ν and μ are two probability measures on $\mathbb{R}$, then for all $q>1$ and p such that $\frac{1}{p}+\frac{1}{q}=1$, we have*

$$W_2(\nu,\mu) \leq W_1^{\frac{1}{2q}}(\nu,\mu)\,W_{(2-\frac{1}{q})p}^{1-\frac{1}{2q}}(\nu,\mu). \tag{78}$$

Furthermore the result tensorises in the sense, that if for $i = 1, ..., n$, μ_i and ν_i are probability measures on $\mathbb{R}$, we have for some constant $c(n)$

$$W_2(\otimes_1^n \nu_i, \otimes_1^n \mu_i) \le c(n)\, W_1^{\frac{1}{2q}}(\otimes_1^n \nu_i, \otimes_1^n \mu_i)\, W_{(2-\frac{1}{q})p}^{1-\frac{1}{2q}}(\otimes_1^n \nu_i, \otimes_1^n \mu_i).$$

Proof. The proof is indeed quite simple and relies mainly on Hölder's inequality. Indeed, in dimension one the optimal transport plan is the same for every convex cost (see for example Villani [43]), so that there exists a transport plan π such that

$$\begin{aligned} W_2^2(\nu,\mu) &= \int\int |x-y|^2 d\pi \\ &\le \left(\int\int |x-y| d\pi\right)^{1/q} \left(\int\int |x-y|^{(2-\frac{1}{q})p} d\pi\right)^{1/p} \\ &= W_1^{\frac{1}{q}}(\nu,\mu)\, W_{(2-\frac{1}{q})p}^{2-\frac{1}{q}}(\nu,\mu). \end{aligned}$$

The case of product probability measure is deduced using the result in dimension one and the following two direct assertions

$$W_2^2(\otimes_1^n \nu_i, \otimes_1^n \mu_i) = \sum_{i=1}^n W_2^2(\nu_i, \mu_i),$$

and if ν and μ have for *ith* marginal ν_i and μ_i

$$W_p(\nu_i, \mu_i) \le W_p(\nu, \mu).$$

Remark 22. We failed at the present time to get the general version of this lemma, i.e. does there exists a constant c only depending on the dimension such that for two probability measures on $\mathbb{R}^n$, we have

$$W_2(\nu,\mu) \le c(n)\, W_1^{\frac{1}{2q}}(\nu,\mu)\, W_{(2-\frac{1}{q})p}^{1-\frac{1}{2q}}(\nu,\mu)?$$

In fact, as will be seen from our applications, even if c does depend of ν and μ (in a nice way), it would be sufficient to get new gradient commutation property. ♢

We are now in position to prove various gradient commutation properties in non standard cases. For simplicity, we suppose here that the diffusion coefficient is constant, i.e.

$$dX_t = dB_t + b(X_t)dt.$$

Theorem 14. *Let us suppose here that either (X_t) lives in $\mathbb{R}$ or that starting from $X_0 = x \in \mathbb{R}^n$, $X_t = (X_t^1, ..., X_t^n)$ is composed of independent component. Assume moreover that $(H.\kappa)$ is satisfied and that $\kappa(r) \ge -L$ then, with λ*

defined in Theorem 11,

$$W_2(\mathcal{L}(X_t^x), \mathcal{L}(X_t^y)) \leq c(n) \left(\frac{2}{\phi_{min}}\right)^{\frac{1}{2q}} e^{\left[(1-\frac{1}{2q})L - \frac{\lambda}{2q}\right]t} |x-y|$$

so that the weak gradient commutation property holds

$$|\nabla P_t f|^2 \leq c(n) \left(\frac{2}{\phi_{min}}\right)^{\frac{1}{2q}} e^{\left[(1-\frac{1}{2q})L - \frac{\lambda}{2q}\right]t} P_t |\nabla f|^2$$

and thus a local Poincaré inequality holds.

Note that this theorem is the first one to give the commutation gradient property in non strictly convex case with a good behaviour at infinity.

Proof. Using synchronous coupling as previously explained and the fact that $\kappa(r) \geq -L$ we have that

$$W_{(2-\frac{1}{q})p}(\mathcal{L}(X_t^x), \mathcal{L}(X_t^y)) \leq e^{Lt}|x-y|.$$

In the same time, by Theorem 11, we have that

$$W_1(\mathcal{L}(X_t^x), \mathcal{L}(X_t^y)) \leq \frac{2}{\phi_{min}} e^{-\lambda t} |x-y|.$$

We then use Lemma 78 to get the first assertion. The second one is obtained as in Proposition 8.

Example 3. Consider for example the log-concave case $b(x) = -4x^3$ for which Bakry-Emery theory enables us to get that we are in 0-curvature and thus

$$|\nabla P_t f|^2 \leq P_t |\nabla f|^2.$$

However, using Theorem 14 and this last inequality, we easily get that there exists $\lambda > 0$ such that

$$|\nabla P_t f|^2 \leq \min\left(1, \frac{2}{\phi_{min}} e^{-\lambda t}\right) P_t |\nabla f|^2,$$

which is completely new. It captures both the short time behavior equivalent to the Γ_2 0-curvature criterion and the long time behavior for which $P_t f \to \mu(f)$ and thus $\nabla P_t f|$ is expected to decay to 0.
Note that we may extend this example to a double well potential, in the case when the height of the well is not too large.

8 Preserving curvature.

A natural question about curvature is the following: is curvature preserved by a diffusion process ? According to a result by Kolesnikov [32], the Ornstein-Uhlenbeck process is essentially the only one, among diffusion processes, preserving log-concavity (i.e. if ν_0 is log-concave, so is ν_T for all $T > 0$). One may also wonder if ν_t may satisfy other "curvature" like inequality as HWI for example. It would have important applications on local inequalities, indeed transportation inequalities together with a HWI inequality may imply logarithmic Sobolev inequality.

In the spirit of the previous remark, consider a standard Ornstein-Uhlenbeck process $X_.$, i.e. the solution of

$$dX_t = dB_t - \frac{\lambda}{2}\, X_t\, dt\,. \tag{79}$$

The curvature K is thus equal to $\lambda \in \mathbb{R}$. If $\mathcal{L}(X_0) = \nu$, it is known that the law ν_T of X_T is the same as the law of

$$e^{-\lambda T/2}\left(Z + \sqrt{\frac{e^{\lambda T}-1}{\lambda}}\, G\right)\,,$$

where G and Z are independent random variables, G being a standard gaussian variable and Z having distribution ν. Hence

$$C_{LS}(\nu_T) \leq e^{-\lambda T} C_{LS}(\nu) + \frac{2(1-e^{-\lambda T})}{\lambda}\,,$$

or, if we use the notation $C_{LS}(Y) = C_{LS}(\eta)$ for a random variable Y with distribution η,

$$C_{LS}\left(e^{-\lambda T/2}\left(Z + \sqrt{\frac{e^{\lambda T}-1}{\lambda}}\, G\right)\right) \leq e^{-\lambda T} C_{LS}(Z) + \frac{2(1-e^{-\lambda T})}{\lambda}\,. \tag{80}$$

But if A is a random variable it is clear that $C_{LS}(\lambda A) = \lambda^2\, C_{LS}(A)$. It follows

$$C_{LS}\left(Z + \sqrt{\frac{e^{\lambda T}-1}{\lambda}}\, G\right) \leq C_{LS}(Z) + \frac{2(e^{\lambda T}-1)}{\lambda}\,. \tag{81}$$

This is not surprising since a more general result can be obtained directly (extending a similar result for the Poincaré inequality in [6]):

Proposition 18. *Let X and Y be independent random variables and $\lambda \in [0,1]$ then,*

$$C_{LS}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \leq \lambda C_{LS}(X) + (1-\lambda) C_{LS}(Y)\,, \text{ the same holds with } C_P\,.$$

Proof. The first result for C_P is proved in [6] proposition 1. For C_{LS} the proof is very similar. Let f be smooth. Then

$$\mathbb{E}\left((f^2\,\log f^2)(\sqrt{\lambda}X+\sqrt{1-\lambda}Y)\right)\leq$$

$$\begin{aligned}
&\leq \int\left(\int f^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\right)\,\log\left(\int f^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\right)\,dP_Y(y)\\
&\quad+\int\left(C_{LS}(X)\int\lambda\,|\nabla f|^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\right)\,dP_Y(y)\\
&\leq\left(\int f^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\,d\mathbb{P}_Y(y)\right)\,\log\left(\int f^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\,d\mathbb{P}_Y(y)\right)\\
&\quad+C_{LS}(Y)\int\left|\nabla\sqrt{\int f^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)}\right|^2\,d\mathbb{P}_Y(y)\\
&\quad+\lambda\,C_{LS}(X)\int|\nabla f|^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\,dP_Y(y)\\
&\leq\mathbb{E}(f^2(\sqrt{\lambda}X+\sqrt{1-\lambda}Y))\,\log\left(\mathbb{E}(f^2(\sqrt{\lambda}X+\sqrt{1-\lambda}Y))\right)\\
&\quad+\lambda\,C_{LS}(X)\int|\nabla f|^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)\,dP_Y(y)\\
&\quad+(1-\lambda)\,C_{LS}(Y)\int\left(\frac{\int|\nabla f^2|(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)}{2\sqrt{\int f^2(\sqrt{\lambda}x+\sqrt{1-\lambda}y)\,d\mathbb{P}_X(x)}}\right)^2\,d\mathbb{P}_Y(y)\,.
\end{aligned}$$

Since $\nabla f^2=2f\,\nabla f$, we may use Cauchy-Schwarz inequality in order to bound the last term in the latter sum. This yields exactly the desired result.

Remark 23. It is known that any log-concave probability measure satisfies some Poincaré inequality. The result is due to Bobkov [7] (a short proof is contained in [2]). If Z is a log-concave random variable, $Z+\alpha G$ is still log-concave according to the Prekopa-Leindler theorem.

Here is an amusing proof of the consequence of Prekopa's result when one variable is gaussian.

Let X (resp. Y) be a random variable with law $e^{-V(x)}dx$ (resp. a standard gaussian variable). We assume that X and Y are independent. The density of $X+\sqrt{\lambda}\,Y$ is thus given by

$$q(x)=(2\pi\lambda)^{-n/2}\int e^{-V(u)}\,e^{-\frac{|x-u|^2}{2\lambda}}\,du\;=\;(2\pi\lambda)^{-n/2}\,p(x)\,.$$

Let $H(x)$ be the hessian matrix of $\log p$. Then

$$p^2(x)\,\langle \xi, H(x)\,\xi\rangle = p(x)\,\frac{1}{\lambda^2}\,\left(\int \langle \xi, (x-u)\rangle^2\, e^{-V(u)}\, e^{-\frac{|x-u|^2}{2\lambda}}\, du\right) - \frac{1}{\lambda}\,p^2(x)|\xi|^2$$
$$- \frac{1}{\lambda^2}\,\left(\int \langle \xi, (x-u)\rangle\, e^{-V(u)}\, e^{-\frac{|x-u|^2}{2\lambda}}\, du\right)^2 .$$

Hence

$$\langle \xi, H(x)\,\xi\rangle = -\frac{1}{\lambda}\,|\xi|^2 + \frac{1}{\lambda^2}\,\left(\int \langle \xi, (x-u)\rangle^2\, e^{-V(u)}\, e^{-\frac{|x-u|^2}{2\lambda}}\, \frac{du}{p(x)}\right)$$
$$- \frac{1}{\lambda^2}\,\left(\int \langle \xi, (x-u)\rangle\, e^{-V(u)}\, e^{-\frac{|x-u|^2}{2\lambda}}\, \frac{du}{p(x)}\right)^2 .$$

Now assume that V satisfies (H.C.K) for some $K \in \mathbb{R}$. The probability measure

$$e^{-V(u)}\, e^{-\frac{|x-u|^2}{2\lambda}}\, \frac{du}{p(x)}$$

(or if one prefers its potential) satisfies (H.C.$K + (1/\lambda)$). If $K + (1/\lambda) > 0$, it thus satisfies a Poincaré inequality with constant $\lambda/(1+K\lambda)$. Applying this Poincaré inequality to the function $u \mapsto \langle \xi, (x-u)\rangle$, we obtain

$$\langle \xi, H(x)\,\xi\rangle \leq -\,\frac{K}{1+K\lambda}\,|\xi|^2 .$$

Thanks to simple scales we may thus state

Proposition 19. *Let X be a random variable with law $e^{-V(x)}dx$ and Y a standard gaussian variable independent of X. If V satisfies (H.C.K) for $K \in \mathbb{R}$, then for $0 \leq \lambda \leq 1$, the distribution of $\sqrt{\lambda}\,X + \sqrt{1-\lambda}\,Y$ satisfies (H.C.$\frac{K}{\lambda+K(1-\lambda)}$) as soon as $\lambda + K(1-\lambda) > 0$.*
In particular if X is log-concave ($K = 0$) so is $\sqrt{\lambda}\,X + \sqrt{1-\lambda}\,Y$.

References

1. 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.
2. 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.
3. D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré. *J. Func. Anal.*, 254:727–759, 2008.
4. D. Bakry, I. Gentil, and L. Ledoux. On Harnack inequalities and optimal transportation. Preprint, available on ArXiv, 2012.

5. D. Bakry, I. Gentil, and L. Ledoux. Analysis and Geometry of Markov diffusion operators. Book in preparation, 2013.
6. K. Ball, F. Barthe, and A. Naor. Entropy jumps in the presence of a spectral gap. *Duke Math. J.*, 119:41–63, 2003.
7. S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Prob.*, 27(4):1903–1921, 1999.
8. S. G. Bobkov. Spectral gap and concentration for some spherically symmetric probability measures. In *Geometric aspects of functional analysis, Israel Seminar 2000-2001,*, volume 1807 of *Lecture Notes in Math.*, pages 37–43. Springer, Berlin, 2003.
9. S. G. Bobkov, I. Gentil, and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pu. Appli.*, 80(7):669–696, 2001.
10. F. Bolley, I. Gentil, and A. Guillin. Convergence to equilibrium in Wasserstein distance for Fokker-Planck equation. *J. Func. Anal.*, 263(8):2430–2457, 2012.
11. F. Bolley, I. Gentil, and A. Guillin. Uniform convergence to equilibrium for granular media. To appear in Archive for Rational Mechanics and Analysis, 2012.
12. F. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.*, 44(5):867–884, 2010.
13. C. Borell. Diffusion equations and geometric inequalities. *Potential Analysis*, 12:49–71, 2000.
14. P. Cattiaux. A pathwise approach of some classical inequalities. *Potential Analysis*, 20:361–394, 2004.
15. P. Cattiaux. Hypercontractivity for perturbed diffusion semi-groups. *Ann. Fac. des Sc. de Toulouse*, 14(4):609–628, 2005.
16. P. Cattiaux, I. Gentil, and A. Guillin. Weak logarithmic-Sobolev inequalities and entropic convergence. *Probab. Theory Relat. Fields*, 139:563–603, 2007.
17. P. Cattiaux and A. Guillin. On quadratic transportation cost inequalities. *J. Math. Pures Appl.*, 88(4):341–361, 2006.
18. P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Probab. Theory Relat. Fields*, 140:19–40, 2008.
19. P. Cattiaux, A. Guillin, and P. A. Zitt. Poincaré inequalities and hitting times. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 49(1):95–118, 2013.
20. P. Cattiaux and C. Léonard. Minimization of the Kullback information of diffusion processes. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 30(1):83–132, 1994. and correction in *Ann. Inst. Henri Poincaré* vol.31, p.705-707, 1995.
21. J.F. Collet and F. Malrieu. Logarithmic Sobolev inequalities for inhomogeneous semigroups. *ESAIM Probability and Statistics*, 12:492–504, 2008.
22. H. Djellout, A. Guillin, and L. Wu. Transportation cost information inequalities for random dynamical systems and diffusions. *Ann. Prob.*, 334:1025–1028, 2002.
23. A. Eberle. Reflection coupling and Wasserstein contractivity without convexity. *C. R. Acad. Sci. Paris, Ser. I*, 349:1101–1104, 2011.
24. A. Eberle. Couplings, distances and contractivity for diffusion processes revisited . Available on Math. ArXiv 1305.1233 [math.PR], 2013.
25. J. Fontbona and B. Jourdain. A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. Available on Math. ArXiv 1107.3300 [math.PR], 2011.
26. N. Gozlan and C. Léonard. Transport inequalities - a survey. *Markov Processes and Related Fields*, 16:635–736, 2010.
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. A. Guillin and F-Y Wang. Degenerate Fokker-Planck equations: Bismut formula, gradient estimate and Harnack inequality. *J. Differential Equations*, 253(1):20–40, 2012.

29. N. Huet. Isoperimetry for spherically symmetric log-concave probability measures. *Rev. Mat. Iberoamericana*, 27(1):93–122, 2011.
30. N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland, Amsterdam, 2nd edition, 1988.
31. B. Jourdain. Contraction of one-dimensional stochastic differential equations in wasserstein distance. Personnal Communication, 2013.
32. A. V. Kolesnikov. On diffusion semigroups preserving the log-concavity. *J. Funct. Anal.*, 186(1):196–205, 2001.
33. M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités XXXV*,, volume 1755 of *Lecture Notes in Math.*, pages 167–194. Springer, 2001.
34. J. Lehec. Representation formula for the entropy and functional inequalities. To appear in Ann. Inst. Henri Poincaré. Prob. Stat. Available on Math. ArXiv 1006.3028v2 [math.PR], 2010.
35. T. Lindvall and L.C.G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Prob.*, 14:860–872, 1986.
36. F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Process. Appl.*, 95(1):109–132, 2001.
37. F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Func. Anal.*, 173:361–400, 2000.
38. J. Pitman and M. Yor. Bessel processes and infinitely divisible laws. In *Stochastic Integrals,*, volume 851 of *Lecture Notes in Math.*, pages 285–370. Springer, 1980.
39. P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, 2005.
40. M. Röckner and F. Y. Wang. Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups. *J. Funct. Anal.*, 185(2):564–603, 2001.
41. M. Röckner and F. Y. Wang. Log-Harnack inequality for stochastic differential equations in Hilbert spaces and its consequences. *Anal. Quant. Probab. Related Top.*, 13:27–37, 2010.
42. C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.
43. C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
44. M. Von Renesse and K. T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.
45. F. Y. Wang. *Functional inequalities, Markov processes and Spectral theory*. Science Press, Beijing, 2004.
46. F. Y. Wang and T. Zhang. Log-Harnack inequality for mild solutions of SPDEs with strongly multiplicative noise. Available on Math. ArXiv 1210.6416 [math.PR], 2012.
47. Chuang Yi. On the first passage time distribution of an Ornstein-Uhlenbeck process. *Quantitative Finance*, 10(9):957–960, 2010.