Logarithmic Sobolev inequalities for some nonlinear PDE’s

F. Malrieu

Laboratoire de Statistique et Probabilités UMR C5583, Toulouse, France

Abstract

The aim of this paper is to study the behavior of solutions of some nonlinear partial differential equations of McKean-Vlasov type. The main tools used are, on one hand, the logarithmic Sobolev inequality and its connections with the concentration of measure and the transportation inequality with quadratic cost; on the other hand, the propagation of chaos for particle systems in mean field interaction.

Key words: Interacting particle system, logarithmic Sobolev inequality, propagation of chaos, relative entropy, concentration of measure

1 Introduction

Let us introduce the first nonlinear PDE (in \( \mathbb{R}^d \)) we want to study. It is a nonlinear McKean-Vlasov PDE:

\[
\begin{aligned}
\frac{\partial u}{\partial t} &= \text{div} \left[ \nabla u + (u \nabla U + u \nabla (u \ast W)) \right], \\
u(0, \cdot) &= u_0,
\end{aligned}
\]

where \( \ast \) stands for the convolution and \( W \) is convex, even, with polynomial growth and \( U \) is uniformly convex (i.e. \( \text{Hess} \ U(x) \geq \beta \mathbb{I} \) for some \( \beta > 0 \)). The solution of (1) can be interpreted as the law of the stochastic process \( (X_t)_{t \geq 0} \)
solution of
\begin{equation}
\begin{aligned}
    &dX_t = \sqrt{2} dB_t - \nabla U(X_t) \, dt - \nabla W \ast u_t(X_t) \, dt, \\
    &\mathcal{L}(X_t) = u_t(x) \, dx, \\
    &u_0 \text{ smooth}
\end{aligned}
\end{equation}

where $\mathcal{L}(X_t)$ is the law of $X_t$.

**Remark 1.1.** Let us motivate in few words the study of (1). Benedetto et al. (1998) give a physical interpretation when
\[ d = 1, \quad U(x) = \beta \frac{x^2}{2} \quad \text{and} \quad W(x) = \lambda |x|^3. \]

Equation (1) is the homogeneous version of a transport equation in a thermal bath with temperature $1/\beta$ that can be derived from inelastic collisions:
\[ (\partial_t + v \partial_x)u(t, x, v) = \partial^2_{vv} + \partial_v [(\beta v + W' \ast_v u(t, x, v))u(t, x, v)] \]

where $\ast_v$ stands for the convolution with respect to the velocity $v$. This kind of equations are called kinetic equations for granular media.

The probabilistic approach of this type of equations is to replace the nonlinear process by a particle system in mean field interaction i.e. to replace, in Equation (2), $u$ by the empirical measure of the particle system. This leads to the following SDE in $\mathbb{R}^{dN}$:
\begin{equation}
\begin{aligned}
    &dX^{i,N}_t = \sqrt{2} dB^i_t - \nabla U(X^{i,N}_t) \, dt - \frac{1}{N} \sum_{j=1}^{N} \nabla W(X^{i,N}_t - X^{j,N}_t) \, dt \\
    &\mathcal{L}(X^0_N) = u_0 \otimes^N dx \quad \text{for} \quad i = 1, \ldots, N
\end{aligned}
\end{equation}

where $(B^i)_1$ are independent Brownian motions on $\mathbb{R}^d$. As it is expected, the particle system is a “good approximation” of the solution of Equation (2).

The usual way to quantify this intuition is to prove the propagation of chaos property. In Section 3, we show that this property is uniform in time.

**Theorem 1.2.** There exists a $K$ such that, for every $N \geq 1$,
\[ \sup_{t \in \mathbb{R}} \mathbb{E} \left( |X^{1,N}_t - X^{1}_t|^2 \right) \leq \frac{K}{N}, \]

where $(X^1_t)$ is solution of Equation (2) where the Brownian motion $(B)$ is replaced by $(B^1)$.

On the other hand, we present a new approach by the use of logarithmic Sobolev inequalities (Section 2 deals with the results needed in this paper concerning this topic). Along this paper, the strategy is to obtain, via the
Bakry-Emery criterion, logarithmic Sobolev inequalities for the law of the particle systems with constant that do not depend on the size of systems. This provides good estimates in the limit $N \to \infty$. As a consequence, we derive the exponential rate for the convergence to equilibrium for the nonlinear PDE.

**Theorem 1.3.** There exists a constant $K$ such that for every positive $t$,

$$\|u_t - \overline{u}\|_1 \leq Ke^{-\beta t/2}$$

where $\overline{u}$ is the unique solution of

$$\overline{u}(x) = \frac{1}{Z} \exp (-U(x) - W * \overline{u}(x))$$

with $Z = \int \exp (-U(x) - W * \overline{u}(x)) \, dx$.

At last we obtain confidence intervals for the convergence of the empirical measure at time $t$ to $\overline{u}$: for every $r \geq 0$ and every function $f$ from $\mathbb{R}$ to $\mathbb{R}$ with Lipschitz (semi)norm bounded by 1,

$$\mathbb{P}\left( \left| \frac{1}{N} \sum_{i=1}^{N} f(X_t^{i:N}) - \int f(x) \overline{u}(x) \, dx \right| \geq r + \sqrt{\frac{K}{N} + Ke^{-\beta t}} \right) \leq 2 \exp \left[ -\frac{N \beta r^2}{2} \right].$$

These confidence intervals are a straightforward consequence of the phenomenon of the concentration of measure induced by the logarithmic Sobolev inequality.

Section 4 is dedicated to the study of the case $U = 0$ which has been investigated from the probabilistic point of view by Benachour et al. (1998a,b). Let us mention that the trend to equilibrium for granular media is also investigated by Carrillo et al. (2001). They are only interested in the convergence for the nonlinear equation but their results are much more general.

Section 5 deals with the general case of McKean-Vlasov type equations i.e. the nonlinear diffusion is solution of

$$\left\{ \begin{array}{l}
    d\overline{X}_t = \sqrt{2}\sigma(\overline{X}_t, \kappa * u(t, \overline{X}_t)) \, dB_t + b(\overline{X}_t, \kappa * u(t, \overline{X}_t)) \, dt \\
    \mathcal{L}(\overline{X}_t) = u(t, dy) \\
    \overline{X}_{t=0} = X_0.
\end{array} \right. \quad (4)$$

where $\sigma$, $b$ and $\kappa$ are smooth and bounded functions. In this case, the self-stabilizing phenomenon disappears but we still get estimates at time $T$. 

3
2 Some classical and useful results

A probability measure $\mu$ on $\mathbb{R}^n$ satisfies a logarithmic Sobolev inequality with constant $C$ if

$$\text{Ent}_\mu(f^2) \leq C \int |\nabla f|^2 \, d\mu$$

(5)

for all smooth enough functions $f$ where

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

Let us recall two consequences of this property for $\mu$. First, for every $r \geq 0$, and every Lipschitz function $f$ on $\mathbb{R}^n$ (equipped with the Euclidean topology) with

$$\|f\|_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} \leq 1,$$

one has

$$\mu(|f - \mathbb{E}_\mu(f)| \geq r) \leq 2e^{-r^2/C}.$$

This is the so-called concentration phenomenon, see Ledoux (1999).

Secondly, denote by $W_2(\mu, \nu)$ the Wasserstein distance with quadratic cost i.e.

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

where the infimum is running over all probability measure $\pi$ on $\mathbb{R}^n \times \mathbb{R}^n$ with respective marginals $\mu$ and $\nu$. By the Monge-Kantorovitch representation (see Rachev (1991)), we have

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

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

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

for every $x, y \in \mathbb{R}^n$. Following Otto and Villani (2000), if $\mu$ is absolutely continuous and satisfies (5), then, for every probability measure $\nu$ absolutely continuous with respect to $\mu$,

$$W_2(\mu, \nu)^2 \leq \frac{C}{2} \text{H}((\nu \mid \mu))$$
where $H$ is the relative entropy:

$$H(\nu | \mu) = \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu = \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise.} \end{cases} \quad (6)$$

Another approach of this theorem has been studied by Bobkov et al. (2001).

Consider now a diffusion process with infinitesimal generator $L$ and semigroup $(P_t)_t$. As usually, we associate to $L$ its carré du champ $\Gamma$ defined by

$$\Gamma(f, g) = \frac{1}{2} (L(fg) - fLg - gLf)$$

and the operator $\Gamma_2$

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

The diffusion generator $L$ satisfies a curvature-dimension inequality (with curvature $\rho$ and dimension $n$) if for every smooth function $f$,

$$\mathfrak{L}(f) \geq \rho \Gamma(f) + \frac{1}{n} (Lf)^2.$$

In this paper, we only deal with the infinite dimensional case. A general study is performed by Bakry and Emery (1985). The curvature inequality provides logarithmic Sobolev inequalities: following Bakry (1997), for any $\rho \in \mathbb{R}$, the following properties are equivalent

- For every $f$, $\mathfrak{L}(f) \geq \rho \Gamma(f)$
- For every $f$, $\sqrt{\Gamma(P_tf)} \leq e^{-\rho t} P_t \sqrt{\Gamma f}$
- The probability measures $(P_t(\cdot)(x))_x$ satisfies a logarithmic Sobolev inequality with constant

$$C_t = \frac{2}{\rho} \left(1 - e^{-\rho t}\right).$$

Besides, if $\rho$ is positive, the last property remains true for the invariant measure $\mu$ with constant $2/\rho$. This provides an exponential rate for the ergodicity of the semigroup:

$$H(P_t | \mu) \leq K e^{-2\rho t}.$$
3 Dissipative equations

3.1 The model

Consider the nonlinear differential equation (1) in $\mathbb{R}^d$ where $W$ is convex, even, with polynomial growth and $U$ is uniformly convex (i.e. $\text{Hess } U(x) \geq \beta I$ for some $\beta > 0$). The real $\gamma$ is defined to be the largest real (nonnegative) satisfying $\text{Hess } W(x) \geq \gamma I$.

**Remark 3.1.** In order to prove existence and uniqueness for Equation (2), it is possible to use the technics of Benachour et al. (1998a). In particular, notice that it can be rewritten as

$$
\begin{align*}
\frac{dX_t}{dt} &= \sqrt{2} dB_t - \nabla U(X_t) dt - \nabla W_t(X_t) dt, \\
\nabla W_t(x) &= E[\nabla W(x - X_t)], \\
u_0 &\text{ smooth.}
\end{align*}
\tag{7}
$$

One can prove first existence of $u_t$ by a fixed point argument using the Wasserstein distance. Then, once $u_t$ is known, strong uniqueness for $(X_t)$ holds.

The associated particle system is solution of

$$
\begin{align*}
\frac{dX_t^{i,N}}{dt} &= \sqrt{2} dB_t^i - \nabla U(X_t^{i,N}) dt - \frac{1}{N} \sum_{j=1}^{N} \nabla W(X_t^{i,N} - X_t^{j,N}) dt \\
\mathcal{L}(X_0^N) &= u_0^\otimes_N dx \quad \text{for } i = 1, \ldots, N
\end{align*}
\tag{8}
$$

**Remark 3.2.** In what follows, $(\overline{X}_t)$ will stand for the solution of (2) where the Brownian motion $(B)$ is replaced by $(B^i)$.

3.2 Uniform propagation of chaos

**Theorem 3.3.** There exists a $K$ such that, for every $N \geq 1$,

$$
\sup_{t \in \mathbb{R}} \mathbb{E}\left(\left|X_t^{1,N} - \overline{X}_t\right|^2\right) \leq \frac{K}{N}
\tag{9}
$$

and

$$
\mathbb{E}\left(\sup_{0 \leq t \leq T} \left|X_t^{1,N} - \overline{X}_t\right|^2\right) \leq K \frac{T}{N}.
$$

**Remark 3.4.** The constant $K$ depends on $\beta$, $\gamma$ and on the $p$-th moment of $u_0$ for a power $p$ that controls the growth of $U$ and $W$. 

6
Proof of Theorem 3.3. The proof is quite similar to the one in Benachour et al. (1998a). Nevertheless the term “$\nabla U$” provides a better estimate in time. For $i = 1, \ldots, N$,

$$X_t^{i,N} - X_t^i = X_s^{i,N} - X_s^i - \int_s^t \left( \nabla U(X_r^{i,N}) - \nabla U(X_r^i) \right) dr$$

$$- \frac{1}{N} \sum_{j=1}^N \int_s^t \left( \nabla W(X_r^{i,N} - X_r^j) - \nabla W * u_r(X_r^i) \right) dr.$$

By Itô’s formula,

$$\sum_{i=1}^N \left| X_t^{i,N} - X_t^i \right|^2 = \sum_{i=1}^N \left| X_s^{i,N} - X_s^i \right|^2$$

$$-2 \sum_{i=1}^N \int_s^t \left( X_t^{i,N} - X_t^i \right) \cdot \left( \nabla U(X_r^{i,N}) - \nabla U(X_r^i) \right) dr$$

$$- \frac{2}{N} \sum_{i,j=1}^N \int_s^t \rho_{ij}^{(1)}(r) dr \quad (10)$$

where

$$\rho_{ij}^{(1)}(r) = \left( X_t^{i,N} - X_t^i \right) \cdot \left( \nabla W(X_r^{i,N} - X_r^j) - \nabla W * u_r(X_r^i) \right)$$

$$= \left( X_t^{i,N} - X_t^i \right) \cdot \left( \nabla W(X_r^{i,N} - X_r^j) - \nabla W(X_r^i - X_r^j) \right)$$

$$+ \left( X_t^{i,N} - X_t^i \right) \cdot \left( \nabla W(X_r^i - X_r^j) - \nabla W * u_r(X_r^i) \right)$$

$$= \rho_{ij}^{(2)}(r) + \rho_{ij}^{(3)}(r).$$

The vector field $\nabla W$ is odd and satisfies

$$\left( \nabla W(x) - \nabla W(y) \right) \cdot (x - y) \geq 0$$

then, by definition of $\rho_{ij}^{(2)}(r)$,

$$\rho_{ij}^{(4)} = \rho_{ij}^{(2)}(r) + \rho_{ji}^{(2)}(r)$$

$$= \left[ X_t^{i,N} - X_t^i - (X_t^j - X_t^j) \right] \cdot \left[ \nabla W(X_r^{i,N} - X_r^j) - \nabla W(X_r^i - X_r^j) \right] \geq 0.$$

It has been shown that

$$\sum_{i,j=1}^N \rho_{ij}^{(2)}(r) = \sum_{1 \leq i < j \leq N} \rho_{ij}^{(4)}(r) \geq 0.$$
On the other hand, Cauchy-Schwarz inequality leads to

\[-E \left[ \sum_{j=1}^{N} \rho_{ij}^{(3)}(r) \right] \leq -E \left[ (X_{i}^{i,N} - \overline{X}_{i}^{i}) \cdot \left( \sum_{j=1}^{N} (\nabla W(X_{i}^{i} - \overline{X}_{j}^{i}) - \nabla W \ast u_{r}(\overline{X}_{j}^{i})) \right) \right] \]

\[\leq \left( E \left[ |X_{i}^{i,N} - \overline{X}_{i}^{i}|^2 \right] \right)^{\frac{1}{2}} \theta_{i}(r)^{\frac{1}{2}} \]

where

\[\theta_{i}(r) = E \left( \left| \sum_{j=1}^{N} (\nabla W(X_{i}^{i} - \overline{X}_{j}^{i}) - \nabla W \ast u_{r}(\overline{X}_{j}^{i})) \right|^2 \right) \]

Then, we get

\[\theta_{i}(r) = \sum_{j=1}^{N} E(\xi_{j}(r)^2) + 2 \sum_{1 \leq j < k \leq N} E(\xi_{j}(r) \cdot \xi_{k}(r)),\]

with the obvious notation

\[\xi_{j}(r) = \nabla W(X_{i}^{i} - \overline{X}_{j}^{i}) - \nabla W \ast u_{r}(\overline{X}_{j}^{i})\]

If \(j\) is not equal to \(k\), one of them is not equal to \(i\) and then,

\[E(\xi_{j}(r) \cdot \xi_{k}(r)) = 0\] if \(j \neq k\)

since the random variables \(\overline{X}_{r}\) are independent copies of \(\overline{X}_{r}^{i}\) with density \(u_{r}\).

At last,

\[E(\xi_{j}(r)^2) = E\left( |\nabla W(X_{i}^{i} - \overline{X}_{j}^{i}) - \nabla W \ast u_{r}(\overline{X}_{j}^{i})|^2 \right) \]

\[\leq E\left( |\nabla W(X_{i}^{i} - \overline{X}_{j}^{i})|^2 \right) \]

\[\leq K E\left( |\overline{X}_{r}^{i}|^{2p} \right) + K E\left( |\overline{X}_{r}^{j}|^{2p} \right) \]

\[\leq 2K M_{2p}.\]

We have established

\[-E \left[ \sum_{j=1}^{N} \rho_{ij}^{(3)}(r) \right] \leq \sqrt{KN} \alpha(r)^{1/2}\]

where \(\alpha\) is defined by

\[\alpha(t) = E \left[ (X_{i}^{i,N} - \overline{X}_{i}^{i})^2 \right]. \quad (11)\]
Let take the expectation of (10). Using the exchangeability of the marginals of the particle system, it comes

$$\alpha(t) \leq \alpha(s) - 2\beta \int_s^t \alpha(r) \, dr + \frac{2\lambda K}{\sqrt{N}} \int_s^t \alpha(r)^{1/2} \, dr. \quad (12)$$

This means that

$$\alpha'(t) \leq -2\beta \alpha(t) + \frac{2\lambda K}{\sqrt{N}} \alpha(t)^{1/2}.$$ 

Gronwall lemma implies (since $\alpha(0)$ is 0) that

$$\alpha(t)^{1/2} \leq \frac{\lambda K}{\beta \sqrt{N}} [1 - e^{-\beta t}]$$

which is (9).

The second estimate in theorem 3.3 follows classically from (9) (see Benachour et al. (1998a)).

**3.3 Curvature of the particle systems**

The main observation is that the infimum (over $N$) of the curvature of the particle system is positive.

**Lemma 3.5.** The curvature of the particle system of size $N$ is bounded below by $\beta + \gamma/N$.

**Proof.** We first study the infinitesimal generator $\mathbf{L}^N$ of the $N$-particle system which is defined by

$$\mathbf{L}^N F = \Delta F - \nabla \Psi \cdot \nabla F$$

where

$$(\nabla \Psi)_i(x) = \nabla U(x_i) + \frac{1}{N} \sum_{j=1}^N \nabla W(x_i - x_j).$$

for $i = 1, \ldots, N$.

**Remark 3.6.** The bold symbols $\Delta$ and $\nabla$ stand for differential operators on $\mathbb{R}^{dN}$ where $\Delta$ and $\nabla$ respectively are the Laplacian and the gradient operators on $\mathbb{R}^d$.

In order to apply the $\mathbf{L}_2$ criterion, we have to show that

$$\text{Hess } \Psi \geq \left( \beta + \frac{\gamma}{N} \right) I_{dN}.$$
Reasoning with blocks of size $d \times d$, one can show that
\[
(Hess \, \Psi)_{ii} = Hess \, U(x_i) + \frac{1}{N} \sum_{j=1}^{N} Hess \, W(x_i - x_j) \quad \text{for } i = 1, \ldots, N
\]
and
\[
(Hess \, \Psi)_{ij} = -\frac{1}{N} Hess \, W(x_i - x_j) \leq 0 \quad \text{for } i, j = 1, \ldots, N \text{ and } i \neq j.
\]
The result follows from the simple observation that the diagonal of $Hess \, \Psi$ is strictly dominating. In a more precise way,
\[
(Hess \, \Psi)_{ii} = \beta + \frac{\gamma}{N} - \sum_{j \neq i} (Hess \, \Psi)_{ij} \quad \text{for } i = 1, \ldots, N
\]
implies that the eigenvalues of $Hess \, \Psi$ are greater or equal than $\beta + \gamma/N$. One can notice that $\beta + \gamma/N$ is an eigenvalues of $Hess \, \Psi$ with eigenvector $(1, \ldots, 1)$.

This ensures that the semigroup of the $N$-particle system at time $t$ (i.e. the law of $X_t^N$ starting at $x \in \mathbb{R}^{dN}$) satisfies a logarithmic Sobolev inequality with constant
\[
C_t = \frac{2}{\beta + \gamma/N} \left(1 - e^{-2(\beta + \gamma/N)t}\right).
\]
One can formulate an analogous result for more general initial conditions.

**Corollary 3.7.** Let $P_t$ be a diffusion semigroup with curvature $\rho$ and a probability measure $u_0(dy)$ satisfying a logarithmic Sobolev inequality with constant $c_0$. Then the law of $X_t$ (with $L(X_0) = u_0$) satisfies a logarithmic Sobolev inequality with constant
\[
D_t = \frac{2}{\rho} \left(1 - e^{-2\rho t}\right) + c_0 e^{-2\rho t}.
\]

**Proof.** Let $f$ by a smooth function. It is clear that
\[
\text{Ent}_{P_t(x)du_0}(f^2) = \int \text{Ent}_{P_t(x)}(f^2)du_0(x) + \text{Ent}_{u_0}(P_t(f^2)).
\]
The logarithmic Sobolev inequality for the semigroup leads to
\[
\int \text{Ent}_{P_t(x)}(f^2)du_0(x) \leq \frac{2}{\rho} \left(1 - e^{-2\rho t}\right) \int P_t(\Gamma(f))(x)du_0(x).
\]
On the other hand, we have too
\[
\text{Ent}_{u_0}(P_t(f^2)) \leq \frac{c_0}{4} \int \frac{\Gamma P_t(f^2)(x)}{P_t(f^2)(x)} \, du_0(x).
\]
Now,

$$\sqrt{\Gamma P_t(f^2)} \leq e^{-\rho t} P_t \sqrt{\Gamma(f^2)} = 2e^{-\rho t} P_t \left( f \sqrt{\Gamma f} \right)$$

$$\leq 2e^{-\rho t} \left( P_t(f^2) \right)^{1/2} \left( P_t(\Gamma f) \right)^{1/2}$$

by the commutation relation and Cauchy-Schwarz inequality.

As a consequence, the law of $X_t^N$ is concentrated around its mean as a Gaussian variable with variance $D_t/2$. This provides confidence intervals for the convergence of the empirical measure: for all $g$ from $\mathbb{R}^d$ in $\mathbb{R}$ with Lipschitz norm less than 1, we have for every $r \geq 0$

$$\mathbb{P} \left( \left| \frac{1}{N} \sum_{i=1}^N g(X_{t_i}^N) - \mathbb{E}(g(X_{t_i}^N)) \right| \geq r \right) \leq 2 \exp \left( -\frac{N r^2}{D_t} \right).$$

This is just a consequence of the following remark.

**Remark 3.8.** If the Lipschitz norm of $g$ from $\mathbb{R}^d$ in $\mathbb{R}$ is bounded by 1, then the Lipschitz norm of

$$G(x) = \frac{1}{N} \sum_{i=1}^N g(x_i)$$

from $\mathbb{R}^d$ in $\mathbb{R}$ is bounded by $1/\sqrt{N}$.

**Corollary 3.9 (Concentration).** Suppose that $u_0(dy)$ satisfies a logarithmic Sobolev inequality with constant $c_0$. For all $g$ from $\mathbb{R}^d$ in $\mathbb{R}$ with Lipschitz norm less than 1, we have

$$\mathbb{P} \left( \left| \frac{1}{N} \sum_{i=1}^N g(X_{t_i}^N) - \int g(y)u(t,y) \, dy \right| \geq r + \sqrt{\frac{K}{N}} \right) \leq 2 \exp \left( -\frac{N r^2}{D_t} \right).$$

for $t > 0$, $N \in \mathbb{N}$ and $r \geq 0$.

**Remark 3.10.** It is also possible to establish the concentration property assuming that $u_0$ is compactly supported (even if $u_0$ does not satisfy a logarithmic Sobolev inequality).

### 3.4 Curvature of the nonlinear semigroup

Keeping in mind that, once its law at time $t$ is known for all $t$, the nonlinear process is an inhomogeneous Markov process with generator

$$\overline{L}_s f = \Delta f - [\nabla U + \nabla W \ast u_s] \cdot \nabla f ,$$

it is possible to establish a logarithmic Sobolev inequality for the nonlinear semigroup. Let us sketch the adaptation to the inhomogeneous case of the
proof presented in Bakry (1997). First, define the operator carré du champ at time \( s \) by
\[
\Gamma(s)(f, g) = \frac{1}{2}[\mathcal{L}_s(fg) - f\mathcal{L}_s g - g\mathcal{L}_s f]
\]
and the operator \( \overline{\Gamma}_2(s) \) by
\[
\overline{\Gamma}_2(s)(f, g) = \frac{1}{2}[\mathcal{L}_s(\Gamma(s)(f, g)) - \Gamma(s)(f, \mathcal{L}_s g) - \Gamma(s)(g, \mathcal{L}_s f)].
\]

**Remark 3.11.** For simplicity, \( \Gamma(f) \), respectively \( \overline{\Gamma}_2(f) \), stands for \( \Gamma(f, f) \), respectively \( \overline{\Gamma}_2(f, f) \).

In what follows, we assume that \( \Gamma \) does not depend on time \( s \). This is implied by the fact that it is true for the diffusive coefficients of \( \mathcal{L}_s \) and it is true for the granular media equation.

We will say that \( (\mathcal{L}_s)_s \) satisfies the Bakry-Emery criterion if for every \( s \),
\[
\overline{\Gamma}_2(s)(f) \geq \rho \Gamma(f).
\] (13)

In particular, for the solution of (2), a straightforward computation leads to
\[
\Gamma(f) = |\nabla f|^2,
\]
\[
\overline{\Gamma}_2(s)(f) = \|\text{Hess } f\|_2^2 + (\text{Hess } U)(\nabla f, \nabla f) + (\text{Hess } W * u_s)(\nabla f, \nabla f),
\]
where \( (\text{Hess } U)(X, Y) \) stands for the bilinear form associated to the Hessian matrix of \( U \). Then, (13) is true for \( \beta + \gamma \).

The consequence of the curvature inequality (13) is the same as in the homogeneous case.

**Proposition 3.12.** The following properties are equivalent
(i) For every \( s \) and \( f \), \( \overline{\Gamma}_2(s)(f) \geq \rho \Gamma(f) \)
(ii) For every \( u, t \) and \( f \), \( \sqrt{\Gamma(P_{u,t}f)} \leq e^{-\rho(t-u)}P_{u,t}\sqrt{\Gamma f} \)
(iii) The probability measures \( (P_{u,t}(\cdot))(x) \) satisfy a logarithmic Sobolev inequality with constant
\[
\overline{C}_{u,t} = \overline{C}_{t-u} = \frac{2}{\rho}e^{-2\rho(t-u)}.
\]

We are now able to improve the propagation of result (9). Let us define \( u_t^{(k,N)} \) as the density of the law of \( k \) particles among \( N \) at time \( t \) and \( u_t^{(N)} \) will stands for \( u_t^{(N,N)} \).

**Proposition 3.13.** For \( k \leq N \), we have
\[
\sup_{t \geq 0} \|u_t^{(k,N)} - u_t^{\otimes k}\|_1 \leq K \sqrt{\frac{k}{N}}.
\]
Proof. The first point is to obtain
\[
\|u_t^{(k,N)} - u_t^{\otimes k}\|_1 \leq 2\sqrt{\frac{k}{N}} \sqrt{H(u_t^{(N)} | u_t^{\otimes N})}.
\]
This easily follows from the classical relation established by Csiszár and Kullback (see Pinsker (1964))
\[
\|\mu - \nu\|_{TV} \leq \sqrt{2H(\mu | \nu)} \tag{14}
\]
(where \(\|\cdot\|_{TV}\) stands for the total variation norm) and the following one due to Csiszár (1984).

**Lemma 3.14.** Let \((E, \mathcal{E})\) a measurable space and \(\mu^{(N)}\) be an exchangeable probability measure on a product space \(E^N\) such that \(\mu^{(N)} \ll \nu^{\otimes N}\) for some probability measure \(\nu\) on \(E\). Denote by \(\mu^{(k,N)}\) the marginal on the first \(k\) coordinates then, for every \(k \leq N\), we have
\[
H(\mu^{(k,N)} | \nu^{\otimes k}) \leq \frac{2k}{N} H(\mu^{(N)} | \nu^{\otimes N}) \tag{15}
\]
Now let us study the behavior of \(F\) defined by
\[
F(t) = H(u_t^{(N)} | u_t^{\otimes N}).
\]

**Lemma 3.15.** The derivative of \(F\) is given by
\[
F'(t) = -\int \left| \frac{\nabla u_t^{(N)}}{u_t^{(N)}} - \frac{\nabla u_t^{\otimes N}}{u_t^{\otimes N}} \right|^2 u_t^{(N)} - \int \left[ \frac{\nabla u_t^{(N)}}{u_t^{(N)}} - \frac{\nabla u_t^{\otimes N}}{u_t^{\otimes N}} \right] [\nabla \Psi_N - \nabla \Psi_t] u_t^{(N)}
\]
where
\[
(\nabla \Psi_N)_i = \nabla U(x_i) + \frac{1}{N} \sum_{j=1}^N \nabla W(x_i - x_j)
\]
and
\[
(\nabla \Psi_t)_i = \nabla U(x_i) + \nabla W \ast u_t(x_i).
\]
The proof of this result follows from a straightforward computation and an integration by parts. By the basic property of any bilinear form:
\[
2\langle a, b \rangle \leq \langle a, a \rangle + \langle b, b \rangle
\]
we get
\[
F'(t) \leq -\frac{1}{2} \int \left| \frac{\nabla u_t^{(N)}}{u_t^{(N)}} - \frac{\nabla u_t^{\otimes N}}{u_t^{\otimes N}} \right|^2 u_t^{(N)} + \frac{1}{2} \int |\nabla \Psi_N - \nabla \Psi_t|^2 u_t^{(N)}.
\]
At last, \( u_t \) satisfies a logarithmic Sobolev inequality with constant \( C_t \). By tensorisation (see Gross (1975)), it is still true (with the same constant) for \( u_{t}^{\otimes N} \). As a consequence, we obtain

\[
F(t) \leq \frac{2}{\beta + \gamma} \int \left| \frac{\nabla u_t^{(N)}}{u_t^{(N)}} - \frac{\nabla u_{t}^{\otimes N}}{u_{t}^{\otimes N}} \right|^2 u_t^{(N)}.
\]

Then the function \( F \) satisfies the following differential inequation:

\[
F'(t) \leq -\frac{\beta}{4} F(t) + \frac{1}{2} \int |\nabla \Psi_N - \nabla \Psi_t|^2 u_t^{(N)}.
\]

By the result (9) and the uniform control of the moments of \( u_s \) for \( s \) in \([0, +\infty)\), we get

\[
\int |\nabla \Psi_N - \nabla \Psi_t|^2 u_t^{(N)} \leq K.
\]

Then one has to use the polynomial growth of \( U \) and \( W \) and the propagation of chaos inequality to obtain

\[
F(t) \leq K
\]

which achieves the proof. \( \square \)

### 3.5 Ergodicity

A straightforward adaptation of Theorem 2.2 in Benedetto et al. (1998) leads to the following qualitative result.

**Theorem 3.16.** There exists a unique \( \overline{u} \) such that for any \( u_0 \) we have

\[
\lim_{t \to \infty} \| u_t - \overline{u} \|_1 = 0.
\]

The best way to characterize the stationary measure \( \overline{u} \) is to see it as the unique minimum of the free energy functional defined by:

\[
\eta(f) = \int f(x) \log f(x) \, dx + \int U(x)f(x) \, dx + \frac{1}{2} \int \int W(x-y)f(x)f(y) \, dx \, dy.
\]

It is now possible to give an exponential rate of convergence in Theorem 3.16. First, we need a preliminary proposition which deals with the convergence to equilibrium for the particle system.

**Proposition 3.17.** For all \( u_0 \), there exists a constant \( K \) such that for all \( N \) in \( \mathbb{N} \) and \( t \geq 0 \),

\[
H(u_t^{(N)} \mid \mu_N) \leq K N e^{-\beta t/2}
\]

where \( \mu_N \) is the probability measure with density

\[
\frac{1}{Z_N} \exp \left( -\sum_{i=1}^{N} U(x_i) - \frac{1}{2N} \sum_{i,j=1}^{N} W(x_i - x_j) \right).
\]

14
Proof. It is well-known (see Bakry (1994)) that
\[ H(u^{(N)}_t \mid \mu_N) \leq H(u^{\otimes N}_0 \mid \mu_N) e^{-\beta t/2}. \]

On the other hand, using the explicit expression (17), we have
\[
H(u^{\otimes N}_0 \mid \mu_N) = \int u^{\otimes N}_0 \log f^{\otimes N}_0 - \int u^{\otimes N}_0 \log \mu_N \\
= N \int u_0 \log u_0 + \int u^{\otimes N}_0 \Psi_N + \log Z_N \\
\leq K N + K \sum_{i=1}^N \int u^{\otimes N}_0 P(|x_1|) \\
+ N \log \left( \int e^{-U(x_1)} \, dx_1 \right),
\]
where \( P \) is a polynomial that dominates the growth of \( U \) and \( W \). Then the result follows. \( \Box \)

Denote by \( \mu_{1,N} \) the first marginal of \( \mu_N \). Then, for every integer \( N \), we have by Propositions 3.13 and 3.17,
\[
\|u_t - \pi\|_1 \leq \|u_t - u^{(1,N)}_t\|_1 + \|u^{(1,N)}_t - \mu_{1,N}\|_1 + \|\mu_{1,N} - \pi\|_1 \\
\leq \frac{K}{\sqrt{N}} + K \sqrt{N} e^{-\beta t}.
\]
Taking \( N \) of the order of \( e^{\beta t/2} \), we have established the following result.
**Theorem 3.18.** There exists a constant \( K \) (depending on \( c \)) such that for every positive \( t \) and \( u_0 \) in \( P_c \),
\[ \|u_t - \pi\|_1 \leq K e^{-\beta t/2}. \]

**Remark 3.19.** This result is a somehow disappointing since the rate of convergence for the nonlinear process is not the one of the particle system.

Let us show that the hope in Remark 3.19 is natural in some sense.
**Lemma 3.20.** Let \( \mu \) and \( \nu \) be two probability measures on \( \mathbb{R}^n \). Then, for every probability measures \( \mu_N \) and \( \nu_N \) (on \( \mathbb{R}^{Nn} \)) with respective marginals \( \mu, \ldots, \mu \) and \( \nu, \ldots, \nu \),
\[ W_2(\mu_N, \nu_N)^2 \geq N \, W_2(\mu, \nu)^2. \]

This lemma implies that
\[ W_2(u^{(1,N)}_t, \mu_{1,N}) \leq \frac{1}{\sqrt{N}} W_2(u^{(N)}_t, \mu_N). \]
Besides, as $\mu_N$ satisfies a logarithmic Sobolev inequality with constant $2/\beta$, $\mu_N$ satisfies a transportation inequality for the quadratic cost: for every measure $\nu_N$ on $\mathbb{R}^{dN}$,  
\[ W_2(\nu_N, \mu_N)^2 \leq \frac{1}{\beta} H(\nu_N | \mu_N). \]
Then we get
\[
W_2(u, \pi) \leq W_2(u, u^{(1,N)}) + W_2(u^{(1,N)}, \mu_{1,N}) + W_2(\mu_{1,N}, \pi)
\leq 2 \sup_{t \geq 0} \left( \mathbb{E} \left[ |X_t^{1,N} - \mu_t|^2 \right] \right)^{1/2} + \sqrt{\frac{K}{N}} H(u^{(N)} | \mu_N)
\leq \frac{K}{\sqrt{N}} + K e^{-\beta t}.
\]
Let $N$ goes to infinity to get the following result.

**Proposition 3.21.**
\[ W_2(u, \pi) \leq K e^{-\beta t}. \]
This is the same rate as in the linear case. This result provides efficient exact confidence intervals for the convergence of the empirical measure at time $t$ to the invariant measure $\pi$.

**Proposition 3.22.** For every $r \geq 0$,
\[
P \left( \frac{1}{N} \sum_{i=1}^{N} f(X_i^{t,N}) - \int f(x) \pi(x) \, dx \geq r + \sqrt{\frac{K}{N}} + K e^{-\beta t} \right) \leq 2 \exp \left[ - \frac{Nt^2}{D_t} \right].
\]

### 4 Self-stabilizing process with non negative curvature

In this section, we study the case where $U = 0$ in Equation (2). This is the equation studied by Benachour et al. (1998a,b). Let us describe quickly the results they have established. There is propagation of chaos and more precisely, there exists a $K > 0$ such that
\[
\sup_{0 \leq s \leq T} \mathbb{E} \left[ |X_s^{t,N} - \mu_s|^2 \right] \leq K \frac{T^2}{N}. \tag{18}
\]
They also establish the propagation of chaos at the level of processes.
\[
\mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_s^{t,N} - \mu_s|^2 \right] \leq \frac{Kt}{N}.
\]

**Remark 4.1.** Notice that, once again, the dependence with respect to time in Equation (18) is of the order of the square of the log-Sobolev constant at time $t$. 

16
4.1 The Ricci curvature of the particle system

We first study the $N$-particle system. Its curvature is bounded below by $\gamma/N$ (which is nonnegative): there is no positive uniform (in $N$) lower bound. Nevertheless, we still have a result at time $t$. In particular, we have that, for every $T$, $N$ and $r \geq 0$,}

$$
\sup_{t \leq T} \mathbb{P} \left( \left| \frac{1}{N} \sum_{i=1}^{N} f(X_{t}^{i,N}) - \mathbb{E}[f(\bar{X}_{t})] \right| \geq r + t \sqrt{\frac{K}{N}} \right) \leq 2 \exp \left[ -\frac{N r^2}{C_{T}(N)} \right]
$$

where $K$ is the constant in (18) and $C_{T}(N)$ is the log-Sobolev constant of the particle system of size $N$:

$$
C_{T}(N) = \frac{2N}{\gamma} (1 - \exp(-\gamma T/N)) \leq 4T.
$$

We have now to present another result established by Benachour et al. (1998b). They show that there exists a constant $K$ such that for every function $f$ of uniform norm and of Lipschitz seminorm bounded by 1,

$$
\left| \mathbb{E}[f(\bar{X}_{t})] - \int f(y) u(y) \, dy \right| \leq \frac{K}{(1+t)^2}.
$$

Then, we have a concentration inequality.

**Theorem 4.2.** For every $r \geq 0$,

$$
\mathbb{P} \left( \left| \frac{1}{N} \sum_{i=1}^{N} f(X_{t}^{i,N}) - \int f(y) u(y) \, dy \right| \geq r + t \sqrt{\frac{K}{N}} + \frac{K}{(1+T)^2} \right) \leq 2 \exp \left[ -\frac{N r^2}{4T} \right]
$$

where $u$ is the solution

$$
\overline{u}(x) = \frac{1}{Z} \exp (-W \ast \overline{u}(x))
$$

with $Z = \int \exp (-W \ast \overline{u}(x)) \, dx$.

4.2 Curvature of the nonlinear process

As in section 3.4, we can associate to the nonlinear process a curvature which is bounded below by $\gamma > 0$. Then, we still have the propagation of chaos for
blocks of size $o(N)$.

**Proposition 4.3.** For $k \leq N$, we have

$$\sup_{t \geq 0} \| u_t^{(k,N)} - u_t^{\otimes k} \|_1 \leq K \sqrt{\frac{k}{N}}.$$  

5 The general case

This part is dedicated to the study of the generic model of nonlinear diffusions that can be approximated by particles in mean field interaction. The beginning consists in a straightforward adjustment of the classical method presented by Sznitman (1991).

5.1 The model

Let us start with Borel-measurable functions $b_i(x, y), \sigma_{ij}(x, y), 1 \leq i, j \leq d,$ from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^d$ and $\kappa$ from $\mathbb{R}$ to $\mathbb{R}$. Let $(B_t)_{t \geq 0}$ be a $d$-dimensional Brownian motion. We study the following equation:

$$\begin{aligned}
    dX_t &= \sqrt{2}\sigma(X_t, \kappa * u(t, X_t)) dB_t + b(X_t, \kappa * u(t, X_t)) dt \\
    L(X_t) &= u(t, dy) \\
    X_{t=0} &= X_0.
\end{aligned}$$

The density of a solution at time $t$ is known to be a weak solution of:

$$\frac{\partial}{\partial t} u(s, x) = \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(x, \kappa * u(s, x)) u(s, x)] - \sum_{i=1}^{d} \frac{\partial}{\partial x_i} [b_i(x, \kappa * u(s, x)) u(s, x)],$$

where $a$ is equal to $\sigma \sigma^*$.

**Theorem 5.1.** If $\sigma$, $b$ and $\kappa$ are bounded and globally Lipschitz functions, strong existence and uniqueness hold for equation (19).
5.2 Convergence of the associated particle system

Let \((B^i)_{i \in \mathbb{N}}\) be independent Brownian motions in \(\mathbb{R}^d\). The interacting particle system associated to (19) is the solution of

\[
\begin{cases}
  dX^{i,N}_t = \sqrt{2} \sigma(X^{i,N}_s, \kappa \ast \Pi^N_s(X^{i,N}_s)) \, dB^i_t \\
  + b(X^{i,N}_s, \kappa \ast \Pi^N_s(X^{i,N}_s)) \, ds \text{ for } i = 1, \ldots, N \\
  X^{i,N}_0 = X^i_0
\end{cases}
\]

(20)

where \(\Pi^N_t\) is the empirical measure of the system at time \(t\) i.e. it is equal to

\[
\Pi^N_t = \frac{1}{N} \sum_{k=1}^N \delta_{X^{k,N}_t}.
\]

How it is expected, the propagation of chaos holds. Indeed we have the following estimate:

**Theorem 5.2.** If \(\sigma\) and \(b\) are bounded and globally Lipschitz, for all \(T \in \mathbb{R}^+\), there exists a \(K_T \in \mathbb{R}\) such that for all \(i\) and \(N\) in \(\mathbb{N}\),

\[
E \left[ \sup_{t \leq T} \left| X^{i,N}_t - X^i_t \right|^2 \right] \leq \frac{K_T}{N}.
\]

(21)

**Remark 5.3.** It can be shown first that for some \(K'_T\),

\[
\sup_{t \leq T} \left( E \left[ \left| X^{i,N}_t - X^i_t \right|^2 \right] \right)^{1/2} \leq \frac{K'_T}{\sqrt{N}}.
\]

(22)

Once again, \(K'_T\) is of the order of \(\exp(KT)\) like the Log-Sobolev constant \(C_T\) of the particle system at time \(T\) (see section 5.3).

5.2.1 An improvement in a special case

In this section, we will suppose that \(\sigma\) is equal to \(I\). Then, it is still true that the nonlinear process has a bounded-below (in \(\mathbb{R}\)) curvature. The its law at time \(t\) satisfies a logarithmic Sobolev inequality. This ensures that, for \(k \leq N\), we have

\[
\sup_{t \leq T} \left\| u^{(k,N)}_t - u^{(k)}_t \right\|_1 \leq K_T \sqrt{\frac{k}{N}}.
\]

In fact, it is possible to get a best estimate (at the level of processes). Following a suggestion of Professor Del Moral we use a method developed by Ben Arous and Zeitouni (1999) in an abstract framework. Nevertheless our case is simpler because we already have established Theorem 5.2. Let us start with few notations:
• $\mathbb{P}_T^N$ stands for the law of $(X_t^N)_{0 \leq t \leq T} = \left((X_t^{i,N}, ..., X_t^{k,N})\right)_{0 \leq t \leq T}$ the particle system until time $T$
• $\mathbb{P}_T^{k,N}$ denotes the law of $\left((X_t^{i,N}, ..., X_t^{k,N})\right)_{0 \leq t \leq T}$. By exchangeability, the $k$-marginals of $\mathbb{P}_T^N$ do not depend on the choice of coordinates,
• let $\overline{P}_T$ be the law until time $T$ of the nonlinear process solution of (19).

The key point is that we are able to show that the relative entropy $H\left(\mathbb{P}_T^N \mid \mathbb{P}_T^{\otimes N}\right)$ is bounded in $N$.

**Proposition 5.4.** There exists a constant $K_T$ such that for all $N$,

$$H\left(\mathbb{P}_T^N \mid \mathbb{P}_T^{\otimes N}\right) \leq K_T.$$

**Proof.** By Girsanov theorem, $\mathbb{P}_T^N$ has a density with respect to $\mathbb{P}_T^{\otimes N}$ given by

$$\frac{d\mathbb{P}_T^N}{d\mathbb{P}_T^{\otimes N}}(Y) = \exp(H_T^N)$$

where

$$H_T^N = \sum_{i=1}^N \int_0^T \left[ b(Y_s^i, \frac{1}{N} \sum_{j=1}^N \kappa(Y_s^i - Y_s^j)) - b(Y_s^i, \kappa * u_s(Y_s^i)) \right] dB_s^i$$

$$- \frac{1}{2} \sum_{i=1}^N \int_0^T \left[ b(Y_s^i, \frac{1}{N} \sum_{j=1}^N \kappa(Y_s^i - Y_s^j)) - b(Y_s^i, \kappa * u_s(Y_s^i)) \right] ds.$$

Under the measure $\mathbb{P}_T^{\otimes N}$, $((B_t^i)_{0 \leq t \leq T})_{0 \leq i \leq N}$ is a $N$-dimensional Brownian motion and $((\overline{B}_t^i)_{0 \leq t \leq T})_{0 \leq i \leq N}$ defined by

$$\overline{B}_t^i = B_t^i - \int_0^T \left[ b(Y_s^i, \frac{1}{N} \sum_{j=1}^N \kappa(Y_s^i - Y_s^j)) - b(Y_s^i, \kappa * u_s(Y_s^i)) \right] ds$$

is a $N$-dimensional Brownian motion under $\mathbb{P}_T^N$. Then it follows from (6)

$$H\left(\mathbb{P}_T^N \mid \mathbb{P}_T^{\otimes N}\right) = \mathbb{E}_{\mathbb{P}_T^N} \left[ \log \left( \frac{d\mathbb{P}_T^N}{d\mathbb{P}_T^{\otimes N}} \right) \right]$$

$$= \frac{1}{2} \sum_{i=1}^N \mathbb{E} \int_0^T \left[ b(X_s^{i,N}, \kappa * \Pi_s^{N}(X_s^{i,N})) - b(X_s^{i,N}, \kappa * u_s(X_s^{i,N})) \right]^2 ds$$

$$\leq \frac{K}{2} \sum_{i=1}^N \mathbb{E} \int_0^T \left[ |\kappa * \Pi_s^{N}(X_s^{i,N}) - \kappa * u_s(X_s^{i,N})| \right]^2 ds$$

$$\leq \frac{K}{2N} \sum_{i,j=1}^N \mathbb{E} \int_0^T \left[ |\kappa(X_s^{i,N} - X_s^{j,N}) - \kappa * u_s(X_s^{i,N})| \right]^2 ds.$$
We write
\[ \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa \ast u_s(X^i_{s,N}) = \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa(X^i_{s,N} - X^j_{s,N}) + \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa(X^i_{s,N} - X^j_{s,N}) + \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa(X^i_{s,N} - X^j_{s,N}) + \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa(X^i_{s,N} - X^j_{s,N}) + \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa(X^i_{s,N} - X^j_{s,N}) \]
and get, by Theorem 5.2,
\[
\mathbb{E} \left[ \left| \kappa(X^i_{s,N} - X^j_{s,N}) - \kappa \ast u_s(X^i_{s,N}) \right|^2 \right] \leq \frac{K_T}{N}
\]
which achieves the proof.

Now we just have to use (15) and the Csiszár and Kullback inequality (14) to establish a strong convergence estimation.

Theorem 5.5. For all $T$ and $N$, there exists a constant $K_T$ such that, for all $k \leq N$,
\[
\| P^k_{T,N} - P^k_T \| \leq K_T \sqrt{\frac{k}{N}}.
\]
This implies a strong version of propagation of chaos for blocks of size $o(N)$.

5.3 Concentration of the empirical measure

We now want to establish that the particle system (in the generic case) has a bounded-below curvature. This is much more complicated than in the case where $\sigma$ is equal to $I$. Indeed, the second order part of the infinitesimal generator equips $\mathbb{R}^{dN}$ with a Riemannian metric which changes the notion of gradient.

5.3.1 Control of the Ricci curvature

For simplicity’s sake, we will suppose, in this section, $d = 1$. The case $d \geq 2$ can be treated by the same way but with complicated notations. In Remark 5.9, we sketch the changes that occur when $d$ is greater than 2. We treat here the case of Equation (19). An advanced study of Riemannian geometry can be found in Gallot et al. (1990) but let us try to describe the setting in few words. Consider the differential operator $L$ on $\mathbb{R}^N$ defined by
\[
L = \sum_{i,j=1}^{N} g^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} h_i(x) \frac{\partial}{\partial x_i}.
\]
Its second order part equips $\mathbb{R}^N$ with an intrinsic metric $g$ which depends on $x \in \mathbb{R}^N$ and which is characterized by the matrix $(g_{ij})$ (the inverse of $(g^{ij})$). The geometry of $(\mathbb{R}^N, g)$ is quite different from the usual one and it appears a curvature. We now show how to control Ricci curvature of the differential operators $L^N$ associated to (20) and defined by

$$L^N = \sum_{i=1}^{N} g^i(x) \frac{\partial^2}{\partial x_i^2} + \sum_{i=1}^{N} b_i(x) \frac{\partial}{\partial x_i}$$

(23)

where

$$g^i(x) = a \left( x_i, \frac{1}{N} \sum_{j=1}^{N} \kappa(x_i - x_j) \right), \quad g_i(x) = \frac{1}{g^i(x)}$$

(24)

and

$$b_i(x) = b \left( x_i, \frac{1}{N} \sum_{j=1}^{N} \kappa(x_i - x_j) \right).$$

Now let us recall the general definition of the Ricci curvature.

**Definition 5.6.** Let us define

1. **Christophel symbols** for $i, j, k = 1, \ldots, N$: $\Gamma_{ki}^j = \frac{1}{2} \sum_{p=1}^{N} g^{ip} \left( \frac{\partial g_{pj}}{\partial x_k} + \frac{\partial g_{kp}}{\partial x_i} - \frac{\partial g_{ki}}{\partial x_p} \right)$;

2. **the Riemann curvature tensor** $(R_{ql}^i)$ where

$$R_{ql}^i = \frac{\partial \Gamma_{qi}^k}{\partial x_l} - \frac{\partial \Gamma_{ql}^k}{\partial x_i} + \sum_{p=1}^{N} \Gamma_{pl}^i \Gamma_{qk}^p - \sum_{p=1}^{N} \Gamma_{pk}^i \Gamma_{ql}^p;$$

3. **the Ricci curvature tensor** $(R_{ql})$ where

$$R_{ql} = \sum_{i=1}^{N} R_{qi}^i;$$

4. **the drift tensor** $(m_{ij})$ where

$$M_{ij} = \frac{1}{2} (\nabla_i b_j + \nabla_j b_i) \quad \text{or} \quad \nabla_i b_j = \frac{\partial b_j}{\partial x_i} + \sum_{k=1}^{N} \Gamma_{ki}^j b_k.$$

In our context, the matrix $(G_{ij})$ is diagonal and then many terms in the previous definitions are equal to 0. For example, $\Gamma_{ki}^j$ is equal to 0 as soon as the three index are different. By a straightforward computation we have the explicit expression of the Ricci and drift tensors.
Lemma 5.7. The entries of Ricci tensor are

\[ R_{lq} = \frac{1}{2} \frac{\partial^2}{\partial x_l \partial x_q} \log \frac{g_1 \ldots g_N}{g_l g_q} + \frac{1}{4} \sum_{i \neq q, l} \left( \frac{\partial}{\partial x_i} \log g_i \right) \left( \frac{\partial}{\partial x_q} \log g_q \right) \]

and

\[ R_{qq} = \frac{1}{2} \left[ \frac{\partial^2}{\partial x_q^2} \log \frac{g_1 \ldots g_N}{g_q} + \sum_{i \neq q} g_i^2 \frac{\partial^2}{\partial x_i^2} g_q + \sum_{i \neq q} \frac{\partial g_i}{\partial x_i} \frac{\partial g_q}{\partial x_q} + \frac{1}{2} \sum_i \left( \frac{\partial \log g_i}{\partial x_q} \right)^2 \right] \]

\[ - \frac{1}{4} \sum_{i \neq q} g_q g_i \left( \frac{\partial \log g_q}{\partial x_i} \right)^2 + \frac{1}{8} \sum_i g_q g_i \left( \frac{\partial \log g_q}{\partial x_i} \right) \left( \frac{\partial \log (g_1 \ldots g_N)}{\partial x_i} \right) . \]

Moreover, the entries of the drift tensor have the following form:

\[ M_{ii} = \nabla_i b_i = \frac{\partial b_i}{\partial x} - \frac{1}{N} \sum_{k=1}^N \kappa'(x_k - x) \frac{\partial b_i}{\partial u} + \frac{1}{2} \sum_{k=1}^N \frac{\partial \log g_i}{\partial x_k} \]

et

\[ M_{ij} = \frac{1}{2} \left( \frac{\partial b_i}{\partial x_i} + \frac{\partial b_i}{\partial x_j} \right) + \left( 1 - g_i g_j \right) \frac{\partial \log g_i}{\partial x_j} b_i + \left( 1 - g_j g_i \right) \frac{\partial \log g_j}{\partial x_i} b_j . \]

We have now to control the spectrum of the matrix \( T \), equal to \( R - M \).

**Proposition 5.8.** If there exist a \( c \) such that \( 1/c \leq a(x, y) \leq c \) and if the first (resp the first and second) derivatives of \( b \) (resp \( a \)) are bounded by \( c \), the spectrum of \( T \) is uniformly bounded (above and below) in \( N \in \mathbb{N} \) by a real number \( \rho \).

**Proof.** By Gershgorin-Haddamard Theorem (see Grifone (1990)), if \( \lambda \) is an eigenvalue of \( T \),

\[ |\lambda| \leq \max_{1 \leq i \leq N} \left( \sum_{j=1}^N |T_{ij}| \right) . \]

We are going to show that, under the assumptions of proposition 5.8, \( R_{qq} \) et \( M_{qq} \) are uniformly bounded in \( N \) whereas \( R_{ql} \) and \( M_{ql} \) are of order \( 1/N \). In order to deal with simple notations, \( \alpha \) will be a common bound greater than \( 1 \) for \( \kappa \) and its derivatives. Then we have
\[
\left| \frac{\partial^2 \log g_1 \cdots g_N}{\partial x_l \partial x_q} \right| \leq \sum_{i \neq l, q} \left| \frac{\partial^2 \log g_i}{\partial x_l \partial x_q} \right|
\]
\[
\leq \frac{1}{N^2} \sum_{i \neq l, q} |\kappa'(x_i - x_l)\kappa'(x_i - x_q)| \frac{\partial^2 \log a}{\partial y_2^2} \]
\[
\leq \frac{1}{N^2} \sum_{i \neq l, q} 2\alpha^2 c^4
\]
and
\[
\sum_{i \neq q, l} \left| \frac{\partial}{\partial x_l} \log g_i \right| \left| \frac{\partial}{\partial x_q} \log g_i \right| \leq \frac{1}{N^2} \sum_{i \neq q, l} |\kappa'(x_i - x_l)\kappa'(x_i - x_q)| \left( \frac{\partial \log a}{\partial y_2} \right)^2
\]
\[
\leq \frac{1}{N^2} \sum_{i \neq q, l} \alpha^2 c^4.
\]
Then a short computation gives
\[
|R_{lq}| \leq \frac{7\alpha^2 c^4}{4N}.
\]
By the same way, it can be shown that
\[
\left\{
\begin{array}{l}
|R_{qq}| \leq 7\alpha^2 c^6 \\
|M_{ii}| \leq 3\alpha c^2 \\
|M_{ij}| \leq 5\alpha c^5/N.
\end{array}
\right.
\]
Then we have obtained the bound we were looking for. All the operators \( L^N \) have a curvature bounded below by \( \rho = 17\alpha^2 c^6 \).
\[\square\]

Of course, a more precise result could be obtained by specifying the bound of each derivative but this work is not very essential from a theoretic point of view.

**Remark 5.9.** When \( d \geq 2 \), \( g_i \) is a \( d \times d \) matrix. In order to control the spectrum of \( T \), one has to consider blocks of size \( d \times d \) as in the proof of Lemma 3.5.

5.3.2 Concentration of measure

**Theorem 5.10.** For all Lipschitz function with \( \|f\|_{Lip} \leq 1 \), we have, for all \( r \geq 0 \),
\[
\sup_{t \leq T} \mathbb{P} \left( \left| \frac{1}{N} \sum_{i=1}^N f(X_t^{i,N}) - \int f(y)u(t, y) \, dy \right| \geq r + \sqrt{\frac{K_T}{N}} \right) \leq 2 \exp \left( - \frac{N r^2}{C_T} \right).
\]
where
\[ C_T = \frac{2}{\rho} \left(1 - e^{-2\rho T}\right), \]
and \( u \) solution of (20).

**Remark 5.11.** Explicit forms of the constants \( \rho \) and \( K_T \) can be given. With the notations we have introduce in section 5.3.1,

\[
\begin{align*}
\rho &= 17\alpha^2 c^6 \\
K_T &= 6\alpha^4 c^2 \exp(12\alpha^2 c^2 T).
\end{align*}
\]

**Acknowledgments** This work is a part of my PhD thesis. I would like to thank my advisor Professor D. Bakry for his useful advices and encouragements.

**References**


