Entropy and mean field models

Arnaud Guillin
With: Wei Liu, Liming Wu, Chaoen Zhang, Pierre Le Bris, Pierre Monmarché

About Entropy in Large Classical Particle Systems

Université Clermont Auvergne / IUF

I. Introduction

The model(s)

We will be interested here with particles system in mean field interactions

$$dX_t^i = \sqrt{2\sigma_N}dB_t^i - \nabla U(X_t^i)dt - \frac{1}{N}\sum_{j\neq i} K(X_t^i - X_t^j)dt,$$

where

- σ_N diffusion coefficient, $\sigma_N = \sqrt{2\sigma/N}$ or σ,
- $((B_t^i)_{t\geq 0})_i$ independent Brownian motions,
- U confining potential, e.g. $\nabla U(x) = \lambda x$ with $\lambda > 0$ or 0,
- W is an interaction potential :
    - $K(x) = \nabla W$, regular, but not too large,
    - in dimension 2, Biot Savart kernel

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

Particles system and its long time behavior

Focus on

$$dX_t^i = \sqrt{2}dB_t^i - \nabla U(X_t^i)dt - \frac{1}{N}\sum_{j\neq i}\nabla W(X_t^i - X_t^j)dt,$$

whose invariant measure is

$$d\mu^N = e^{-\sum_{i=1}^N U(x_i) - \frac{1}{N}\sum_{ij} W(x_i - x_j)}dx$$

How can we study the convergence to equilibrium, if possible uniform in N ?

- Meyn-Tweedie's techniques... very general but poorly quantitative and never uniform in N.
- Coupling... story for another day
- functional inequalities !

Extensions to the kinetic case possible.

Particles system and its limit in N

$$dX_t^i = \sqrt{2\sigma_N} dB_t^i - U'(X_t^i)dt - \frac{1}{N}\sum_{j \neq i} W'(X_t^i - X_t^j)dt.$$

Particles system and its limit in N

$$dX_t^i = \sqrt{2\sigma_N} dB_t^i - U'(X_t^i)dt - \frac{1}{N}\sum_{j\neq i} W'(X_t^i - X_t^j)dt.$$

Formally, notice $\frac{1}{N}\sum_{j=1}^N W'(X_t^i - X_t^j) = V' * \mu_t^N(X_t^i)$, where

$$\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}.$$

Denote also $\rho_t^N = Law(X_t^1, \cdots, X_t^N)$.

Particles system and its limit in N

$$dX_t^i = \sqrt{2\sigma_N} dB_t^i - U'(X_t^i)dt - \frac{1}{N}\sum_{j\neq i} W'(X_t^i - X_t^j)dt.$$

Formally, notice $\frac{1}{N}\sum_{j=1}^{N} W'(X_t^i - X_t^j) = V' * \mu_t^N(X_t^i)$, where

$$\mu_t^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^i}.$$

Denote also $\rho_t^N = Law(X_t^1, \cdots, X_t^N)$.

Assuming $\sigma_N \to \sigma$, we should get a nonlinear McKean-Vlasov equation

$$\begin{cases} dX_t = \sqrt{2\sigma} dB_t - U'(X_t)dt - W' * \bar{\rho}_t(X_t)dt, \\ \bar{\rho}_t = \text{Law}(X_t), \end{cases}$$

Particles system and its limit in N

$$dX_t^i = \sqrt{2\sigma_N} dB_t^i - U'(X_t^i)dt - \frac{1}{N}\sum_{j\neq i} W'(X_t^i - X_t^j)dt.$$

Formally, notice $\frac{1}{N}\sum_{j=1}^N W'(X_t^i - X_t^j) = V' * \mu_t^N(X_t^i)$, where

$$\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}.$$

Denote also $\rho_t^N = Law(X_t^1, \cdots, X_t^N)$.

Assuming $\sigma_N \to \sigma$, we should get a nonlinear McKean-Vlasov equation

$$\begin{cases} dX_t = \sqrt{2\sigma} dB_t - U'(X_t)dt - W' * \bar{\rho}_t(X_t)dt, \\ \bar{\rho}_t = \text{Law}(X_t), \end{cases}$$

which is linked to

$$\partial_t \bar{\rho}_t = \partial_x \left(\left(U' + V' * \bar{\rho}_t \right) \bar{\rho}_t \right) + \sigma \partial_{xx}^2 \bar{\rho}_t.$$

Idea

In a system of N exchangeable interacting particles, as N increases, two particles become more and more statistically independent.

Mark Kac introduced the terminology *Propagation of chaos* to describe this phenomenon.

There is equivalence between

- **Local estimate** $\rho_t^{N,k} = \mathcal{L}(X_t^1, ..., X_t^k) \to \bar{\rho}_t^{\otimes k}$
- **Global estimate** $\mu_t^N \to \bar{\rho}_t$

Usual methods

Goal : Show $\mu_t^N \to \bar{\rho}_t$, or $\rho_t^1 \to \bar{\rho}_t$ *quantitatively and uniformly in time*.

Usual methods

Goal : Show $\mu_t^N \to \bar{\rho}_t$, or $\rho_t^1 \to \bar{\rho}_t$ *quantitatively and uniformly in time*.

Some methods :

Usual methods

Goal : Show $\mu_t^N \to \bar{\rho}_t$, or $\rho_t^1 \to \bar{\rho}_t$ *quantitatively and uniformly in time*.

Some methods :

- **Tightness** *(Rogers-Shi, Méléard, Cépa-Lépingle, Fournier-Hauray-Mischler, ...)*.

Usual methods

Goal : Show $\mu_t^N \to \bar{\rho}_t$, or $\rho_t^1 \to \bar{\rho}_t$ *quantitatively and uniformly in time*.

Some methods :

- **Tightness** *(Rogers-Shi, Méléard, Cépa-Lépingle, Fournier-Hauray-Mischler, ...)*.
- **Coupling methods** *(McKean, Sznitman, Malrieu, Durmus, Eberle,...) :*

$$\mathcal{W}_2\left(\mu, \nu\right)^2 = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\left(|X-Y|^2\right).$$

Usual methods

Goal : Show $\mu_t^N \to \bar{\rho}_t$, or $\rho_t^1 \to \bar{\rho}_t$ *quantitatively and uniformly in time*.

Some methods :

- **Tightness** *(Rogers-Shi, Méléard, Cépa-Lépingle, Fournier-Hauray-Mischler, ...)*.
- **Coupling methods** *(McKean, Sznitman, Malrieu, Durmus, Eberle,...) :*

$$\mathcal{W}_2(\mu, \nu)^2 = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\left(|X - Y|^2\right).$$

- **Energy/Entropy estimates** *(Malrieu, Mischler, Mouhot, Rosenzweig, Serfaty, Jabin, Wang, Lacker, Bresch, Soler, Poyato, Delgadino, Carrillo, Pavliotis, Gvalani, Tugaut,...)*.
  For example the rescaled relative entropy

$$\mathcal{H}_N(\nu, \mu) = \begin{cases} \frac{1}{N}\mathbb{E}_\mu\left(\frac{d\nu}{d\mu}\log\frac{d\nu}{d\mu}\right) & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

Usual methods

Goal : Show $\mu_t^N \to \bar{\rho}_t$, or $\rho_t^1 \to \bar{\rho}_t$ *quantitatively and uniformly in time*.

Some methods :

- **Tightness** *(Rogers-Shi, Méléard, Cépa-Lépingle, Fournier-Hauray-Mischler, ...)*.
- **Coupling methods** *(McKean, Sznitman, Malrieu, Durmus, Eberle,...) :*

$$\mathcal{W}_2\left(\mu,\nu\right)^2 = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\left(|X-Y|^2\right).$$

- **Energy/Entropy estimates** *(Malrieu, Mischler, Mouhot, Rosenzweig, Serfaty, Jabin, Wang, Lacker, Bresch, Soler, Poyato, Delgadino, Carrillo, Pavliotis, Gvalani, Tugaut,...)*.
  For example the rescaled relative entropy

$$\mathcal{H}_N(\nu,\mu) = \begin{cases} \frac{1}{N}\mathbb{E}_\mu\left(\frac{d\nu}{d\mu}\log\frac{d\nu}{d\mu}\right) & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

- **Weak norm and Lions derivatives calculus** (Chassagneux, Szpruch, Tse, Delarue, ...)

Logarithmic Sobolev inequality

Entropy and long time convergence of diffusions

Let us first focus on the simple case

$$dX_t = \sqrt{2}dB_t - \nabla U(x_t)dt$$

reversible wrt $\mu = e^{-U}$, semigroup denoted P_t and generator L.

Definition

We say that μ satisfies a logarithmic Sobolev inequality if for all nice function f

$$\mathrm{Ent}_\mu(f^2) = \int f^2 \log\left(\frac{f^2}{\int^2 f d\mu}\right) d\mu \leq 2C_{LS} \int |\nabla f|^2 d\mu.$$

We say that μ satisfies a Poincaré inequality if for all nice function f

$$\mathrm{Var}_\mu(f) = \int (f - \int f d\mu)^2 d\mu \leq C_P \int |\nabla f|^2 d\mu.$$

Pinsker inequality. Let f be a density wrt μ

$$\|f\mu - \mu\|_{TV} \leq \sqrt{2\mathrm{Ent}_\mu(f)}.$$

Some nice properties of LSI

Theorem

1. *LSI is equivalent to*

$$Ent_\mu(P_t f) \le e^{-2t/C_{LS}} Ent_\mu(f).$$

2. *LSI is equivalent to hypercontractivity.*
3. *LSI implies Talagrand Inequality :* $\forall f,\ W_2^2(f\mu, \mu) \le 2C\, Ent_\mu(f)$*.*
4. *LSI implies Poincaré inequality (equivalent to* L^2 *convergence to equilibrium).*
5. *LSI implies Gaussian concentration*

$$\mu(f - \mu(f) > r) \le e^{-cr^2}.$$

6. *Perturbation : if* μ *satisfies LSI and* $d\nu = e^V d\mu$ *with* V *bounded then* ν *satisfies LSI.*
7. *Tensorization : if* μ *satisfies LSI(*C_{LS}*) then so does* $\mu^{\otimes N}$*.*

(see the excellent book of Bakry-Gentil-Ledoux)

How to prove LSI ?

There are some general well known sufficient conditions

- Bakry-Emery Γ_2 condition : Hess$U \geq \kappa > 0$
  (refined multi-scale Bakry-Emery condition by Bauerschmidt-Bodineau)
- Capacity-measure condition : $Cap_\mu(A) \geq c\,\mu(A)\log(1/\mu(A))$ which can be transfered to Hardy's type condition in dimension 1.
- Lyapunov condition : $\exists V \geq 1, c > 0, b > 0$ such that

$$LV(x) \leq -cU(x)\,V(x) + b.$$

Combined with tensorisation and perturbation, it leads to nice examples...

However for our mean field model where

$$d\mu^N = e^{-\sum_{i=1}^N U(x_i) + \frac{1}{N}\sum_{i,j} W(x^i, x^j)}$$

it is harder to get adimensional LSI.

Some litterature

- spin systems
  - Stroock-Zegarlinski (92) on LSI, Dobrushin condition and conitnuous spin system,
  - Zegarlinski (92, 96) on Dobrushin uniqueness theorem and LSI,
  - Bodineau-Helffer (99) on LSI for unbounded spin systems,
  - Yoshida (00,01) on LSI and mixing condition,
  - Ledoux (01) for nice review/results,
  - Guionnet-Zegarlinski (04) for a survey,
  - Bodineau-Bauerschmidt (19) for the introduction of the multi-scale Bakry Emery condition, via stochastic localization,
- mean field models
  - Malrieu (01,03) via Bakry-Emery
  - Carrillo-McCann-Villani (03) for the non linear limit,
  - Cattiaux-G-Malrieu (08), Eberle-G-Zimmer (19) via coupling
  - Bolley-Gentil-G (13) via functional inequalities for the non linear limit

LSI for mean field model

Theorem (G-Liu-Wu-Zhang)

Recall $d\mu^N = e^{-\sum_{i=1}^N U(x_i) + \frac{1}{N}\sum_{i,j} W(x^i, x^j)}$. *Assume*

1. *that the one-dimensional conditional distribution* $\mu_{\hat{x}_i}$ *satisfies a LSi(*$\hat{C}_{LS,m}$*) independently of* $\hat{x}_i$*,*
2. *that* $\gamma_0 := C_{lip,m} \sup_{x,y,|z|=1} |\nabla^2 1_{x,y} W(x,y) z| < 1$ *where*

$$C_{lip,m} := \frac{1}{4}\int_0^\infty \exp\left(\frac{1}{4}\int_0^s b_0(u)du\right) sds$$

$$b_0(r) = \sup_{|x-y|=r} -\frac{x-y}{|x-y|}.((\nabla U(x) - \nabla U(y)) + (\nabla_x W(x,z) - \nabla_x (W(y,z)))$$

then μ^N *satisfies a LSI(*$\hat{C}_{LS,m} : (1-\gamma_0)^2$*).*

Some elements of proof

We have to control Zegarlinski's interdependance coefficients c_{ij}^Z defined as

$$|\nabla_i(\mu_{\hat{x}^j}(f^2))^{1/2}| \leq (\mu_{\hat{x}^j}(|\nabla_i f|^2))^{1/2} + c_{ij}^{\mathbf{Z}}(\mu_{\hat{x}^j}(|\nabla_j f|^2))^{1/2}.$$

Step 1.

First remark that for C_b^1 function, if

$$|\nabla_i \mu_{\hat{x}^j}(g)| \leq c_{ij}\mu_{\hat{x}^j}(|\nabla g|), \ \forall \hat{x}^j$$

then $c_{ij}^Z \leq c_{ij}$.

Some elements of proof

Step 2.

$$\begin{aligned}\nabla_j \mu_{\hat{x}^i}(g) &= \nabla_j \left(\int g(x_i) e^{-H(x_1,\cdots,x_N)} dx_i / \int e^{-H(x_1,\cdots,x_N)} dx_i \right) \\ &= \frac{\int g(x_i)(-\nabla_j H) e^{-H} dx_i}{\int e^{-H} dx_i} + \frac{\int g(x_i) e^{-H} dx_i \int \nabla_j H \cdot e^{-H} dx_i}{(\int e^{-H} dx_i)^2} \\ &= -\int g(x_i) \nabla_j H d\mu_{\hat{x}^i} + \int g(x_i) d\mu_{\hat{x}^i} \int \nabla_j H d\mu_{\hat{x}^i} \\ &= \mathrm{Cov}_{\mu_{\hat{x}^i}} \left(g, -\frac{1}{N-1} (\nabla_y W)(x_i, x_j) \right)\end{aligned}$$

and then

$$\begin{aligned} z \cdot \nabla_{x_j} \mu_{\hat{x}^i}(g) &= \mathrm{Cov}_{\mu_{\hat{x}^i}}(g, -\frac{1}{N-1} (\nabla_y W)(x_i, x_j) \cdot z) \\ &= -\frac{1}{N-1} < g, (\nabla_y W)(x_i, x_j) \cdot z - \mu_{\hat{x}^i}((\nabla_y W)(\cdot, x_j) \cdot z) >_{\mu_{\hat{x}^i}} \\ &= -\frac{1}{N-1} < (-L_i) g, (-L_i)^{-1} [\text{same}] >_{\mu_{\hat{x}^i}} \\ &= -\frac{1}{N-1} \int \nabla_i g \cdot \nabla_i (-L_i)^{-1} [\text{same}] d\mu_{\hat{x}^i}. \end{aligned}$$

Some elements of proof

Step 3.

Now we have that

$$||(-L_i)^{-1}||_{\text{Lip}} \le c_{Lip,m}$$

and thus

$$\begin{aligned}
&\|\nabla_i(-L_i)^{-1}((\nabla_y W)(\cdot, x_j)\cdot z - \mu_{\hat{x}^i}((\nabla_y W)(\cdot, x_j)\cdot z))\|_{L^\infty(\mu_{\hat{x}^i})} \\
&\quad \le c_{Lip,m} \sup_{x_i, x_j} |\nabla_{x_i}((\nabla_y W)(x_i, x_j)\cdot z)| \\
&\quad = c_{Lip,m} \sup_{x,y} |\nabla^2_{x,y} W(x,y) z| \\
&\quad \le c_{Lip,m} \|\nabla^2_{x,y} W\|_\infty .
\end{aligned}$$

Example and application

- **1d Curie-Weiss model.** $U(x) = \beta(\frac{x^4}{4} - \frac{x^2}{2})$, $W(x,y) = -\beta Kxy$ then

$$C_{lip,m} \leq \sqrt{\frac{\pi}{\beta}} e^{\beta/4} \qquad \gamma_0 \leq \sqrt{\pi\beta} e^{\beta/4}|K|.$$

- **Application.** for $d\nu = fdx$

$$E(\nu) = \int f \log f + \frac{1}{2}\int\int W(x,y)d\nu(x)d\nu(y), \qquad H_W(\nu = E(\nu) - \inf E$$

$$I_W := \frac{1}{4}\int \left|\frac{\nabla f}{f} + \nabla V + \nabla_x \int W(x,y)d\nu(y)\right|^2 d\nu$$

Theorem

Under the previous assumptions,

1. *H_W has a unique minimizer ν_∞*
2. *the non-linear LSI holds : $H_W(\nu) \leq 2C_{LS} I_W(\nu)$.*
3. *$H_W(\nu_t) \leq e^{-t/(2C_{LS})} H_W(\nu_0)$.*

Let us refer also to the work of Delgadino-Pavliotis-Rabshani-Smith who relates the (uniform in N) LSI to uniform in time propagation of chaos.

LSI and uniform in time propagation of chaos for vortex 2D case

Motivation : the 2D vortex model

The Biot-Savart kernel, defined in $\mathbb{T}^2$ by

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

Motivation : the 2D vortex model

The Biot-Savart kernel, defined in $\mathbb{T}^2$ by

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

Consider the 2D incompressible Navier-Stokes system on $u \in \mathbb{T}^2$

$$\begin{aligned} \partial_t u =& - u \cdot \nabla u - \nabla p + \Delta u \\ \nabla \cdot u =& 0, \end{aligned}$$

where p is the local pressure. Curl of the equation leads to $\omega(t,x) = \nabla \times u(t,x)$ satisfies

$$\partial_t \omega = -\nabla \cdot ((K * \omega)\,\omega) + \Delta\omega.$$

Motivation : the 2D vortex model

The Biot-Savart kernel, defined in $\mathbb{T}^2$ by

$$K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

Consider the 2D incompressible Navier-Stokes system on $u \in \mathbb{T}^2$

$$\begin{aligned} \partial_t u =& - u \cdot \nabla u - \nabla p + \Delta u \\ \nabla \cdot u =& 0, \end{aligned}$$

where p is the local pressure. Curl of the equation leads to $\omega(t,x) = \nabla \times u(t,x)$ satisfies

$$\partial_t \omega = -\nabla \cdot ((K * \omega)\,\omega) + \Delta \omega.$$

Goal : Use a particle system to approximate the solution of the 2D vortex model !

the 2D vortex model

The Biot-Savart kernel, defined in $\mathbb{T}^2$ by

$$K(x) = \frac{1}{2\pi}\frac{x^{\perp}}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right).$$

and the 2D vortex equation

$$\partial_t \bar{\rho} = -\nabla \cdot ((K * \bar{\rho})\,\bar{\rho}) + \Delta\bar{\rho}.$$

and the particles system

$$dX_t^i = \sqrt{2}dB_t^i + \frac{1}{N}\sum_{j\neq i} K(X_t^i - X_t^j)dt$$

with law ρ^N.

Theorem (adapted from Jabin-Wang ('18))

Under some assumptions (satisfied by the Biot-Savart kernel) there are positive C_1 and C_2 such that for all $N \in \mathbb{N}$, all exchangeable probability density ρ_0^N and all $t \geq 0$

$$\mathcal{H}_N(\rho_t^N, \bar{\rho}_t^N) \leq e^{C_1 t} \left(\mathcal{H}_N(\rho_0^N, \bar{\rho}_0^N) + \frac{C_2(t)}{N} \right)$$

where $\bar{\rho}^N$ stands for the law of N independent copies of non linear particles.

Results

Theorem (adapted from Jabin-Wang ('18))

Under some assumptions (satisfied by the Biot-Savart kernel) there are positive C_1 and C_2 such that for all $N \in \mathbb{N}$, all exchangeable probability density ρ_0^N and all $t \geq 0$

$$\mathcal{H}_N(\rho_t^N, \bar{\rho}_t^N) \leq e^{C_1 t} \left(\mathcal{H}_N(\rho_0^N, \bar{\rho}_0^N) + \frac{C_2(t)}{N} \right)$$

where $\bar{\rho}^N$ stands for the law of N independent copies of non linear particles.

Theorem (G-Le Bris-Monmarché)

Under some assumptions (satisfied by the Biot-Savart kernel) there are positive constants C_1, C_2 and C_3 such that for all $N \in \mathbb{N}$, all exchangeable probability density ρ_0^N and all $t \geq 0$

$$\mathcal{H}_N(\rho_t^N, \bar{\rho}_t^N) \leq C_1 e^{-C_2 t} \mathcal{H}_N(\rho_0^N, \bar{\rho}_0^N) + \frac{C_3}{N}$$

Various distances

Corollary

Under some assumptions (satisfied by the Biot-Savart kernel), assuming moreover that $\rho_0^N = \bar{\rho}_0^N$, there is a constant C such that for all $k \leq N \in \mathbb{N}$ and all $t \geq 0$,

$$\|\rho_t^{k,N} - \bar{\rho}_t^k\|_{L^1} + \mathcal{W}_2\left(\rho_t^{k,N}, \bar{\rho}_t^k\right) \leq C\left(\left\lfloor \frac{N}{k} \right\rfloor\right)^{-\frac{1}{2}}$$

Step one : Time evolution of the relative entropy

We write

$$\mathcal{H}_N(t) = \mathcal{H}_N(\rho_t^N, \bar{\rho}_t^N), \;\; \mathcal{I}_N(t) = \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_t^N \left| \nabla_{x_i} \log \frac{\rho_t^N}{\bar{\rho}_t^N} \right|^2 d\mathbf{X}^N.$$

Step one : Time evolution of the relative entropy

We write

$$\mathcal{H}_N(t) = \mathcal{H}_N(\rho_t^N, \bar\rho_t^N)\,, \;\; \mathcal{I}_N(t) = \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_t^N \left| \nabla_{x_i} \log \frac{\rho_t^N}{\bar\rho_t^N} \right|^2 d\mathbf{X}^N.$$

By derivating

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_N(t) \le & -\mathcal{I}_N(t) \\ & - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(K(x_i - x_j) - K * \rho(x_i)\right) \cdot \nabla_{x_i} \log \bar\rho_t^N d\mathbf{X}^N \\ & - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(\operatorname{div} K(x_i - x_j) - \operatorname{div} K * \bar\rho_t(x_i)\right) d\mathbf{X}^N.\end{aligned}$$

Assumptions ?

$$\textbf{Goal : } K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$$

Justifying the calculations

- $\bar{\rho} \in \mathcal{C}^\infty(\mathbb{R}^+ \times \mathbb{T}^d)$

Assumptions ?

Goal : $K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$

Justifying the calculations

- $\bar\rho \in \mathcal{C}^\infty(\mathbb{R}^+ \times \mathbb{T}^d)$ and there is $\lambda > 1$, s.t $\frac{1}{\lambda} \leq \bar\rho \leq \lambda$

Assumptions ?

$$\textbf{Goal :} K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$$

Justifying the calculations

- $\bar{\rho} \in \mathcal{C}^\infty_\lambda(\mathbb{R}^+ \times \mathbb{T}^d)$

Assumptions ?

$$\textbf{Goal :}\ K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$$

Justifying the calculations

- There is $\lambda > 1$ such that $\bar{\rho}_0 \in \mathcal{C}^\infty_\lambda(\mathbb{T}^d)$
  $\implies \bar{\rho} \in \mathcal{C}^\infty_\lambda(\mathbb{R}^+ \times \mathbb{T}^d)$ (Ben-Artzi ('94))

Assumptions ?

$$\textbf{Goal :}\ K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$$

Justifying the calculations

- There is $\lambda > 1$ such that $\bar{\rho}_0 \in \mathcal{C}^\infty_\lambda(\mathbb{T}^d)$
  $\implies \bar{\rho} \in \mathcal{C}^\infty_\lambda(\mathbb{R}^+ \times \mathbb{T}^d)$ (Ben-Artzi ('94))
- $\rho^N \in \mathcal{C}^\infty_\lambda(\mathbb{R}^+ \times \mathbb{T}^{Nd})$ (? ? ?)

Assumptions ?

Goal : $K(x) = \frac{1}{2\pi}\frac{x^{\perp}}{|x|^2} = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)$

Justifying the calculations

- There is $\lambda > 1$ such that $\bar{\rho}_0 \in \mathcal{C}^{\infty}_{\lambda}(\mathbb{T}^d)$
  $\implies \bar{\rho} \in \mathcal{C}^{\infty}_{\lambda}(\mathbb{R}^+ \times \mathbb{T}^d)$ (Ben-Artzi ('94))
- $\rho^N \in \mathcal{C}^{\infty}_{\lambda}(\mathbb{R}^+ \times \mathbb{T}^{Nd})$ (? ? ?)

Dealing with the terms

- In the sense of distributions, $\nabla \cdot K = 0$.

Step one : Time evolution of the relative entropy

We write

$$\mathcal{H}_N(t) = \mathcal{H}_N(\rho_t^N, \bar{\rho}_t^N)\,, \quad \mathcal{I}_N(t) = \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_t^N \left|\nabla_{x_i} \log \frac{\rho_t^N}{\bar{\rho}_t^N}\right|^2 d\mathbf{X}^N.$$

By properly derivating

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_N(t) \le & -\mathcal{I}_N(t) \\ & - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(K(x_i - x_j) - K * \rho(x_i)\right)\cdot \nabla_{x_i} \log \bar{\rho}_t^N d\mathbf{X}^N \\ & - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(\operatorname{div} K(x_i - x_j) - \operatorname{div} K * \bar{\rho}_t(x_i)\right) d\mathbf{X}^N.\end{aligned}$$

Step one : Time evolution of the relative entropy

We write

$$\mathcal{H}_N(t) = \mathcal{H}_N(\rho_t^N, \bar{\rho}_t^N)\,, \quad \mathcal{I}_N(t) = \frac{1}{N}\sum_i \int_{\mathbb{T}^{dN}} \rho_t^N \left|\nabla_{x_i} \log \frac{\rho_t^N}{\bar{\rho}_t^N}\right|^2 d\mathbf{X}^N.$$

By properly derivating

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_N(t) \leq & -\mathcal{I}_N(t) \\ & - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(K(x_i - x_j) - K * \rho(x_i)\right) \cdot \nabla_{x_i} \log \bar{\rho}_t^N d\mathbf{X}^N \\ & - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(\cancel{\operatorname{div} K(x_i - x_j)} - \cancel{\operatorname{div} K * \bar{\rho}_t(x_i)}\right) d\mathbf{X}^N.\end{aligned}$$

Step two : Integration by part

We are left with

$$\frac{d}{dt}\mathcal{H}_N(t) \leq -\mathcal{I}_N(t) - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(K(x_i - x_j) - K * \rho(x_i)\right) \cdot \nabla_{x_i} \log \bar{\rho}_t^N d\mathbf{X}^N.$$

Idea : Use the regularity of $\bar{\rho}$ to deal with the singularity of K

Step two : Integration by part

We are left with

$$\frac{d}{dt}\mathcal{H}_N(t) \leq -\mathcal{I}_N(t) \\ \qquad - \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(K(x_i - x_j) - K * \rho(x_i)\right) \cdot \nabla_{x_i} \log \bar{\rho}_t^N d\mathbf{X}^N.$$

Idea : Use the regularity of $\bar{\rho}$ to deal with the singularity of K

Remark : Notice that, for the Biot-Savart kernel on the whole space $\mathbb{R}^2$

$$\tilde{K}(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2},$$

we have $\tilde{K} = \nabla \cdot \tilde{V}$ with $\tilde{V}$ bounded :

$$\tilde{V}(x) = \frac{1}{2\pi}\begin{pmatrix} -\arctan\left(\frac{x_1}{x_2}\right) & 0 \\ 0 & \arctan\left(\frac{x_2}{x_1}\right) \end{pmatrix}.$$

Step two : Integration by part

For all $t \geqslant 0$,

$$\frac{d}{dt}\mathcal{H}_N(t) \leq A_N(t) + \frac{1}{2}B_N(t) - \frac{1}{2}\mathcal{I}_N(t),$$

with

$$A_N(t) := \frac{1}{N^2}\sum_{i,j}\int_{\mathbb{T}^{dN}} \rho_t^N \left(V(x_i - x_j) - V * \bar\rho(x_i)\right) : \frac{\nabla^2_{x_i}\bar\rho_t^N}{\bar\rho_t^N} d\mathbf{X}^N$$

$$B_N(t) := \frac{1}{N}\sum_{i}\int_{\mathbb{T}^{dN}} \rho_t^N \frac{\left|\nabla_{x_i}\bar\rho_t^N\right|^2}{|\bar\rho_t^N|^2} \left|\frac{1}{N}\sum_j V(x_i - x_j) - V * \bar\rho(x_i)\right|^2 d\mathbf{X}^N.$$

Note that we would prefer to deal with the non linear particles which are i.i.d. We want also to get rid of the derivatives $\left|\nabla_{x_i}\bar\rho_t^N\right|$ or $\left|\nabla^2_{x_i}\bar\rho_t^N\right|$.

Step three : Change of reference measure and large deviation estimates

Lemma

For two probability densities μ *and* ν *on a set* Ω*, and any* $\Phi \in L^\infty(\Omega)$*,* $\eta > 0$ *and* $N \in \mathbb{N}$*,*

$$\mathbb{E}^\mu \Phi \leq \eta \mathcal{H}_N(\mu, \nu) + \frac{\eta}{N} \log \mathbb{E}^\nu e^{N\Phi/\eta}.$$

Theorem (Jabin-Wang '18)

For a p.m. μ *on* $\mathbb{T}^d$ *and* $\phi \in L^\infty(\mathbb{T}^d \times \mathbb{T}^d)$ *with*

$$\gamma := \left(1600^2 + 36e^4\right)\left(\sup_{p \geq 1} \frac{\| \sup_z |\phi(\cdot, z)| \|_{L^p(\mu))}}{p}\right)^2 < 1.$$

If ϕ *satisfies* $\int_{\mathbb{T}^d} \phi(x,z)\mu(dx) = 0 = \int_{\mathbb{T}^d} \phi(z,x)\mu(dx)$*. Then, for all* $N \in \mathbb{N}$*,*

$$\int_{\mathbb{T}^{dN}} \exp\left(\frac{1}{N}\sum_{i,j=1}^N \phi(x_i, x_j)\right)\mu^{\otimes N} d\boldsymbol{X}^N \leq \frac{2}{1-\gamma} < \infty.$$

Conclusion

For all $t \geqslant 0$,

$$\frac{d}{dt}\mathcal{H}_N(t) \leq C\left(\mathcal{H}_N(t) + \frac{1}{N}\right) - \frac{1}{2}\mathcal{I}_N(t),$$

with

$$C = \hat{C}_1 \|\nabla^2 \bar{\rho}_t\|_{L^\infty} \|V\|_{L^\infty} \lambda + \hat{C}_2 \|V\|_{L^\infty}^2 \lambda^2 d^2 \|\nabla \bar{\rho}_t\|_{L^\infty}^2$$

where $\hat{C}_1, \hat{C}_2$ are universal constants.

We have to compensate $\mathcal{H}_N(t)$ by $\mathcal{I}_N(t)$!

Step four : Uniform bounds and logarithmic Sobolev inequality

Two goals :

- A logarithmic Sobolev inequality for $\bar{\rho}^N$: $\mathcal{H}_N(t) \leq C_{LS}\, \mathcal{I}_N(t)$

Step four : Uniform bounds and logarithmic Sobolev inequality

Two goals :

- A logarithmic Sobolev inequality for $\bar{\rho}^N$: $\mathcal{H}_N(t) \leq C_{LS}\,\mathcal{I}_N(t)$
- Uniform in time bounds on $\|\nabla\bar{\rho}_t\|_{L^\infty}$ and $\|\nabla^2\bar{\rho}_t\|_{L^\infty}$

A logarithmic Sobolev inequality

Lemma (Tensorization)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, then for all $N \geq 0$, $\nu^{\otimes N}$ satisfies a LSI with constant C_ν^{LS}

A logarithmic Sobolev inequality

Lemma (Tensorization)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, then for all $N \geq 0$, $\nu^{\otimes N}$ satisfies a LSI with constant C_ν^{LS}

Lemma (Perturbation)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, and μ is a probability measure with density h with respect to ν such that, for some constant $\lambda > 0$, $\frac{1}{\lambda} \leq h \leq \lambda$, then μ satisfies a LSI with constant $C_\mu^{LS} = \lambda^2 C_\nu^{LS}$.

A logarithmic Sobolev inequality

Lemma (Tensorization)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, then for all $N \geq 0$, $\nu^{\otimes N}$ satisfies a LSI with constant C_ν^{LS}

Lemma (Perturbation)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, and μ is a probability measure with density h with respect to ν such that, for some constant $\lambda > 0$, $\frac{1}{\lambda} \leq h \leq \lambda$, then μ satisfies a LSI with constant $C_\mu^{LS} = \lambda^2 C_\nu^{LS}$.

Lemma (LSI for the uniform distribution)

The uniform distribution u on $\mathbb{T}^d$ satisfies a LSI with constant $\frac{1}{8\pi^2}$.

A logarithmic Sobolev inequality

Lemma (Tensorization)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, then for all $N \geq 0$, $\nu^{\otimes N}$ satisfies a LSI with constant C_ν^{LS}

Lemma (Perturbation)

If ν is a probability measure on $\mathbb{T}^d$ satisfying a LSI with constant C_ν^{LS}, and μ is a probability measure with density h with respect to ν such that, for some constant $\lambda > 0$, $\frac{1}{\lambda} \leq h \leq \lambda$, then μ satisfies a LSI with constant $C_\mu^{LS} = \lambda^2 C_\nu^{LS}$.

Lemma (LSI for the uniform distribution)

The uniform distribution u on $\mathbb{T}^d$ satisfies a LSI with constant $\frac{1}{8\pi^2}$.

For all $N \in \mathbb{N}$, $t \geq 0$ and all probability density $\mu_N \in \mathcal{C}^\infty_{>0}(\mathbb{T}^{dN})$,

$$\mathcal{H}_N\left(\mu_N, \bar{\rho}_t^N\right) \leq \frac{\lambda^2}{8\pi^2} \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{T}^d} \mu_N \left| \nabla_{x_i} \log \frac{\mu_N}{\bar{\rho}_t^N} \right|^2 d\mathbf{X}^N$$

Uniform in time bounds on the derivatives

Lemma

For all $n \geqslant 1$ and $\alpha_1, ..., \alpha_n \in [\![1, d]\!]$, there exist $C_n^u, C_n^\infty > 0$ such that for all $t \geqslant 0$,

$$\|\partial_{\alpha_1,\ldots,\alpha_n}\bar{\rho}_t\|_{L^\infty} \leq C_n^u \quad \textit{and} \quad \int_0^t \|\partial_{\alpha_1,\ldots,\alpha_n}\bar{\rho}_s\|_{L^\infty}^2 ds \leq C_n^\infty$$

Thanks to Morrey's inequality and Sobolev embeddings, it is sufficient to prove such bounds in the Sobolev space H^m for all m, i.e in L^2

Step five : Conclusion

There are constants $C_1, C_2^\infty, C_3 > 0$ and a function $t \mapsto C_2(t) > 0$ with $\int_0^t C_2(s)ds \leq C_2^\infty$ for all $t \geq 0$ such that for all $t \geq 0$

$$\frac{d}{dt}\mathcal{H}_N(t) \leq -(C_1 - C_2(t))\mathcal{H}_N(t) + \frac{C_3}{N}.$$

Multiplying by $\exp(C_1 t - \int_0^t C_2(s)ds)$ and integrating in time we get

$$\begin{aligned}\mathcal{H}_N(t) &\leq e^{-C_1 t + \int_0^t C_2(s)ds}\mathcal{H}_N(0) + \frac{C_3}{N}\int_0^t e^{C_1(s-t)+\int_s^t C_2(u)du}ds \\ &\leq e^{C_2^\infty - C_1 t}\mathcal{H}_N(t) + \frac{C_3}{C_1 N}e^{C_2^\infty},\end{aligned}$$

which concludes.

Open problems

There are of course a lot of problems remaining

- Adapted LSI and uniform in time propagation of chaos ? see Rozensweig-Serfaty.
- For vortex 2D case : up to now restricted to the torus... extend it to whole space ?
- Keller-Segel model ? (see Bresch-Jabin, Serfaty & al. for modulated energy, Tomasevic &al. for parabolic-parabolic case)
- Vlasov-Fokker-Planck equation with singular kernel as in Bresch-Jabin-Soler very recent paper. Quantitative and uniform in time ?
- Optimal rate *à la* Lacker for singular models ?

Open problems

There are of course a lot of problems remaining

- Adapted LSI and uniform in time propagation of chaos ? see Rozensweig-Serfaty.
- For vortex 2D case : up to now restricted to the torus... extend it to whole space ?
- Keller-Segel model ? (see Bresch-Jabin, Serfaty & al. for modulated energy, Tomasevic &al. for parabolic-parabolic case)
- Vlasov-Fokker-Planck equation with singular kernel as in Bresch-Jabin-Soler very recent paper. Quantitative and uniform in time ?
- Optimal rate *à la* Lacker for singular models ?

Thank you !