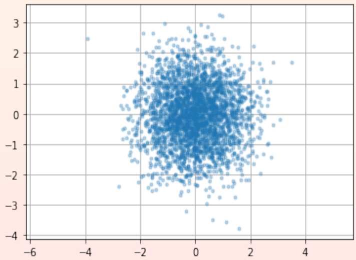
Mean-Field Control for Diffusion Aggregation system with Coulomb Interaction

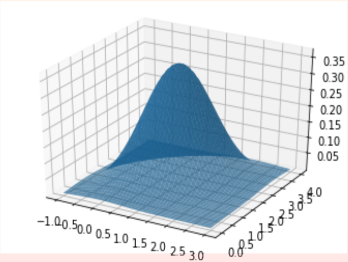
Li Chen
Universität Mannheim

Gauss Mini Workshop, 23.01.2025 Braunschweig

"Mean-field", physics and probability, applications in other sciences

- System of ODEs (or SDEs) are used to describe the dynamics of particle systems.
- Problem: Very high number of the particles, high computational costs.
- Key structure: particles are not distinguishable, interaction forces are symmetric and of order $\frac{1}{N}$
- Idea: Derive a one particle equation with average effect from the others
- Solution: The PDE for probability density. the mean-field limit.





Idea of mean field limit

A toy model, deterministic case

Mean field type N interaction particle system in dimension d with interaction force $F : \mathbb{R}^d \to \mathbb{R}^d$

$$\frac{dX_i^N}{dt} = \frac{1}{N} \sum_{j \neq i}^{N} F(X_i^N - X_j^N),$$

$$X_i^N|_{t=0} = X_i^N(0), \quad 1 \leq i \leq N, \qquad \text{are } N \text{ different } d\text{-D vectors.}$$

If F is Lipschitz, then there exists a global unique solution.

For $N \gg 1$, high computational costs. Can one take $N \to \infty$?

Idea: Empirical measure of $(X_1^N, X_2^N, \cdots, X_N^N)$,

$$\mu_N(\cdot, t) = \frac{1}{N} \sum_{i=1}^{N} \delta(\cdot - X_i^N(t)).$$

Idea of mean field limit

For any test function $\varphi \in C_0^\infty$, we have "formally",

$$\begin{aligned}
\frac{d}{dt}\int \mu_N(x,t)\varphi(x)dx &= \frac{d}{dt}\Big(\frac{1}{N}\sum_{i=1}^N \varphi(X_i^N)\Big) = \frac{1}{N}\sum_{i=1}^N \nabla\varphi(X_i^N)\cdot\frac{d}{dt}X_i^N \\
&= \frac{1}{N}\sum_{i=1}^N \nabla\varphi(X_i^N)\cdot\frac{1}{N}\sum_{j=1}^N F(X_i^N - X_j^N) \\
&= \frac{1}{N}\sum_{i=1}^N \nabla\varphi(X_i^N)\cdot\int F(X_i^N - y)\mu_N(y,t)dy \\
&= \int \mu_N(x,t)\int F(x-y)\mu_N(y,t)dy\cdot\nabla\varphi(x)dx.
\end{aligned}$$

μ_N satisfies the following transport equation in the sense of distributions

$$\partial_t \mu_N + \nabla\cdot(\mu_N(F * \mu_N)) = 0.$$

Idea of mean field limit

The corresponding characteristics satisfies the same particle system,

$$\frac{dX_i^N}{dt} = \frac{1}{N}\sum_{j\neq i}^{N} F(X_i^N - X_j^N) = \int F(X_i^N - y)\mu_N(y,t)dy, \quad X_i^N|_{t=0} = X_i^N(0).$$

If $\mu_N \to \rho$ in "some sense". The particle model

$$\frac{dX(t)}{dt} = \int F(X(t)-y)\rho(y,t)dy, \quad \rho(x,t) = X(t)\#\rho_0(x)$$

$$X|_{t=0} = X_0, \text{ for all } X_0 \in \mathbb{R}^d, \ \rho_0 \text{ a given nonnegative measure.}$$

where ρ is the push-forward measure of ρ_0 by the flow $X(t)$, i.e.

$$\rho(B) = \rho_0(X(t)^{-1}(B)), \quad \forall B \subset \mathbb{R}^d.$$

Specifically, if ρ_0 is an empirical measure, these two problems are the same. For general nonnegative measure ρ_0, the solution of

$$\partial_t \rho + \nabla\cdot(\rho(F*\rho)) = 0, \quad \rho(x,0) = \rho_0$$

should be the "right" $N \to \infty$ limit for the N particle system.

Idea for N particle SDE system

The evolution of $(X_1^N(t), X_2^N(t), \cdots, X_N^N(t)) \in \mathbb{R}^{dN}$ is given by

$$dX_i^N(t) = \frac{1}{N}\sum_{j\neq i}^{N} F(X_i^N(t) - X_j^N(t))dt + \sqrt{2}d\mathbf{B}^i(t),$$
$$X_i^N|_{t=0} = X_i^N(0), \quad N \text{ i.i.d. random variables with law } \rho_0$$

The formal mean field limit is (Mckean Vlasov)

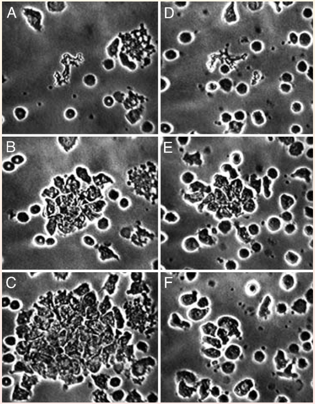
$$d\hat{X}_i^N(t) = \int_{\mathbb{R}^d} F(\hat{X}_i^N(t) - y)\rho(t,y)dy + \sqrt{2}d\mathbf{B}^i(t), \quad \rho(t,x) = law\{\hat{X}_i^N(t)\},$$
$$\hat{X}_i^N|_{t=0} = \hat{X}_i^N(0) \quad N \text{ i.i.d. random variables with law } \rho_0.$$

Ito's formula and expectation shows that ρ satisfies the nonlocal PDE

$$\partial_t\rho + \nabla\cdot(\rho(F*\rho)) = \Delta\rho, \quad \rho(x,0) = \rho_0.$$

Patlak-Keller-Segel system: the basic mathematical model for Chemotaxis. It is PDE model to describe the movement of cells in response to chemical signals.

$$\rho_t = \Delta\rho - \mathrm{div}(\rho\nabla c), \quad -\Delta c = \rho, \quad \text{or} \quad \boxed{\rho_t + \mathrm{div}(\rho\nabla\Phi * \rho) = \Delta\rho}$$



- $\Phi(x) = 2\pi \log|x|$ for $d = 2$
  $\Phi(x) = C(d)/|x|^{d-2}$ for $d \geq 3$
- The cells move towards regions of higher signal concentration.
- As the cells produce the signal substance, the movement may lead to an aggregation of cells.
- Aggregation is counter-balanced by cell diffusion.

Typical quantities of the P-K-S System

- Conservation of mass

$$\boxed{m_0(t) = \int \rho(x,t)dx = \int \rho_0(x)dx = m_0.}$$

- Entropy (Lyapunov functional) dissipation relation,

$$\text{Entropy: } H(\rho) = \int (\rho \ln \rho - \frac{\rho c}{2})dx,$$

$$\boxed{\frac{d}{dt}H(\rho) + \int \rho |\nabla \ln \rho - \nabla c|^2 dx = 0.} \Rightarrow \boxed{H(\rho) \le H(\rho_0)}$$

Key feature of the system: Global existence vs. finite time blow up. Since 1990's, Jäger and Luckhaus, Biler, Herrero, Horstmann, Medina, Nagai, Stevens, Velazquez, Winkler......

The critical mass 8π in 2-D

$$c = -\frac{1}{2\pi}\int \log\frac{1}{|x-y|}\rho(y)dy$$

Entropy: $H(\rho) = \int \rho \ln \rho dx - \frac{1}{8\pi}\int\int \rho(x,t)\rho(y,t)\log\frac{1}{|x-y|^2}dxdy.$

FACT 1: Entropy dissipation relation

$$\boxed{H(\rho) \le H(\rho_0).}$$

FACT 2: Logarithmic Hardy-Littlewood-Sobolev inequality

If $\rho \ge 0$ in L^1 and $\rho \log \rho \in L^1$, then

$$\boxed{\int \rho \log \rho dx - \frac{1}{m_0}\int\int \rho(x)\rho(y)\log\frac{1}{|x-y|^2}dxdy + C(m_0) \ge 0,}$$

where $m_0 = \int \rho(x)dx,\ C(m_0) := m_0(1 + \log \pi - \log m_0).$

A direct application is that

$$H(\rho(\cdot,t)) \geq (1-\frac{m^0}{8\pi})\int \rho\log\rho dx - \frac{m^0}{8\pi}C(m^0),$$

$$H(\rho(\cdot,t)) \geq (\frac{1}{m^0}-\frac{1}{8\pi})\int\int \rho(x,t)\rho(y,t)\log\frac{1}{|x-y|^2}dxdy - C(m^0).$$

$m_0 < 8\pi$: **global existence**, Blanchet, Dolbeault, Perthame, 2006.
Another proof: Carrillo, Chen, Liu, and Wang, based on Delort's theory of 2-D incompressible Euler equation, 2012.
$m_0 > 8\pi$: **blow up in finite time**, Dolbeault, Perthame, 2004.

Idea: Second moment, $m_2(t) := \int \frac{|x|^2}{2}\rho dx$. $m_2'(t) = 2m_0(1-\frac{m_0}{8\pi}) < 0$.

$m_0 = 8\pi$: Blanchet, Carlen, Carrillo, Masmoudi....

$$d = 2 \; \boxed{\textbf{Diffusion}} \sim \boxed{\textbf{Aggregation}}.$$

Remark $d \geq 3$ Diffusion is weaker than Aggregation, boundedness of the solution only for small initial data.

Question: Slow down or speed up the blow-up phenomena by adding control function with minimal costs? When yes, Where to add the control function?

Idea: add the control function in the chemical potential equation by adjusting the density of the cells:

$$\partial_t \rho = \Delta\rho - \nabla\cdot(\rho\chi\nabla c), \quad -\Delta c = \rho - f, \quad \rho_0 \in L^1(\mathbb{R}^d;(1+|x|^2)\,\mathrm{d}x).$$

Admissable control space $X = \left\{ f : \|f\|_{W^{1,q}(\mathbb{R}^d)\cap L^1(\mathbb{R}^d)} \le l(t),\ l(t) \in L^r(0,T), r > 1+\frac{d}{2}, q > d \right\}.$

The cost functional

$$J(\rho[f], f) = \int_0^T \|\rho[f] - z\|_{L^p(\mathbb{R}^d)}\mathrm{d}t + \int_0^T \langle f, \rho[f]\rangle \mathrm{d}t.$$

- $\rho[f]$ is the solution of K-S with control function f.
- the objective probability density $z \in L^1 \cap L^p(\mathbb{R}^d)$ with $p \in [2,\infty)$

Goal: $\exists\ \bar{f} \in X$ **s.t.**

$$J(\bar{f}) = \min_{f\in X} J(f).$$

Understand the control problem on the particle level prove the existence of $\bar{f}$ through the corresponding micro-scopic system

$$dX_i^{N,\varepsilon} = \frac{\chi}{N}\sum_{j=1}^{N} \nabla\tilde{\Phi}_\varepsilon(X_i^{N,\varepsilon} - X_j^{N,\varepsilon})\,dt - \chi\nabla\tilde{\Phi}_\varepsilon * f(X_i^{N,\varepsilon})\,dt + \sqrt{2}dW_t^i$$

$$X_i^N(0) = \xi_i,\ \text{i.i.d},\ \rho_0 \text{ is the probability density function of } \xi_i.$$

With mollification function j_ε, $\tilde{\Phi}_\varepsilon$ is given by

$$\widetilde{\Phi} = \begin{cases} -C_d\frac{1}{|x|^{d-2}}, & |x| \geq 2\varepsilon \\ -C_d(2-d)(2\varepsilon)^{1-d}|x| + C_d(1-d)(2\varepsilon)^{2-d}, & |x| < 2\varepsilon \end{cases}, \quad \widetilde{\Phi}_\varepsilon = j_\varepsilon * \widetilde{\Phi},$$

Cost functional

$$J_N(\mu_N[f], f) = \int_0^T \mathbb{E}\left(\left\|j_\varepsilon * \mu_N[f] - z\right\|_{L^p(\mathbb{R}^d)}\right)\mathrm{d}t + \int_0^T \mathbb{E}\left(\langle f, \mu_N[f]\rangle\right)\mathrm{d}t.$$

Cut-off scaling

$$\varepsilon \sim N^\beta \to 0,\ \textbf{when}\ N \to \infty.$$

Main result Let $d \geq 2$, $\rho_0 \in W^{1,q}$ satisfies the previous conditions,

- For any fixed ε and N, there exists $\widetilde{f} \in X$, i.e.

$$J_N(\mu_N[\widetilde{f}], \widetilde{f}) = \min_{f \in X} J_N(\mu_N[f], f).$$

- Let $\chi < 8\pi$ for $d = 2$ or $\left\|\rho_0\right\|_{L^{\frac{d}{2}}(\mathbb{R}^d)} \leq \Theta$ (for some small Θ) for $d \geq 3$. Then $\forall T > 0$ and $\forall\ f \in X$, there is a solution

$$\rho[f] \in L^\infty(0, T; L^1 \cap W^{1,q}).$$

- For $\varepsilon = N^{-\beta}$ with $\beta \in (0, \beta_*)$ for some small $\beta^* > 0$. Furthermore, let $f_N \in X$ be any minimizer of $J_N(\mu_N[f], f)$ and $\bar{f}$ be any weak accumulation point of it, then

$$J(\rho[\bar{f}], \bar{f}) = \min_{f \in X} J(\rho[f], f).$$

Selected Literature Review

Mean field optimal control

Fornasier and Solombrino (2014) Introduce the concept, 2nd order equation, weak convergence
Fornasier, Lisini, Orrieri, and Savara (2019) 1st order, weak convergence
Cavagnari, Lisini, Orrieri, Savara (2022)
Carrillo, Pimentel, and Voskanyan (2020)
Cesaroni and Cirant. (2021) $\cdots\cdots$

Mean field optimal control for stochastic particle models

Lacker, (2017) rigorous consistent with the limit of optimal controls of N-particle systems.
Djete, Possama, and Tan (2022): problems with common noise
Djete (2020): state dynamics depend on the joint distribution.
Averboukh, Pontryagin (2022): optimality conditions by the Pontryagin Maximum Principle
Bayraktar, Cosso, Pham, (2018), Djete, Possamai, and Tan (2022): optimality conditions by HJB equations

Selected Literature Review

Optimal Control for Keller-Segel system

P. Braz e Silva1, F. Guillan-Gonzalez, F. Perusato, M. A. Rodraguez-Bellido, (2023): parabolic and bounded domain

S. -U. Ryu, A. Yagi, (2001 JMAA): parabolic and bounded domain

Mean field limit (for non-Lipschitz potentials)

K. Oelschlager (1984,87,89): Porus medium, moderate interacting, L^2 estimate with convergence rate $o(N^{-1/2})$

A. Stevens (2000) Keller-Segel, moderate interacting

Jabin and Wang,, (2018): relative entropy

Lazarovic and Pickl (2017): Convergence in probability

Rosenzweig and Serfaty (2023): Reisz potential

S. Serfaty (2020):Coulomb potential

......

Step 1. The existence of minimizer of N-particle system

Proposition (Existence of minimizer)

For fixed N and ε, there exists $\widetilde{f} \in X$ s.t. $J_N(\widetilde{f}) = \min_{f\in X} J_N(f)$.

Idea: PDE method. Consider the high dimensional linear PDE of the joint distribution of particles $X_i^{N,\varepsilon}$, $i = 1, \cdots, N$, with factorized initial data $\rho_0^{\otimes N}$.

$$\partial_t \rho^{N,\varepsilon} - \sum_{i=1}^{N} \Delta_{x_i} \rho^{N,\varepsilon} + \sum_{i=1}^{N} \nabla_{x_i} \Big(\chi \rho^{N,\varepsilon} \big(\frac{1}{N} \sum_{j=1}^{N} \nabla \widetilde{\Phi}_\varepsilon (x_i - x_j) - \nabla \widetilde{\Phi}_\varepsilon * f(x_i) \big) \Big) = 0$$

Cost function can be reformulated without probability language.

$$J_N(\mu_N[f], f) = \int_0^T \int_{\mathbb{R}^{dN}} \Big\| \frac{1}{N} \sum_{i=1}^{N} j_\varepsilon(\cdot - x_i) - z(\cdot) \Big\|_{L^p(\mathbb{R}^d)} \rho^{N,\varepsilon}[f](t, x_1, \cdots, x_N) \mathrm{d}X_N \mathrm{d}t$$
$$+ \int_0^T \int_{\mathbb{R}^d} f(t,x) \rho^{N,\varepsilon;1}[f](t,x) \mathrm{d}x \mathrm{d}t,$$

Compactness argument of the minimizing sequence.
Key technical steps are uniform estimate of the PDE solution, then apply Aubin-Lions compact embedding lemma.

Key Idea to connect the particle system and PDE

Introducing the intermediate problem (ε-problems)

$$\partial_t \rho^\varepsilon - \Delta \rho^\varepsilon = \nabla \cdot (\chi \rho^\varepsilon \nabla \widetilde{\Phi}_\varepsilon * (\rho^\varepsilon - f)),$$

The corresponding McKean-Vlasov equation is

$$\begin{aligned} d\bar{X}_i^\varepsilon =& \chi \nabla \widetilde{\Phi}_\varepsilon * (\rho^\varepsilon - f)(\bar{X}_i^\varepsilon)\, dt + \sqrt{2}\, dW_t^i. \\ X_i^\varepsilon(0) =& \xi_i, \text{ i.i.d, } \rho_0 \text{ is the probability density function of } \xi_i. \end{aligned}$$

Step 2. Key results on the PDE level

- For any fixed $f \in X$, we have $\rho[f], \rho^\varepsilon[f] \in L^\infty(0,T; L^1 \cap L^\infty)$ and

$$\|\rho[f] - \rho^\varepsilon[f]\|_{L^\infty(0,T;L^1 \cap L^\infty)} \leq C\varepsilon.$$

- If $f_N \rightharpoonup f$ in X, then there is a convergence subsequence of $\rho[f_N]$ such that

$$\rho[f_{N_k}] \rightharpoonup \rho[f] \text{ in } L^1(0,T; L^p(\mathbb{R}^d)), \quad \rho[f_{N_k}] \to \rho[f] \text{ in } L^2(0,T; L^1(\mathbb{R}^d)).$$

Step 3. Mean field limit for given control with $\varepsilon = N^{-\beta}$

- Convergence in probability
  $\forall \alpha \in (0, \frac{1}{2})$, $\exists \beta_* > 0$, s.t. $\forall \beta \in (0, \beta_*)$, it holds for $\varepsilon = N^{-\beta}$

$$\mathbb{P}(\mathcal{A}_\alpha) = \mathbb{P}\Big(\omega \in \Omega : \max_{1 \le i \le N} \Big| X_i^{N,\varepsilon}[f] - \overline{X}_i^{\varepsilon}[f] \Big| \ge N^{-\alpha}\Big) \le \frac{C(\gamma)}{N^\gamma}, \quad \forall \gamma > 0.$$

- Mean field limit
  For any fixed control function $\widetilde{f}_N$, let $\rho^{N,\varepsilon;1}[\widetilde{f}_N]$ be the 1-particle marginal of $\rho^{N,\varepsilon}[\widetilde{f}_N]$ then it holds

$$\|\rho^{N,\varepsilon;1}[\widetilde{f}_N] - \rho[\widetilde{f}_N]\|_{L^\infty(0,T;L^1(\mathbb{R}^d))} \le \frac{C(t)}{N^\beta}.$$

Idea: Relative entropy estimate for the moderate interacting system. Using convergence in probability to close the relative entropy inequality.

Step 4. Limit of cost function for given control $\widetilde{f}_N \in X$

$$\lim_{N\to\infty} \left(J_N(\widetilde{f}_N) - J(\widetilde{f}_N)\right) = 0.$$

Remember

$$J_N(\mu_N[f], f) = \int_0^T \mathbb{E}(\|j_\varepsilon * \mu_N[f] - z\|_{L^p(\mathbb{R}^d)})\,\mathrm{d}t + \int_0^T \mathbb{E}(\langle f, \mu_N[f]\rangle)\,\mathrm{d}t$$

$$J(\rho[f], f) = \int_0^T \|\rho[f] - z\|_{L^p(\mathbb{R}^d)}\mathrm{d}t + \int_0^T \langle f, \rho[f]\rangle \mathrm{d}t.$$

Idea: $|J_N(\widetilde{f}_N) - J(\widetilde{f}_N)| \le A_1 + A_2$.

$$\begin{aligned} A_1 \le & \int_0^T \mathbb{E}\Big(\|j_\varepsilon * \mu_N[\widetilde{f}_N] - j_\varepsilon * \overline{\mu}_N[\widetilde{f}_N]\|_{L^p(\mathbb{R}^d)}\Big)\mathrm{d}t \\ & + \int_0^T \mathbb{E}\Big(\|j_\varepsilon * \overline{\mu}_N[\widetilde{f}_N] - j_\varepsilon * \rho^\varepsilon[\widetilde{f}_N]\|_{L^p(\mathbb{R}^d)}\Big)\mathrm{d}t \\ & + \int_0^T \|j_\varepsilon * \rho^\varepsilon[\widetilde{f}_N] - \rho^\varepsilon[\widetilde{f}_N]\|_{L^p(\mathbb{R}^d)}\mathrm{d}t + \int_0^T \|\rho^\varepsilon[\widetilde{f}_N] - \rho[\widetilde{f}_N]\|_{L^p(\mathbb{R}^d)}\mathrm{d}t \end{aligned}$$

$$\begin{aligned} A_2 \le & \int_0^T \int_{\mathbb{R}^d} \left|\widetilde{f}_N \rho^{N,\varepsilon;1}[\widetilde{f}_N] - \widetilde{f}_N \rho[\widetilde{f}_N]\right| \mathrm{d}x\,\mathrm{d}t \\ \le & \|\widetilde{f}_N\|_{L^1(0,T;L^\infty(\mathbb{R}^d))} \|\rho^{N,\varepsilon;1}[\widetilde{f}_N] - \rho[\widetilde{f}_N]\|_{L^\infty(0,T;L^1(\mathbb{R}^d))} C N^{-\beta} \to 0. \end{aligned}$$

where in A_2 we used the propagation of chaos result in strong L^1 sense.

Convergence for A_1

The third and forth terms can be easily proved converging to 0. The second term can be controlled by $\|j_\varepsilon\|_{L^\infty} N^{-1} = N^{2\beta-1}$ by using the law of large numbers. Use the convergence of probability to show that convergence of the first term :

$$
\begin{aligned}
A_{11} \leq & T^{\frac{1}{p}} \int_0^T \left| \mathbb{E}\left(\left\| j_\varepsilon * \mu_N[\widetilde{f}_N] - j_\varepsilon * \overline{\mu}_N[\widetilde{f}_N] \right\|_{L^\infty(\mathbb{R}^d)}^{\frac{p-1}{p}} \right) \right| \mathrm{d}t \\
\leq & C(T) \int_0^T \mathbb{E}\left((\mathbb{I}_{\mathcal{A}_\alpha} + \mathbb{I}_{\mathcal{A}_\alpha^c}) \left\| \frac{1}{N} \sum_{i=1}^N (j_\varepsilon(\cdot - X_i^{N,\varepsilon}[\widetilde{f}_N]) - j_\varepsilon(\cdot - \overline{X}_i^\varepsilon[\widetilde{f}_N])) \right\|_{L^\infty}^{\frac{p-1}{p}} \right) \mathrm{d}t \\
\leq & C(T) \left(\|\nabla j_\varepsilon\|_{L^\infty(\mathbb{R}^d)}^{\frac{p-1}{p}} N^{-\frac{(p-1)\alpha}{p}} + \|j_\varepsilon\|_{L^\infty(\mathbb{R}^d)}^{\frac{p-1}{p}} N^{-\frac{(p-1)\gamma}{p}} \right) \\
\leq & C(T) \left(N^{\frac{p-1}{p}((d+1)\beta-\alpha)} + N^{\frac{p-1}{p}(d\beta-\gamma)} \right) \leq C(T) N^{\frac{p-1}{p}((d+1)\beta-\alpha)} \to 0.
\end{aligned}
$$

where $\mathbb{P}\left(\mathcal{A}_\alpha = \left\{ \omega \in \Omega : \max_{1\leq i\leq N} |X^{N,\varepsilon}[\widetilde{f}_N] - \bar{X}^\varepsilon[\widetilde{f}_N]| \geq N^{-\alpha} \right\} \right) < CN^{-\gamma}$.

Step 5. Idea of Γ-convergence

Let f_N be the minimizer of $J_N(f)$, which has a convergent subsequence

$$f_{N_k} \to \bar{f} \text{ weakly in } L^r(L^1 \cap W^{1,q}).$$

we can prove the following two facts

- $\liminf_{k\to\infty} J_{N_k}(f_{N_k}) \geq J(\bar{f})$. (done by lower semicontinuity of norm, and the L^1 strong convergence in mean field limit discussion)
- $\lim_{N\to\infty} J_N(\bar{f}) = J(\bar{f})$. (Limit of cost function for given control)

Then it follows

$$\limsup_{N\to\infty} \min_{f\in X} J_N(f) \leq \lim_{N\to\infty} J_N(\bar{f}) = J(\bar{f}) \leq \liminf_{k\to\infty} J_{N_k}(f_{N_k}) \leq \liminf_{k\to\infty} \min_{f\in X} J_{N_k}(f).$$

Therefore we have $\lim_{k\to\infty} J_{N_k}(f_{N_k}) = \lim_{N_k\to\infty} \min_{f\in X} J_{N_k}(f) = J(\bar{f})$.

Show that $\bar{f}$ is a minimizer of $J(f)$ in X

Since f_{N_k} is the minimizer of J_{N_k}, we have proved that

$$\lim_{k\to\infty} J_{N_k}(f_{N_k}) = J(\bar{f}).$$

On the other hand, for any $\hat{f} \in X$, since $J_N(\hat{f}) \geq J_N(f_N)$ and

$$\lim_{N\to\infty} J_N(\hat{f}) = J(\hat{f}).$$

we obtain automatically that $J(\hat{f}) \geq J(\bar{f})$, which means that $\bar{f}$ is a minimizer of $J(f)$ in X.

Attachment: Relative entropy

A key tool: Relative entropy method Consider the high dimensional PDEs with factorized initial data $\rho_0^{\otimes N}$.

$$\partial_t \rho^{N,\varepsilon} - \sum_{i=1}^{N} \Delta_{x_i} \rho^{N,\varepsilon} + \sum_{i=1}^{N} \nabla_{x_i} \cdot \Big(\chi \rho^{N,\varepsilon} \big(\frac{1}{N} \sum_{j=1}^{N} \nabla \widetilde{\Phi}_\varepsilon (x_i - x_j) - \nabla \widetilde{\Phi}_\varepsilon * \widetilde{f}_N(x_i) \big) \Big) = 0$$

$$\partial_t \rho^{\otimes N} - \sum_{i=1}^{N} \Delta_{x_i} \rho^{\otimes N} + \sum_{i=1}^{N} \nabla_{x_i} \cdot \Big(\chi \rho^{\otimes N} \big(\nabla \Phi * \rho(x_i) - \nabla \Phi * \widetilde{f}_N f(x_i) \big) \Big) = 0$$

The relative entropy of N-particle system $\mathcal{H}$ defined by

$$\mathcal{H}(\rho^{N,\varepsilon} | \rho^{\otimes N}) = \frac{1}{N} \mathcal{H}_N(\rho^{N,\varepsilon} | \rho^{\otimes N}) = \frac{1}{N} \int_{\mathbb{R}^{dN}} \frac{\rho^{N,\varepsilon}}{\rho^{\otimes N}} \log \frac{\rho^{N,\varepsilon}}{\rho^{\otimes N}} \cdot \rho^{\otimes N} \, dX_N.$$

Proposition (Relative entropy estimate (Using convergence in probability))

If $\varepsilon = N^{-\beta}$, $0 < \beta < \beta^*$, *then*

$$\mathcal{H}(\rho^{N,\varepsilon} | \rho^{\otimes N}) \le \frac{C(t)}{N^{2\beta}},$$

Compute the time evolution of the relative entropy, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(\rho^{N,\varepsilon}|\rho^{\otimes N}) + \frac{1}{2N}\int_{\mathbb{R}^{dN}}\sum_{i=1}^{N}\left|\nabla_{x_i}\log\left(\frac{\rho^{N,\varepsilon}}{\rho^{\otimes N}}\right)\right|^2\rho^{N,\varepsilon}\,\mathrm{d}x_1\cdots\mathrm{d}x_N$$

$$\leq \chi\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\Phi*\rho(t,X_i^{N,\varepsilon}) - \frac{1}{N}\sum_{j=1}^{N}\nabla\widetilde{\Phi}_\varepsilon(X_i^{N,\varepsilon}-X_j^{N,\varepsilon})\Big|^2\Big)$$

$$+\chi\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\Phi*f(t,X_i^{N,\varepsilon}) - \nabla\widetilde{\Phi}_\varepsilon*f(t,X_i^{N,\varepsilon})\Big|^2\Big)$$

$$:=\chi M_1 + \chi M_2.$$

Estimate for M_2 is relatively easy,

$$
\begin{aligned}
M_2 \leq & \|(\nabla\widetilde{\Phi}_\varepsilon - \nabla\Phi_\varepsilon) * \widetilde{f}_N\|_{L^\infty}^2 + \|(\nabla\Phi_\varepsilon - \nabla\Phi) * \widetilde{f}_N\|_{L^\infty}^2 \\
\leq & \varepsilon^2 \|\widetilde{f}_N\|_{L^\infty}^2 + \varepsilon^2 \|D^2\Phi * \widetilde{f}_N\|_{L^\infty}^2 \leq \varepsilon^2 \|\widetilde{f}_N\|_{W^{1,q}}^2 \leq \varepsilon^2 l^2(t).
\end{aligned}
$$

Use convergence in probability to close the estimate of M_1

$$
\begin{aligned}
M_1 \leq & \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\Phi * \rho(t, X_i^{N,\varepsilon}) - \nabla\widetilde{\Phi}_\varepsilon * \rho^\varepsilon(t, X_i^{N,\varepsilon})\Big|^2\Big) \\
& + \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\widetilde{\Phi}_\varepsilon * \rho^\varepsilon(t, X_i^{N,\varepsilon}) - \nabla\widetilde{\Phi}_\varepsilon * \rho^\varepsilon(t, \overline{X}_i^\varepsilon)\Big|^2\Big) \\
& + \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\widetilde{\Phi}_\varepsilon * \rho^\varepsilon(t, \overline{X}_i^\varepsilon) - \frac{1}{N}\sum_{j=1}^{N}\nabla\widetilde{\Phi}_\varepsilon(\overline{X}_i^\varepsilon - \overline{X}_j^\varepsilon)\Big|^2\Big) \\
& + \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\frac{1}{N}\sum_{j=1}^{N}\big(\nabla\widetilde{\Phi}_\varepsilon(\overline{X}_i^\varepsilon - \overline{X}_j^\varepsilon) - \nabla\widetilde{\Phi}_\varepsilon(X_i^{N,\varepsilon} - X_j^{N,\varepsilon})\big)\Big|^2\Big) := \sum_{i=1}^{4} M_{1i}.
\end{aligned}
$$

M_{11} can be estimated by pure PDE analysis,

$$
\begin{aligned}
M_{11} \leq & \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\Phi * \rho(t, X_i^{N,\varepsilon}) - \nabla\Phi_\varepsilon * \rho(t, X_i^{N,\varepsilon})\Big|^2\Big) \\
& + \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\Phi_\varepsilon * \rho(t, X_i^{N,\varepsilon}) - \nabla\widetilde{\Phi}_\varepsilon * \rho(t, X_i^{N,\varepsilon})\Big|^2\Big) \\
& + \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\Big|\nabla\widetilde{\Phi}_\varepsilon * \rho(t, X_i^{N,\varepsilon}) - \nabla\widetilde{\Phi}_\varepsilon * \rho^\varepsilon(t, X_i^{N,\varepsilon})\Big|^2\Big) \\
\leq & \|\nabla(\Phi - \Phi_\varepsilon) * \rho\|_{L^\infty(\mathbb{R}^d)}^2 + \|\nabla(\Phi_\varepsilon - \widetilde{\Phi}_\varepsilon) * \rho\|_{L^\infty(\mathbb{R}^d)}^2 + \|\nabla\widetilde{\Phi}_\varepsilon * (\rho - \rho^\varepsilon)\|_{L^\infty(\mathbb{R}^d)}^2 \\
\leq & 2\varepsilon^2\|\rho\|_{L^\infty(0,T;W^{1,q}(\mathbb{R}^d))}^2 + \|\rho - \rho^\varepsilon\|_{L^\infty(0,T;L^1(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d))}^2 \leq C\varepsilon^2 = \frac{C}{N^{2\beta}}.
\end{aligned}
$$

By splitting $\Omega = \mathcal{A}_\alpha \cup \mathcal{A}_\alpha^c$, we have that for $\beta \in (0, \beta_*)$ and $\gamma = 2(\alpha - \beta)$ it holds

$$
\begin{aligned}
M_{12} + M_{14} \leq & C \frac{\|\nabla^2 \widetilde{\Phi}_\varepsilon\|^2_{L^\infty(\mathbb{R}^d)}}{N^{2\alpha}} + C(\gamma) \frac{\|\nabla \widetilde{\Phi}_\varepsilon\|^2_{L^\infty(\mathbb{R}^d)}}{N^\gamma} \\
\leq & C(N^{2(\beta d - \alpha)} + N^{2\beta(d-1)-\gamma}) = N^{2(\beta d - \alpha)}.
\end{aligned}
$$

The term M_{13} can be estimated by directly applying the law of large number, namely

$$
M_{13} \leq C \frac{\|\nabla \widetilde{\Phi}_\varepsilon\|^2_{L^\infty(\mathbb{R}^d)}}{N} \leq C \frac{\varepsilon^{2(1-d)}}{N} = CN^{2\beta(d-1)-1}.
$$

The L^1 convergence follows from:

- Relative entropy estimate:

$$\mathcal{H}(\rho^{N,\varepsilon}|\rho^{\otimes N}) \leq C \min\{N^{-\beta}, N^{-\frac{2\alpha}{3}}\} \to 0.$$

- Super-Additivity of Relative Entropy:

$$\mathcal{H}(\rho^{N,\varepsilon}|\rho^{\otimes N}) = \frac{1}{N}\mathcal{H}_N(\rho^{N,\varepsilon}|\rho^{\otimes N}) \geq \frac{1}{2k}\mathcal{H}_k(\rho^{N,\varepsilon;k}|\rho^{\otimes k})$$

- Csiszàr-Kullback-Pinsker:

$$\|f-g\|^2_{L^1(\mathbb{R}^d)} \leq 2\mathcal{H}(f|g)$$

Therefore, we obtain

$$\|\rho^{N,\varepsilon;1} - \rho\|_{L^1(\mathbb{R}^d)} \to 0.$$

THANKS FOR YOUR ATTENTION!