Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions

Zimo Hao

A joint work with Michael Röckner and Xicheng Zhang

Bielefeld University

arXiv:2204.07952

14 May, 2022

15th Oxford-Berlin Young Researchers Meeting on Applied Stochastic Analysis

Table of contents

1. N-particle systems and propagation of chaos
   - N-particle systems and McKean-Vlasov equations
   - Propagation of chaos
   - Moderate case and density dependent SDE
2. Well-posedness
   - Motivation
   - Mixed L^p space with permutation
   - Main result for well-posedness
3. Main results for propagation of chaos

1 N-particle systems and propagation of chaos
- N-particle systems and McKean-Vlasov equations
- Propagation of chaos
- Moderate case and density dependent SDE

2 Well-posedness
- Motivation
- Mixed L^p space with permutation
- Main result for well-posedness

3 Main results for propagation of chaos

N-particle systems

- Consider the following N-particle systems on $\mathbb{R}^d$:

$$\mathrm{d}X_t^{N,i} = F\Big(t, X_t^{N,i}, \frac{1}{N}\sum_{j=1}^{N}\varphi(X_t^{N,i} - X_t^{N,j})\Big)\mathrm{d}t + \sigma(t, X_t^{N,i})\mathrm{d}B_t^i, \quad (\mathcal{N})$$

  where $i = 1, 2, .., N$,

  - F: Some nonlinear force;
  - φ: interaction kernel;
  - B^i: independent Brownian motions on $\mathbb{R}^d$ (random phenomena).

McKean-Vlasov equations

- When F, φ and σ are smooth, the solution of the N-particle systems $X^{N,i}$ convergences to the solution to the following McKean-Vlasov SDE, which is also called Distribution Dependent SDE:

$$\mathrm{d}X_t = F\Big(t, X_t, \int_{\mathbb{R}^d} \varphi(X_t - y)\mu_t(\mathrm{d}y)\mathrm{d}t\Big) + \sigma(t, X_t)\mathrm{d}B_t, \qquad (1)$$

  where μ_t is the distribution of X_t and B_t is a standard BM.

McKean-Vlasov equations

- When F, φ and σ are smooth, the solution of the N-particle systems $X^{N,i}$ convergences to the solution to the following McKean-Vlasov SDE, which is also called Distribution Dependent SDE:

$$\mathrm{d}X_t = F\Big(t, X_t, \int_{\mathbb{R}^d} \varphi(X_t - y)\mu_t(\mathrm{d}y)\mathrm{d}t\Big) + \sigma(t, X_t)\mathrm{d}B_t, \tag{1}$$

  where μ_t is the distribution of X_t and B_t is a standard BM.

- Well-known results for well-posedness

  i) Linear growth F and φ:
  Mishura-Veretennikov (arXiv-16), Lacker (arXiv-21).

  ii) $L^q_T L^p$ interaction:
  Röckner-Zhang (Bernoulli-21), Zhao (arXiv-20)
  $\|f\|_{L^q_T L^p} := \left(\int_0^T \|f(t)\|_p^q \mathrm{d}t \right)^{1/q}$

Propagation of chaos

- Denote by $P_t^{N,k}$ and P_t the distribution of $(X_t^{N,1},...,X_t^{N,k})$ for $k = 1,2,..,N$ and X_t respectively. It is natural to ask whether we have

$$P_t^{N,k} \to P_t^{\otimes k}, \quad \text{if} \quad P_0^{N,k} \to P_0^{\otimes k} \text{ (which is called } P_0\text{-chaotic).}$$

  This is called propagation of chaos, which originally goes as far back as Maxwell and Boltzmann, and was formalized by Kac in 1950s. Nowadays, it become more and more popular and there are many results on it. We only list some of them with $F(t,x,r) = r$ and singular (non-smooth) interaction φ.

  - Jabin-Wang (JFA-16 & Invent-18): Assume φ is bounded (kinetic case) and $\varphi, \mathrm{div}\varphi \in W^{-1,\infty}$(in $\mathbb{T}^d$),

$$\|P_t^{N,k} - P_t^{\otimes k}\|_{var} \le C\sqrt{\frac{k}{N}}.$$

  - Lacker (EJP-18 & arXiv-21) Assume φ is bounded,

$$\|P_t^{N,k} - P_t^{\otimes k}\|_{var} \le C\frac{k}{N}.$$

Strong propagation of chaos

- Apart from the above convergence results for $P^{N,k}$, assuming that $F(t,x,r) = r$ and φ is Lipschitz, Sznitman in his famous lecture in 1991 showed us the following results:

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |X_t^{N,i} - X_t^i|\Big) \le C\Big(\sqrt{\frac{1}{N}} + \mathbb{E}|X_0^{N,i} - X_0^i|\Big),$$

  where X_t^i is the solution to (1) driven by BM B_t^i.
- We call this type of convergence the strong convergence of propagation of chaos.

Moderate case

- When $\varphi(\cdot) = \varphi_N(\cdot) = \varepsilon_N^{-d}\phi(\varepsilon_N^{-1}\cdot)$, which is called moderately interacting kernel, and $\varepsilon_N \to 0$ as $N \to \infty$, we rewrite our N-particle systems as follow:

$$\mathrm{d}X_t^{N,i} = F\Big(t, X_t^{N,i}, \frac{1}{N}\sum_{j=1}^{N} \varphi_N(X_t^{N,i} - X_t^{N,j})\Big)\mathrm{d}t + \sigma(t, X_t^{N,i})\mathrm{d}B_t^i. \quad (\mathcal{M})$$

- At this time, $\varphi_N \to \delta$. We expect that the limit equation is the following Density Dependent SDE (is also called McKean-Vlasov SDE of Nemytskii-type):

$$\mathrm{d}X_t = F(t, X_t, \rho_t(X_t))\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t, \quad (2)$$

  where ρ_t stands for the density of X_t.
- Here ρ_t solves the following nonlinear Fokker-Planck equation:

$$\partial_t\rho_t = \partial_i\partial_j(a_{ij}\rho_t) + \mathrm{div}(F(\rho_t)\rho_t).$$

Well-known results and question

- (Oelschläger, PTRF-85)
  $F(t,\cdot,\cdot)$ and $rF(r)$ are Lipschitz, $\varphi = W * W$ with some $W \in H^\alpha$
  $\Rightarrow$ Weak convergence when $\varepsilon_N = N^{\beta/d}$ with $\beta \in (0,1)$.
- (Jourdain-Méléard, AIHP-98)
  F, ϕ and σ are smooth $\Rightarrow$ strong convergence rate of propagation of chaos when $\varepsilon_N \asymp (\ln N)^\delta$ with some $\delta > 0$.
- However, Lipschitz assumptions on F, ϕ and σ are too strong in practice. In fact, most of the interesting physical models have bounded measurable or even singular interaction kernels.
- (Question:) Can we establish the strong convergence of propagation of chaos with singular (L^∞ and $L^q_T L^p$) interaction both for classical one ($\mathcal{N}$) and moderate one ($\mathcal{M}$)?

1 N-particle systems and propagation of chaos
- N-particle systems and McKean-Vlasov equations
- Propagation of chaos
- Moderate case and density dependent SDE

2 Well-posedness
- Motivation
- Mixed L^p space with permutation
- Main result for well-posedness

3 Main results for propagation of chaos

Well-posedness for N-particle systems

- Before consider the strong convergence of propagation of chaos, it is nature to ask whether there is a strong well-posedness for both N-particle systems and limit McKean-Vlasov SDE.
- Let $F(t,x,r) = r$ and $\varphi \in L^p$. Consider the following N-particle systems:

$$\mathrm{d}X_t^{N,i} = \frac{1}{N}\sum_{j=1}^{N}\varphi(X_t^{N,i} - X_t^{N,j})\mathrm{d}t + \mathrm{d}B_t^i.$$

- By Krylov-Röckner (PTRF-05), we need

$$\varphi^i(x) := \frac{1}{N}\sum_{j=1}^{N}\varphi(x_i - x_j) \in L^p \text{ with } p > Nd.$$

- (Question) How to obtain strong well-posedness of N-particle systems with $\varphi \in L^p$ and $p > d$.

Property of interaction kernel φ^i

- Notice that for $\mathbf{p_0} = (p, \infty, \infty, ..., \infty)$, $\varphi^1 \in L^{\mathbf{p_0}}$, where

$$\|f\|_{L^{\mathbf{p}}} := \Big(\int_{\mathbb{R}^d} \Big(\cdot \cdot \big(\int_{\mathbb{R}^d} |f(x)|^{p_1} \mathrm{d}x_1 \big)^{p_1/p_2} \cdot \cdot \Big)^{p_N/p_{N-1}} \mathrm{d}x_N \Big)^{1/p_N}.$$

- Ling-Xie (POTA 2021) Strong well-posedness for the following SDE

$$\mathrm{d}X_t = F(X_t)\mathrm{d}t + \mathrm{d}W_t$$

where $F \in L^{\mathbf{p}}$ with $\frac{d}{p_1} + \frac{d}{p_2} + \cdot \cdot + \frac{d}{p_N} < 1$.

Property of interaction kernel φ^i

- Notice that for $\mathbf{p_0} = (p, \infty, \infty, ..., \infty)$, $\varphi^1 \in L^{\mathbf{p_0}}$, where

$$\|f\|_{L^{\mathbf{p}}} := \Big(\int_{\mathbb{R}^d} \Big(\cdot \cdot (\int_{\mathbb{R}^d} |f(x)|^{p_1} \mathrm{d}x_1)^{p_1/p_2} \cdot \cdot \Big)^{p_N/p_{N-1}} \mathrm{d}x_N \Big)^{1/p_N}.$$

- Ling-Xie (POTA 2021) Strong well-posedness for the following SDE

$$\mathrm{d}X_t = F(X_t)\mathrm{d}t + \mathrm{d}W_t$$

where $F \in L^{\mathbf{p}}$ with $\frac{d}{p_1} + \frac{d}{p_2} + \cdots + \frac{d}{p_N} < 1$.

- However, notice that $(\varphi^1, ..., \varphi^N) \notin L^{\mathbf{p_0}}$ because of permutation. Actually, we only have

$$\sup_{x_j, j \neq i} \Big(\int_{\mathbb{R}^d} |\varphi^i(..., x_i, ...)|^p \mathrm{d}x_i \Big)^{1/p} < \infty.$$

And $L^{p_1}L^{p_2} \notin L^{p_2}L^{p_1}$.

Mixed L^p space with permutation

- For multi-index $\mathbf{p} = (p_1, \cdots, p_d) \in [1, \infty]^N$ and any permutation $\mathbf{x} = (x_{i_1}, ..., x_{i_N})$, the mixed $L^{\mathbf{p}}_{\mathbf{x}}$-space is defined by

$$\|f\|_{L^{\mathbf{p}}_{\mathbf{x}}} := \left(\left(\int_{\mathbb{R}^d} \cdots \left(\int_{\mathbb{R}^d} |f(x_1, \cdots, x_d)|^{p_d} \mathrm{d}x_{i_1} \right)^{\frac{p_{d-1}}{p_d}} \cdots \right)^{\frac{p_1}{p_2}} \mathrm{d}x_{i_N} \right)^{\frac{1}{p_1}}.$$

- Note that for general $\mathbf{x} \neq \mathbf{x}'$ and $\mathbf{p} \neq \mathbf{p}'$,

$$L^{\mathbf{p}'}_{\mathbf{x}} \neq L^{\mathbf{p}}_{\mathbf{x}} \neq L^{\mathbf{p}}_{\mathbf{x}'}.$$

  For multi-indices $\mathbf{p} \in [1, \infty]^N$, we shall use the following notations:

$$\left|\frac{d}{\mathbf{p}}\right| = \sum_{i=1}^{d} \frac{d}{p_i}.$$

- $\varphi^i \in L^{\mathbf{p_0}}_{\mathbf{x}_i}$ with $\mathbf{x}_i = (x_i, x_1, ..., x_N)$.

Main result
Assumptions

- Let indices $(q_i, \mathbf{p}_i)$, $i = 0, 1, .., N$ satisfy $\frac{2}{q_i} + \left|\frac{d}{\mathbf{p}_i}\right| < 1.$
- Suppose that $\nabla\sigma \in L_T^{q_0}(L_{\mathbf{x}_0}^{\mathbf{p}_0})$,

$$\kappa_0^{-1}|\xi| \leqslant |\sigma(t,x)\xi| \leqslant \kappa_0|\xi|, \quad \forall x, \xi \in \mathbb{R}^d$$

- For any $T > 0$ and $i = 1, \cdots, N$, there are permutations $\mathbf{x}_i$ such that

$$\sup_{\mu \in C([0,T];\mathcal{P}(\mathbb{R}^d))} \|\sup_{r\geqslant 0} |b^i(\cdot,\cdot,r,\mu.)|\|_{L_T^{q_i}(L_{\mathbf{x}_i}^{\mathbf{p}_i})} \leqslant \kappa_1,$$

and for some $h_i \in L_T^{q_i}(L_{\mathbf{x}_i}^{\mathbf{p}_i})$ and for all $t, x \in [0,T] \times \mathbb{R}^d$, $r, r' \geqslant 0$ and $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$,

$$|b^i(t,x,r,\mu) - b^i(t,x,r',\nu)| \leqslant h_i(t,x)\big(|r - r'| + \|\mu - \nu\|_{\text{var}}\big).$$

Theorem 1 (H., Röckner and Zhang)

Under the assumptions, for any probability measure $\mu_0(\mathrm{d}x) = \rho_0(x)\mathrm{d}x$ with $\rho_0 \in L^\infty$, there is a unique strong solution to the following DDSDE with initial distribution μ_0

$$\mathrm{d}X_t = b(t, X_t, \rho_t(X_t), \mu_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t,$$

where $\rho_t(x)$ and μ_t are the density and dietribution of X_t respectively and $b(t, x, r, \mu) = (b^1, ..., b^N) : \mathbb{R}_+ \times \mathbb{R}^{Nd} \times \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{Nd}$.

If b does not depend on the density variable, then we can drop the assumption $\mu_0(\mathrm{d}x) = \rho_0(x)\mathrm{d}x$.

Theorem 1 (H., Röckner and Zhang)

Under the assumptions, for any probability measure $\mu_0(\mathrm{d}x) = \rho_0(x)\mathrm{d}x$ with $\rho_0 \in L^\infty$, there is a unique strong solution to the following DDSDE with initial distribution μ_0

$$\mathrm{d}X_t = b(t, X_t, \rho_t(X_t), \mu_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t,$$

where $\rho_t(x)$ and μ_t are the density and dietribution of X_t respectively and $b(t, x, r, \mu) = (b^1, ..., b^N) : \mathbb{R}_+ \times \mathbb{R}^{Nd} \times \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{Nd}$.

If b does not depend on the density variable, then we can drop the assumption $\mu_0(\mathrm{d}x) = \rho_0(x)\mathrm{d}x$.

- Strong well-posedness for N-particle systems for some $L^q_T L^p$ interaction kernels.
- Strong well-posedness for limit equation (both distribution dependent case and density dependent case) by taking $N = 1$. This extends the results of H.-Röckner-Zhang (JDE 2021) and Wang (arXiv-2021).

1 N-particle systems and propagation of chaos
- N-particle systems and McKean-Vlasov equations
- Propagation of chaos
- Moderate case and density dependent SDE

2 Well-posedness
- Motivation
- Mixed L^p space with permutation
- Main result for well-posedness

3 Main results for propagation of chaos

Assumptions

- Let $\frac{2}{q} + \frac{d}{p} < 2$.

($\mathbf{H}^\sigma$) There are $\kappa_0 \geqslant 1$ and $\gamma_0 \in (0, 1]$ such that,

$$\kappa_0^{-1}|\xi| \leqslant |\sigma(t,x)\xi| \leqslant \kappa_0|\xi|, \ \|\sigma(t,x) - \sigma(t,x')\|_{HS} \leqslant \kappa_0|x - x'|^{\gamma_0},$$

where $\|\cdot\|_{HS}$ is the usual Hilbert-Schmidt norm of a matrix. Moreover, $\nabla\sigma \in L_T^q L^p$.

($\mathbf{H}^b$) Suppose that $\phi(0) = 0$ and for some measurable $h : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ and $\kappa_1 > 0$,

$$|F(t,x,r)| \leqslant h(t,x) + \kappa_1|r|, \ \ |F(t,x,r) - F(t,x,r')| \leqslant \kappa_1|r - r'|,$$

and for some $p_0 > d$

$$\|h\|_{L_T^q L^p} + \|\varphi\|_{p_0} \leqslant \kappa_1.$$

Main results

Theorem 2 (Strong convergence)

Let $T > 0$. Under the above assumptions, suppose that $P_0^{N,N}$ is symmetric and P_0-chaotic, and

$$\lim_{N\to\infty} \mathbb{E}|X_0^{N,1} - X_0|^2 = 0,$$

then for any $\gamma \in (0, 1)$,

$$\lim_{N\to\infty} \mathbb{E}\left(\sup_{t\in[0,T]} |X_t^{N,1} - X_t|^{2\gamma}\right) = 0.$$

Main results

Theorem 3 (Strong convergence rate)

Let $T > 0$. Assume the same assumptions as the above theorem. Let

$$\kappa_2 := \sup_N \mathcal{H}\left(P_0^{N,N} | P_0^{\otimes N}\right) < \infty. \quad (3)$$

Also assume that h and ϕ are bounded measurable. Then for any $\gamma \in (0,1)$, there is a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t^{N,1} - X_t|^{2\gamma}\right) \leqslant C\left(\mathbb{E}|X_0^{N,1} - X_0|^2 + \frac{\kappa_2}{N}\right)^\gamma.$$

Strong convergence for moderately case

Theorem 4

Let $T > 0$. Suppose that $(\mathbf{H}^\sigma)$ holds, and

$$|F(t,x,r)| \leqslant \kappa_1, \ \ |F(t,x,r) - F(t,x,r')| \leqslant \kappa_1 |r - r'|,$$

and

$$\varphi(x) = \varphi_{\varepsilon_N}(x) = \varepsilon_N^{-d} \phi(x/\varepsilon_N),$$

where ϕ is a bounded probability density function in $\mathbb{R}^d$ with support in the unit ball. Under (3), for any $T > 0$, $\beta \in (0, \gamma_0)$ and $\gamma \in (0,1)$, there is a constant $C > 0$ such that for all N,

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t^{N,1} - X_t|^{2\gamma}\right) \leqslant C\mathrm{e}^{C\varepsilon_N^{-2d}} \left(\mathbb{E}|X_0^{N,1} - X_0|^2 + \frac{\kappa_2}{N}\right)^\gamma + C\varepsilon_N^{2\beta\gamma}.$$

Remarks

- Suppose that for some $C > 0$,

$$\mathbb{E}|X_0^{N,1} - X_0|^2 \leqslant C/N.$$

  If one chooses $\varepsilon_N = (\ln N)^{-1/(2d)}$, then for some $C > 0$,

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t^{N,1} - X_t|^{2\gamma}\right) \leqslant \frac{C}{(\ln N)^{(\beta\gamma)/d}}.$$

- Although we assume F is bounded, once we can establish the existence of bounded solutions to Fokker-Planck equation under linear growth assumption of F on r, then the boundedness of F is no longer a restriction. We illustrate this by the following example.

Example

- Consider the following special case:

$$\partial_t \rho = \Delta\rho + \mathrm{div}(F(\rho)\rho),$$

  where $F : \mathbb{R}_+ \to \mathbb{R}^d$ satisfies $\sum_{i=1}^d |F_i'(r)| \leqslant \kappa_1$.
- Notice that this equation can be written as the following transport form:

$$\partial_t \rho = \Delta\rho + (F(\rho) + F'(\rho)\rho) \cdot \nabla\rho,$$

  $\Rightarrow \|\rho_t\|_\infty \leqslant \|\rho_0\|_\infty := n_0$ by the maximum principle.
- Then $F(\rho) = [(-n_0) \vee F \wedge n_0](\rho)$ with bounded $(-n_0) \vee F \wedge n_0$.
- (Numerical experiment of Burgers equation)
  Consider $d = 1$ and $F(r) = r$ and take $\phi(x) = 1_{[-1,1]}(x)/2$:

$$\frac{1}{2N\varepsilon_N} \sum_{j=1}^N 1_{|X_t^{N,i} - X_t^{N,j}| \leqslant \varepsilon_N},$$

  is easily calculated in a computer.

Thank you!

Danke!