Propagation of chaos for a class of McKean-Vlasov models.

Jean-François Jabir

School of Mathematics, Edinburgh ; HSE, Moscow.

LSA Winter Meeting 2019

• Asymptotic and quantitative propagation of chaos results for general McKean-Vlasov dynamics : From pathwise (strong) sense : Sznitman 1989, Méléard 1995 to propagation of chaos in entropy : Ben Arous and Zeitouni 1999 ; Jabin and Wang 2016.

• Propagation of chaos in entropy and quantitative results via Girsanov transformation for McKean-Vlasov dynamics with path-dependent coefficients :

$$\begin{cases} dX_t = C(t,(X_r)_{0\leq r\leq t})\,dt \\ \qquad\qquad + A(t,(X_r)_{0\leq r\leq t})\Big(B(t,(X_r)_{0\leq r\leq t};\mathcal{L}((X_r)_{0\leq r\leq t}))\,dt + dW_t\Big),\ 0\leq t\leq T, \\ \mathcal{L}((X_r)_{0\leq r\leq t}) = \text{Law of } ((X_r)_{0\leq r\leq t}),\ X_0\sim\mu^0, \end{cases}$$

where W is a $\mathbb{R}^m$ standard Brownian motion, A and C are progressively measurable functionals on $[0,T]\times\mathcal{C}([0,T];\mathbb{R}^d)$ with values on $\mathbb{R}^{d\times m}$ and $\mathbb{R}^d$ respectively, B is of the form : For $0\leq t\leq T$, $x\in\mathcal{C}([0,T];\mathbb{R}^d)$, $P\in\mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d))$,

$$B(t,x,P) = B(t,(\omega_r(x);r\leq t),P\circ(\omega_r;\ r\leq t)^{-1}),\ \omega \text{ canonical process on } \mathcal{C}([0,T];\mathbb{R}^d).$$

Motivation :

• Investigate the explicit rate of convergence of a class of McKean-Vlasov dynamics containing :

∘ Toy model :

$$dX_t = \Big(\int b(t, X_t, x)\, \mu(t, dx)\Big)\, dt + \sigma(t, X_t)\, dW_t,\ \mu(t) = \mathcal{L}(X_t),\ X_0 \sim \mu_0,$$

∘ Langevin dynamic with nonlinearity of McKean type :

$$dY_t = V_t\, dt,\ (Y_0, V_0) \sim \mu_0$$
$$dV_t = \Big(\int b(t, (Y_t, V_t), (y, v))\, \mu(t, dy, dv)\Big)\, dt + \sigma(t, Y_t)\, dW_t,\ \mu(t) = \mathcal{L}(Y_t, V_t).$$

∘ Parabolic-parabolic Keller-Segel probabilistic model (Tomasevic and Talay 2019) :

$$dX_t = \chi\Big(\int_0^t \int \nabla_x g(t-s, X_t - y)\, \mu(s, dx)\, ds\Big)\, dt + dW_t,\ \mu(t) = \mathcal{L}(X_t),$$

where $\chi > 0$ is a scaling parameter.

• Deepening the recent study of Lacker 2018 (strongly inspired by Mishura and Veretennikov 2016) on the propagation of chaos property related to the McKean-Vlasov dynamic :

$$dX_t = B\big(t, (X_r)_{0\le r\le t}, \mathcal{L}((X_r)_{0\le r\le t})\big)\, dt + A\big(t, (X_r)_{0\le r\le t}\big)\, dW_t,$$

Short reminder on propagation of chaos property related to some McKean-Vlasov models :

• **Generic McKean-Vlasov model :**

$$X_t = \xi + \int_0^t b(s, X_s, \mu(s))\,ds + \int_0^t \sigma(s, X_s, \mu(s))\,dW_s,\ \xi \sim \mu_0,$$
$$\mu(t) = \mathcal{L}(X_t)(= \text{Law of}(X_t)),$$

where μ_0 is a given probability measure on $\mathbb{R}^d$, $(W_t;\ t \geq 0)$ is a m-dimensional Brownian motion, and

$$b : [0,\infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d,\ \sigma : [0,\infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{d\times m}$$

are given functions.

• Dynamics introduced in McKean 1966 for the probabilistic interpretation of nonlinear parabolic equations :

$$\begin{cases} \partial_t \mu(t,x) + \nabla_x \cdot (\mu(t,x) b(t,x,\mu(t))) - \frac{1}{2}\text{Trace}(\nabla_x^2 \otimes \mu(t,x)(\sigma\sigma)^*(t,x,\mu(t))) = 0, \\ \mu(t,x) = \mu_0(dx)/dx. \end{cases}$$

Particle approximation :

$$\begin{cases} X_t^{i,N} = \xi^i + \int_0^t b(s, X_s^{i,N}, \overline{\mu}(X_s^{[\![1\,,\,N]\!],N})\, ds + \int_0^t \sigma(s, X_s^{i,N}, \overline{\mu}(X_s^{[\![1\,,\,N]\!],N}))\, dW_s^i, \\ \mu^N(X_t^{[\![1\,,\,N]\!],N})(dx) = \dfrac{1}{N}\displaystyle\sum_{j=1}^N \delta_{\{X_t^{j,N}\in dx\}}, \end{cases}$$

where $(\xi^i,\ (W_t^i;\, ,\, t\geq 0))$, $i=1,\dots,N$, are independent copies of $(\xi,\ (W_t;\, ,\, t\geq 0))$.

Particle approximation :

$$\begin{cases} X_t^{i,N} = \xi^i + \int_0^t b(s, X_s^{i,N}, \overline{\mu}(X_s^{[\![1\,,\,N]\!],N})\,ds + \int_0^t \sigma(s, X_s^{i,N}, \overline{\mu}(X_s^{[\![1\,,\,N]\!],N}))\,dW_s^i, \\ \mu^N(X_t^{[\![1\,,\,N]\!],N})(dx) = \frac{1}{N}\sum_{j=1}^N \delta_{\{X_t^{j,N}\in dx\}}, \end{cases}$$

where $(\xi^i,\,(W_t^i;\,,t\geq 0))$, $i = 1,\ldots,N$, are independent copies of $(\xi,\,(W_t;\,,t\geq 0))$.

Propagation of chaos property (Kac 1956, McKean 1967) :The particle system $(X^{1,N},\ldots,X^{N,N})$ is said to propagate chaos whenever a block tagged stochastic particle systems $X^{[\![1\,,\,k]\!],N} := (X^{1,N},\ldots,X^{k,N})$, satisfies : for all $t\geq 0$,

$$\mathcal{L}(X_t^{[\![1\,,\,k]\!],N}) \xrightarrow[N\to\infty]{\text{weakly}} \mathcal{L}(X_t)^{\otimes k} := \underbrace{\mathcal{L}(X_t)\otimes\cdots\otimes\mathcal{L}(X_t)}_{\text{k times}}$$

Particle approximation :

$$\begin{cases} X_t^{i,N} = \xi^i + \int_0^t b(s, X_s^{i,N}, \overline{\mu}(X_s^{[\![1,N]\!],N})\,ds + \int_0^t \sigma(s, X_s^{i,N}, \overline{\mu}(X_s^{[\![1,N]\!],N}))\,dW_s^i, \\ \mu^N(X_t^{[\![1,N]\!],N})(dx) = \frac{1}{N}\sum_{j=1}^{N} \delta_{\{X_t^{j,N} \in dx\}}, \end{cases}$$

where $(\xi^i, (W_t^i; , t \geq 0))$, $i = 1, \ldots, N$, are independent copies of $(\xi, (W_t; , t \geq 0))$.

Propagation of chaos property (Kac 1956, McKean 1967) :The particle system $(X^{1,N}, ..., X^{N,N})$ is said to propagate chaos whenever a block tagged stochastic particle systems $X^{[\![1,k]\!],N} := (X^{1,N}, \ldots, X^{k,N})$, satisfies : for all $t \geq 0$,

$$\mathcal{L}(X_t^{[\![1,k]\!],N}) \xrightarrow[N\to\infty]{\text{weakly}} \mathcal{L}(X_t)^{\otimes k} := \underbrace{\mathcal{L}(X_t) \otimes \cdots \otimes \mathcal{L}(X_t)}_{\text{k times}}$$

Whenever $(X^{1,N}, ..., X^{N,N})$ is exchangeable and $k \geq 2$, then the property is equivalent to : For all $t \geq 0$,

$$\frac{1}{N}\sum_{j=1}^{N} \delta_{\{X_t^{j,N} \in dx\}} \xrightarrow[N\to\infty]{\text{weakly}} \mathcal{L}(X_t).$$

Toy models : For $0 < T < \infty$ a finite time horizon and $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ smooth, consider the nonlinear McKean SDE

$$dX_t = \left(\int b(X_t, y)\,\mu(t, dy)\right)\,dt + \left(\int \sigma(X_t, y)\,\mu(t, dy)\right)\,dW_t,\; X_0 = \xi \sim \mu_0,$$

and the related particle system :

$$\begin{cases} dX_t^{i,N} = \left(\frac{1}{N}\sum_{j=1}^{N} b(X_t^{i,N}, X_t^{j,N})\right)\,dt + \left(\frac{1}{N}\sum_{j=1}^{N}\sigma(X_t^{i,N}, X_t^{j,N})\right)\,dW_t^i, \\ (X_0^i, (W_t^i;\, t \geq 0))\text{independent copies of } (\xi, (W_t;\, t \geq 0)). \end{cases}$$

Theorem 1 (Quantitative pathwise propagation of chaos, Sznitman 1989 (case σ constant), Méléard 1995)

Assume that $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are Lipschitz continuous. Then $(X^{1,N}, \dots, X^{N,N})$ propagate chaos toward X with

$$W_{2,T}(\mathcal{L}(X^{[\![1,k]\!],N}), \mathcal{L}(X)^{\otimes k}) \leq C\sqrt{k/N},$$

where $W_{2,T}$ is the Monge-Kantorovich–Rubinstein/Wasserstein distance on $\mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d))$:

$$W_{2,T}(P,Q) = \left(\inf_{Y\sim P, Z\sim Q}\mathbb{E}\left[\max_{0\leq t\leq T}|Y_t - Z_t|^2\right]\right)^{1/2},\; P, Q \in \mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d)).$$

Non-exhaustive bibliographic references :

- Probabilistic interpretation of nonlinear pdes in Physics (fluid dynamics, statistical mechanics) and related stochastic particle methods :
  - Propagation of chaos for stochastic particle system with weak (smooth) or moderate interaction : Funaki 1984 ; Oelschlager 1984, 1985 ; Léonard 1986 ; Méléard and Roelly-Coppoletta 1987 ; Dawson and Gartner 1989 ; Bossy and Talay 1996 ; Jourdain and Méléard 1998 ; Antonelli and Kohastu-Higa 2002 ;
  - Burgers equation : Sznitman 1986 ; Bossy and Talay 1997 ;
  - Incompressible Navier-Stokes equations : Marchioro and Pulvirenti 1982 ; Osada 1984 ; Méléard 2001 ; Fontbona 2004 ;
  - Conservative equations : Bossy and Jourdain 2000 ; Jourdain 2002 ; Jourdain and Reygner 2013 ;
  - Viscous Pressureless gas equation : Dermoune 2001 ;
  - Chemotaxis Keller-Segel model : Jourdain and Fournier 2017 ; Cattiaux and Pédéches 2017 ; J., Tomasevic and Talay 2018 ;
- Mean-Field Games, Optimal control problems for McKean-Vlasov dynamics, Machine learning : Lasry and Lions 2006 ; Lions "Cours au collège de France" (2006-2012) ; Carmona and Delarue 2013, 2015, 2018 ; Delarue, Lacker and Ramanan 2019 ; Mei, A. Montanari, and P.-M. Nguyen 2018 ; Hu *et al.* 2019 ; ...

Quantitative weak propagation of chaos : ... For a certain class of functions $f:\mathbb{R}^d\to\mathbb{R}$ and, under some regularity assumptions on b and σ,

$$\left|\mathbb{E}[f(X_t^{i,N})]-\mathbb{E}[f(X_t)]\right|\leq C/N.$$

Ref. : Kolokoltsov 2010 ; Kolokoltsov, Troeva and Wang 2014 ; Mischler and Mouhot 2013 ; Mischler, Mouhot and Wennberg 2014 ; Jourdain and Bencheikh 2019 ; Chassagneux, Szpruch and Tse 2019 ; Chaudru de Raynal and Frikha 2019 ; ...

Fundamental tools : Differential on $\mathcal{P}(\mathbb{R}^d)$:

• A functional $F:\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}$ is said to admit a linear functional/flat derivative on $\mathcal{P}(\mathbb{R}^d)$, if there exists $\frac{\delta F}{\delta\mu}:\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}$ such that

$$F(\nu)-F(\mu)=\int_0^1\int\frac{\delta F}{\delta\mu}(x,(1-\lambda)\mu+\lambda\nu)\big(\mu(dx)-\nu(dx)\big)\,d\lambda,\text{ for all }\nu,\mu\in\mathcal{P}(\mathbb{R}^d).$$

Propagation of chaos in entropy :

$$\lim_{N\to\infty} H\big(\mathcal{L}(X_t^{[\![1,k]\!],N})\,|\,\mathcal{L}(X_t)^{\otimes k}\big) = 0,$$

where H is the relative entropy of Q with respect to P :

$$H(Q\,|\,P) = \begin{cases} \displaystyle\int log(dQ/dP)\, dQ \text{ if } Q \ll P, \\ \infty \text{ otherwise.} \end{cases}$$

Ref. : Aebi 1996 ; Ben Arous and Zeitouni 1999 ; Jabin and Yang 2016 ; Lacker 2018.
• Give a a "stronger" form of propagation of chaos : by Kullback-Csiszár-Pinsker inequality :

$$d_{TV}\big(\mathcal{L}(X_t^{[\![1,k]\!],N}), \mathcal{L}(X_t)^{\otimes k}\big) \leq \sqrt{CH\big(\mathcal{L}(X_t^{[\![1,k]\!],N})\,|\,\mathcal{L}(X_t)^{\otimes k}\big)}$$

and, whenever $\mathcal{L}(X_t)$ satisfies a Talagrand inequality :

$$W_2\big(\mathcal{L}(X_t^{[\![1,k]\!],N}), \mathcal{L}(X_t)^{\otimes k}\big) \leq \sqrt{CH\big(\mathcal{L}(X_t^{[\![1,k]\!],N})\,|\,\mathcal{L}(X_t)^{\otimes k}\big)}$$

Quantitative results :

• Preliminary (Csiszár 1984 ; Hauray and Mischler 2015) : Let E be a metric space, $\mu \in \mathcal{P}(E)$, $\nu^N \in \mathcal{P}(E^N)$, $\nu^{k,N}(dx_1, ..., dx_k) = \int_{E^{N-k}} \nu(dx_1, , ..., dx_k, d\widetilde{x}^{k+1}, ..., d\widetilde{x}^N)$. Then,

$$\frac{1}{k} H(\nu^{k,N} \,|\, \mu^{\otimes k}) \leq \frac{1}{N} H(\nu^{N,N} \,|\, \mu^{\otimes N})$$

Quantitative results :

• Preliminary (Csiszár 1984 ; Hauray and Mischler 2015) : Let E be a metric space, $\mu \in \mathcal{P}(E)$, $\nu^N \in \mathcal{P}(E^N)$, $\nu^{k,N}(dx_1, ..., dx_k) = \int_{E^{N-k}} \nu(dx_1, , ..., dx_k, d\widetilde{x}^{k+1}, ..., d\widetilde{x}^N)$. Then,

$$\frac{1}{k} H(\nu^{k,N} \,|\, \mu^{\otimes k}) \leq \frac{1}{N} H(\nu^{N,N} \,|\, \mu^{\otimes N})$$

• Ben Arous and Zeitouni 1999 :

∘ **Interacting particle systems** :

$$dX_t^{i,N} = -\nabla_x U(X_t^{i,N})\, dt + \frac{1}{N} \sum_{j=1}^{N} \nabla_x \mathcal{U}(X_t^{i,N}, X_t^{j,N})\, dt + dW_t^i,\; X_0^{i,N} = \xi_i$$

∘ **Mean field limit system** :

$$dX_t = -\nabla_x U(X_t)\, dt + \Big(\int \nabla_x \mathcal{U}(X_t, y)\mu(t, dy)\Big)\, dt + dW_t,\; X_0 = \xi.$$

Theorem 2 (Ben Arous and Zeitouni 1999)

Assume that U and $\mathcal{U}$ are of class $\mathcal{C}_b^2$. Then, $H\big(\mathcal{L}(X_t^{1,N}, ..., X_t^{N,N}) \,|\, \mathcal{L}(X_t)^{\otimes N}\big) < \infty$ and, for all $k = k(N)$ such that $\lim_{N\to\infty} k(N)/N = 0$, it holds :

$$\lim_{N\to\infty} H\big(\mathcal{L}(X_t^{1,N}, ..., X_t^{k(N),N};\, t \leq T) \,|\, \mathcal{L}(X_t;\, t \leq T)^{\otimes k(N)}\big) = 0.$$

- Jabin and Wang 2016 : Propagation of chaos for Langevin dynamics :
  - **Interacting particle systems with bounded interacting kernel** b :

$$\begin{cases} dY_t^{i,N} = V_t^{i,N}\,dt, \\ dV_t^{i,N} = \dfrac{1}{N-1}\displaystyle\sum_{j=1,j\neq i}^{N} b(Y_t^{i,N} - Y_t^{j,N})\,dt + \sigma W_t^i. \end{cases}$$

  - **Mean field limit system** :

$$\begin{cases} dY_t = V_t\,dt, \\ dV_t = \left(\displaystyle\int b(Y_t - x)\mu(t,dy)\right)\,dt + \sigma dW_t. \end{cases}$$

Theorem 3 (Jabin and Wang 2016)

Assume that b is bounded, and $\{(X_0^{i,N}, V_0^{i,N})\}_{i=1}^N$ are i.i.d. with distribution $\mu_0(dy,dv) = \rho_0(y,v)dydv$ satisfying, for some $\lambda > 0$,

$$\mathbb{E}\left[|X_0|^{2\lambda} + |U_0|^{2\lambda} + \log(\rho_0(X_0,U_0))\right] < \infty.$$

Then $H\big(\mathcal{L}(X_t^{1,N},...,X_t^{N,N})\,|\,\mathcal{L}(X_t)^{\otimes N}\big) < \infty$ and $H\big(\mathcal{L}(X_t^{1,N},...,X_t^{k,N})\,|\,\mathcal{L}(X_t)^{\otimes k}\big) \leq Ck/N$.

Main approach : PDE analysis. Extensions to McKean-Vlasov dynamics with singular kernel : Jabin and Wang 2018 ; Bresch, Jabin and Wang 2019.

Intuition : Controlling

$$\sup_N H\big(\mathcal{L}(X_t^{[\![1,N]\!],N};\, t\leq T)\,|\,\mathcal{L}(X_t;\, t\leq T)^{\otimes N}\big)$$

(or simply $\sup_N H\big(\mathcal{L}(X_t^{1,N},...,X_t^{N,N})\,|\,\mathcal{L}(X_t)^{\otimes N}\big)<\infty$) can be expressed in terms of the exponential martingale Z_T^N (provided it exists) mapping $\mathcal{L}(X_t;\, t\leq T)^{\otimes N}$ toward $\mathcal{L}(X_t^{[\![1,N]\!],\infty};\, t\leq T)$ as

$$H\big(\mathcal{L}(X_t^{[\![1,N]\!],N};\, t\leq T)\,|\,\mathcal{L}(X_t;\, t\leq T)^{\otimes N}\big)=\mathbb{E}[Z_T^N\log(Z_T^N)].$$

In particular, if $\sup_N \mathbb{E}[|Z_T^N|^{1+\epsilon}]<\infty$ for some $\epsilon>0$, then

$$H\big(\mathcal{L}(X_t^{[\![1,k]\!],N};\, t\leq T)\,|\,\mathcal{L}(X_t;\, t\leq T)^{\otimes N}\big)<\infty$$

and the rate k/N arises.
This program can be done but not directly.

A simple case : For the toy model with $\sigma = I_d$, $m = d$, $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is Borel bounded,

$$dX_t = \left(\int b(X_t, y)\, \mu(t, dy) \right) dt + dW_t,\ X_0 = \xi \sim \mu_0,$$

$$dX_t^{i,N} = \left(\frac{1}{N} \sum_{j=1}^{N} b(X_t^{i,N}, X_t^{j,N}) \right) dt + dW_t^i,\ X_0^{i,N} = \xi_i \sim \mu_0,$$

then

$$Z_T^N = \exp \Bigg\{ \sum_{i=1}^{N} \int_0^T \left(\frac{1}{N} \sum_{j=1}^{N} b(X_t^{i,\infty}, X_t^{j,\infty}) - \int b(X_t^{i,\infty}, y)\, \mu(t, dy) \right) dW_t^i$$
$$- \frac{1}{2} \sum_{i=1}^{N} \int_0^T \left| \frac{1}{N} \sum_{j=1}^{N} b(X_t^{i,\infty}, X_t^{j,\infty}) - \int b(X_t^{i,\infty}, y)\, \mu(t, dy) \right|^2 dt \Bigg\},$$

where $X^{1,\infty}$, $X^{2,\infty}$, ..., $X^{N,\infty}$ are the independent copies of X given by

$$dX_t^{i,\infty} = \left(\int b(X_t^{\infty}, y)\, \mu(t, dy) \right) dt + dW_t^i,\ X_0^{i,\infty} = \xi_i \sim \mu_0.$$

Difficulty : Obtaining a control of $\mathbb{E}[(Z_T^N)^{1+\epsilon}]$ for any time T.

Lemma 4 (Local in time control)

For any time $\kappa > 1$,

$$\sup_N \mathbb{E}\left[(Z_T^N)^\kappa\right] < \infty,$$

provided that $T < 1/(C\|b\|_{L^\infty}\kappa)$ *for* C *some numerical constant. And, more generally,*

$$\sup_N \mathbb{E}\left[(Z_{T+\delta}^N/Z_T^N)^\kappa\right] < \infty,$$

provided that $\delta < 1/(C\|b\|_{L^\infty}^2\kappa)$.

Lemma 4 (Local in time control)

For any time $\kappa > 1$,

$$\sup_N \mathbb{E}\left[(Z_T^N)^\kappa\right] < \infty,$$

provided that $T < 1/(C\|b\|_{L^\infty}\kappa)$ for C some numerical constant. And, more generally,

$$\sup_N \mathbb{E}\left[(Z_{T+\delta}^N/Z_T^N)^\kappa\right] < \infty,$$

provided that $\delta < 1/(C\|b\|^2_{L^\infty}\kappa)$.

Element of proof : For $\triangle b_t^{i,N} = \frac{1}{N}\sum_{j=1}^N b(X_t^{i,\infty}, X_t^{j,\infty}) - \int b(X_t^{i,\infty}, y)\,\mu(t,dy)$,

$$\mathbb{E}[(Z_T^N)^\kappa] \le \mathbb{E}\left[\exp\left\{\kappa\sum_{i=1}^N\int_0^T \triangle b_t^{i,N}\,dW_t^i\right\}\right] \le \sum_{k\ge 0}\frac{\kappa^k}{k!}\mathbb{E}\left[\left(\sum_{i=1}^N\int_0^T \triangle b_t^{i,N}\,dW_t^i\right)^k\right],$$

$$\le \sum_{k\ge 0}\frac{\kappa^{2k}}{(2k)!}(1+2(k+1))\mathbb{E}\left[\left(\sum_{i=1}^N\int_0^T \triangle b_t^{i,N}\,dW_t^i\right)^{2k}\right],$$

and the (sub-Gaussian) moment property :

$$\mathbb{E}\left[\left(\int_0^T \left|\triangle b_t^{i,N}\right|^2 dt\right)^k\right] \le \frac{k!(C\|b\|^2_{L^\infty})^k T^k}{N^k}.$$

To overcome the lack of global in time control on $\mathbb{E}[(Z^N_T)^\kappa]$, introduce local in time Girsanov transformation and use again a metric instead of the relative entropy.

To overcome the lack of global in time control on $\mathbb{E}[(Z^N_T)^\kappa]$, introduce local in time Girsanov transformation and use again a metric instead of the relative entropy.
Alternative : Given $\kappa > 1$, split any finite interval $[0, T]$ into M sub-interval of size $\delta < 1/(C\|b\|_{L^\infty}\kappa)$ and introduce the family of processes $(\{Y^{i,N,m}\}_{i=1}^N)$, $m = 0, .., M$ given by

$$Y^{i,N,m}_t = \begin{cases} X^i_0 + \int_0^t \int b(X^{i,\infty}_s - y)\mu(s, dy)\, ds + W^i_t,\ 0 \le t \le m\delta, \\ Y^{i,N,m}_{m\delta} + \frac{1}{N}\sum_{j=1}^N \int_{m\delta}^t b(X^{i,\infty}_s - X^{j,\infty}_s)\, ds + \sigma(W^i_t - W^i_{\delta m}),\ m\delta < t \le T \wedge \delta M. \end{cases}$$

- $Y^{1,N,m}$, ..., $Y^{1,N,m}$ are independent McKean-Vlasov dynamics up to time $m\delta$;
- Between $m\delta$ and T, $Y^{1,N,m}$, ..., $Y^{1,N,m}$ are interacting particle systems submitted to the effect of the i.i.d. initial states $Y^{1,N,m}_{m\delta}$, ..., $Y^{1,N,m}_{m\delta}$.
- The exponential martingale mapping $Y^{1,N,m+1}$, ..., $Y^{1,N,m+1}$ to $Y^{1,N,m}$, ..., $Y^{1,N,m}$ is equivalent in law to $Z^N_{(m+1)\delta}/Z^N_{m\delta}$.

Using the triangular inequality of the total variation distance :

$$d_{TV}\big(\mathcal{L}(X_t^{[\![1,k]\!],N};\, t\leq T), \mathcal{L}(X_t;\, t\leq T)^{\otimes k}\big)$$
$$\leq \sum_{m=0}^{M-1} d_{TV}\big(\mathcal{L}(Y_t^{[\![1,k]\!],N,m+1};\, t\leq T), \mathcal{L}(Y_t^{[\![1,k]\!],N,m};\, t\leq T)\big).$$

Using the triangular inequality of the total variation distance :

$$d_{TV}\big(\mathcal{L}(X_t^{[\![1,k]\!],N};\, t \leq T), \mathcal{L}(X_t;\, t \leq T)^{\otimes k}\big) \\ \leq \sum_{m=0}^{M-1} d_{TV}\big(\mathcal{L}(Y_t^{[\![1,k]\!],N,m+1};\, t \leq T), \mathcal{L}(Y_t^{[\![1,k]\!],N,m};\, t \leq T)\big).$$

For each m,

$$\begin{aligned}
& d_{TV}\big(\mathcal{L}(Y_t^{[\![1,k]\!],N,m};\, t \leq T), \mathcal{L}(Y_t^{[\![1,k]\!],N,m+1};\, t \leq T)^{\otimes k}\big) \\
&= \sup_{F:\mathcal{C}([0,T];\mathbb{R}^{kd})\to\mathbb{R}\,:\,\|F\|_{L^\infty}\leq 1} \left|\mathbb{E}\left[F(Y_t^{[\![1,k]\!],N,m+1};\, t \leq T)\right] - \mathbb{E}\left[F(Y_t^{[\![1,k]\!],N,m};\, t \leq T)\right]\right| \\
&= \sup_{F:\mathcal{C}([0,T];\mathbb{R}^{kd})\to\mathbb{R}\,:\,\|F\|_{L^\infty}\leq 1} \left|\mathbb{E}\left[\left(Z^N_{((m+1)\delta)\wedge T}/Z^N_{(m\delta)\wedge T} - 1\right) F(Y_t^{[\![1,N]\!],N,m+1};\, t \leq T)\right]\right| \\
&\leq \mathbb{E}\left[\left|\int_{(m\delta)\wedge T}^{((m+1)\delta)\wedge T} \left(\triangle b_t^{1,N}\right)\left(Z^N_t/Z^N_{(m\delta)\wedge T}\right) dW^1_t\right|\right] \\
&\leq \sqrt{k}\frac{p}{p-1}\left(\mathbb{E}\left[(Z^N_{((m+1)\delta)\wedge T}/Z^N_{(m\delta)\wedge T})^p\right]\right)^{1/p} \\
&\quad \times \left(\mathbb{E}\left[\left(\int_{(m\delta)\wedge T}^{((m+1)\delta)\wedge T} \left|\triangle b_t^{1,N}\right|^2 dt\right)^{p/(2(p-1))}\right]\right)^{(p-1)/p} \\
&\leq C_p\sqrt{k/N},
\end{aligned}$$

for any $1 < p < \infty$.

Using the triangular inequality of the total variation distance :

$$d_{TV}\big(\mathcal{L}(X_t^{[\![1,k]\!],N};\, t\leq T), \mathcal{L}(X_t;\, t\leq T)^{\otimes k}\big)$$
$$\leq \sum_{m=0}^{M-1} d_{TV}\big(\mathcal{L}(Y_t^{[\![1,k]\!],N,m+1};\, t\leq T), \mathcal{L}(Y_t^{[\![1,k]\!],N,m};\, t\leq T)\big).$$

For each m,

$$d_{TV}\big(\mathcal{L}(Y_t^{[\![1,k]\!],N,m};\, t\leq T), \mathcal{L}(Y_t^{[\![1,k]\!],N,m+1};\, t\leq T)^{\otimes k}\big)$$
$$= \sup_{F:\mathcal{C}([0,T];\mathbb{R}^{kd})\to\mathbb{R}\,:\,\|F\|_{L^\infty}\leq 1} \left|\mathbb{E}\left[F(Y_t^{[\![1,k]\!],N,m+1};\, t\leq T)\right] - \mathbb{E}\left[F(Y_t^{[\![1,k]\!],N,m};\, t\leq T)\right]\right|$$
$$= \sup_{F:\mathcal{C}([0,T];\mathbb{R}^{kd})\to\mathbb{R}\,:\,\|F\|_{L^\infty}\leq 1} \left|\mathbb{E}\left[\left(Z^N_{((m+1)\delta)\wedge T}/Z^N_{(m\delta)\wedge T} - 1\right) F(Y_t^{[\![1,N]\!],N,m+1};\, t\leq T)\right]\right|$$
$$\leq \mathbb{E}\left[\left|\int_{(m\delta)\wedge T}^{((m+1)\delta)\wedge T} \left(\triangle b_t^{1,N}\right)\left(Z_t^N/Z^N_{(m\delta)\wedge T}\right) dW_t^1\right|\right]$$
$$\leq \sqrt{k}\frac{p}{p-1}\left(\mathbb{E}\left[(Z^N_{((m+1)\delta)\wedge T}/Z^N_{(m\delta)\wedge T})^p\right]\right)^{1/p}$$
$$\times\left(\mathbb{E}\left[\left(\int_{(m\delta)\wedge T}^{((m+1)\delta)\wedge T}\left|\triangle b_t^{1,N}\right|^2 dt\right)^{p/(2(p-1))}\right]\right)^{(p-1)/p}$$
$$\leq C_p\sqrt{k/N},$$

for any $1 < p < \infty$.

McKean-Vlasov dynamic with path dependent coefficients :

$$\begin{cases} dX_t = C(t,(X_r)_{0\leq r\leq t})\,dt \\ \qquad\qquad + A(t,(X_r)_{0\leq r\leq t})\Big(B(t,(X_r)_{0\leq r\leq t};\mathcal{L}((X_r)_{0\leq r\leq t}))\,dt + dW_t\Big),\ 0\leq t\leq T, \\ \mathcal{L}((X_r)_{0\leq r\leq t}) = \text{Law of } ((X_r)_{0\leq r\leq t}),\ X_0\sim\mu^0, \end{cases} \tag{1}$$

where C and A are progressively measurable functionals on $[0,T]\times\mathcal{C}([0,T];\mathbb{R}^d)$ and B is progressively measurable on $[0,T]\times\mathcal{C}([0,T];\mathbb{R}^d)\times\mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d))$.

Related particle system :

$$\begin{cases} dX_t^{i,N} = C(t,(X_r^{i,N})_{0\leq r\leq t})\,dt \\ \qquad\qquad + A(t,(X_r^{i,N})_{0\leq r\leq t})\Big(B(t,(X_r^{i,N})_{0\leq r\leq t};\overline{\mu}_t(X^{[\![1,N]\!],N}))\,dt + dW_t^i\Big),\ 0\leq t\leq T, \\ \overline{\mu}_t(X^{[\![1,N]\!],N}) = \dfrac{1}{N}\displaystyle\sum_{j=1}^N \delta_{\{(X_r^{j,N})_{0\leq r\leq t}\}}, \\ (X_0^{1,N},X_0^{2,N},\dots,X_0^{N,N})\text{ i.i.d. with distribution } \mu_0. \end{cases} \tag{2}$$

Assumptions :

(H_1) : For all $0 \leq t \leq T$, $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, there exists a weak solution to the McKean-Vlasov dynamic (1) such that

$$\mathbb{P}\left(\int_0^T \left|A(t,(X_r)_{0\leq r\leq t})B(t,(X_r)_{0\leq r\leq t};\mathcal{L}((X_r)_{0\leq r\leq t}))\right|^2 dt < \infty\right) = 1.$$

(H_2) : For all $0 \leq t \leq T$, $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, there exists a weak solution to the particle system (2) such that, for all $1 \leq i \leq N$

$$\mathbb{P}\left(\int_0^T \left|A(t,(X_r^{i,N})_{0\leq r\leq t})B(t,(X_r^{i,N})_{0\leq r\leq t};\overline{\mu}_t(X^{[\![1,N]\!],N}))\right|^2 dt < \infty\right) = 1.$$

(H_3) : For all $0 \leq t \leq T$, there exists a unique weak solution to

$$\begin{aligned}
&dY_t = C(t,(Y_r)_{0\leq r\leq t})\,dt + A(t,(Y_r)_{0\leq r\leq t})\,dW_t,\ 0 \leq t \leq T,\\
&Y_0 \sim \mu^0,
\end{aligned}$$

such that

$$\mathbb{P}\left(\int_0^T \left|A(t,(Y_r)_{0\leq r\leq t})B(t,(Y_r)_{0\leq r\leq t};\mathcal{L}((X_r)_{0\leq r\leq t}))\right|^2 dt < \infty\right) = 1.$$

Assumptions :

(H_4) : There exists a constant $0 < \beta < \infty$ such that for any $0 < T_0 < T < \infty$, $0 < \delta < \infty$, and, for all integer $p \geq 1$,

$$\mathbb{E}\left[\left(\int_{T_0}^{(T_0+\delta)\wedge T} \left|\triangle B_t^{i,N,\infty}\right|^2 \, dt\right)^p\right] \leq \frac{p!\beta^p\delta^p}{N^p},$$

where

$$\triangle B_t^{i,N,\infty} := B(t,(X_r^{i,\infty})_{0\leq r\leq t}; \overline{\mu}_t(X^{[\![1,N]\!],\infty})) - B(t,(X_r^{i,\infty})_{0\leq r\leq t}; \mathcal{L}((X_r^{i,\infty})_{0\leq r\leq t})),$$

$$\overline{\mu}_t(X^{[\![1,N]\!],\infty}) = \frac{1}{N}\sum_{j=1}^{N} \delta_{\{(X_r^{j,\infty})_{0\leq r\leq t}\}}$$

for $(X^{1,\infty}, X^{2,\infty}, ..., X^{N,\infty})$ a family of N-independent copies of (1) :

$$\begin{cases} dX_t^{i,\infty} = C(t,(X_r^{i,\infty})_{0\leq r\leq t})\, dt \\ \qquad\qquad + A(t,(X_r^{i,\infty})_{0\leq r\leq t})\Big(B(t,(X_r^{i,\infty})_{0\leq r\leq t}; \mathcal{L}((X_r^{i,\infty})_{0\leq r\leq t}))\, ds + dW_t^i\Big), \\ \mu^{i,\infty}(t) = \mathcal{L}((X_r^{i,\infty})_{0\leq r\leq t}),\ X_0^{i,\infty} \sim \mu^0, \text{ i.i.d..} \end{cases}$$

Theorem 5 (J. 2019)

Assume that the hypotheses (H_1) to (H_4) hold. Then

$$d_{TV}\big(\mathcal{L}(X_t^{[\![1,k]\!],N},\, t \leq T), \mathcal{L}(X_t^{[\![1,k]\!],\infty},\, t \leq T)\big) \leq C(1+\beta T)\sqrt{\frac{k}{N}},$$

where C some constant independent of k and N.

Theorem 6 (J. 2019)

Assume that the hypotheses (H_1) to (H_4) hold, and that for some $\beta'' > 0$, $\mathbb{E}[max_{0\leq t\leq T}|X_t^{1,\infty}|^{\beta''}]$ and $\sup_N \mathbb{E}[max_{0\leq t\leq T}|X_t^{1,N}|^{\beta''}]$ are finite. Then

$$\mathcal{T}_{c_{T,k}}\Big(\mathcal{L}((X_t^{1,N},\cdots,X_t^{k,N})_{0\leq t\leq T}), \mathcal{L}\big((X_t^{1,\infty},\cdots,X_t^{k,\infty})_{0\leq t\leq T}\big)\Big) \leq \frac{C\sqrt{k}(1+\beta T)}{\sqrt{N}}.$$

for $\mathcal{T}_{c_{T,k}}$ the transportation cost functional :

$$\mathcal{T}_{c_{T,k}}(P,Q) = \inf_{Y\sim P, Z\sim Q} \mathbb{E}[c_{T,k}(Y,Z)],\ P,Q \in \mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^{kd})),$$

where $c_{T,k}$ is a cost function satisfying : $c_{T,k}(x,x) = 0$, for all $x \in \mathcal{C}([0,T];\mathbb{R}^{kd})$ and there exists $\widehat{x} \in \mathcal{C}([0,T];\mathbb{R}^{kd})$ and $0 < \beta$, $0 < \beta' < \beta''$ such that

$$c_{T,k}(x,\widehat{x}) \leq \beta(1 + \max_{0\leq t\leq T}|x(t)|^{\beta'}),\ \forall x \in \mathcal{C}([0,T];\mathbb{R}^{kd}).$$

Elements of proof :

• Assumptions (H_1), (H_2) and (H_3) ensures the existence of the Radon-Nikodym density of $\mathcal{L}(X_t^{[\![1,N]\!],N};\, t \leq T)$ w.r.t. $\mathcal{L}(X_t^{[\![1,N]\!],\infty};\, t \leq T)$ for all $0 < T < \infty$:

$$Z_T^N = \exp\left\{ \sum_{i=1}^{N} \int_0^t \triangle B_s^{i,N}\, dW_s^i - \frac{1}{2} \sum_{i=1}^{N} \int_0^T \left| \triangle B_s^{i,N} \right|^2 ds \right\}.$$

• Assumption (H_4) ensures that the local control of moment of $(Z_t^N;\, t \leq T)$: for all $\kappa > 0$,

$$\sup_N \mathbb{E}\left[(Z_{T_0+\delta}^N / Z_{T_0}^N)^\kappa \right] \leq 1 + \exp \kappa^2 + \frac{2}{1 - 8\kappa\delta\beta},$$

provided that $\delta < (8\kappa\beta)^{-1}$.

Comment : The preceding result gives an optimal rate of convergence covering the case of the family of metrics : $0 \leq \alpha \leq 1$,

$$W_{\alpha,T}(P,Q) = \left(\inf_{Y \sim P, Z \sim Q} \mathbb{E}\left[\max_{0 \leq t \leq T} |Y_t - Z_t|^\alpha \right] \right),\; P, Q \in \mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d)),$$

But the rate is sub-optimal in the case $\mathcal{T}_{c_{T,k}} = W_{2,T}^2$.

Lemma 7

Assume that (H_1), (H_2) and (H_3) hold. Assume also that, for all $0 \leq t < \infty$, $x \in \mathcal{C}([0,\infty);\mathbb{R}^d)$ $\nu \in \mathcal{P}(\mathcal{C}([0,\infty);\mathbb{R}^d)) \mapsto B(t,x;P)$ is Lipschitz continuous w.r.t. the total variation distance ; that is there exists $0 < K < \infty$ such that $P, Q \in \mathcal{P}(\mathcal{C}([0,\infty);\mathbb{R}^d))$, $0 \leq t < \infty$, $x \in \mathcal{C}([0,\infty);\mathbb{R}^d)$,

$$|B(t,x;P) - B(t,x;Q)| \leq K d_{TV}(P,Q). \tag{3}$$

Assume finally that the following centering (conditional) property holds :

$$\mathbb{E}\left[B\big(t, X^{i,\infty}; \frac{1}{N-1}\sum_{j=1, j\neq i}^{N} \delta_{\{X^{i,\infty}\}}\big) \,\Big|\, X^{i,\infty} \right] = B(t, X^{i,\infty}; \mathcal{L}(X^{i,\infty}))$$

Then the condition (H_4) is satisfied.

The functional B is said to admit a bounded first and second order flat derivative if, for all $1 \leq l \leq m$ there exist two progressively measurable bounded functionals : $\frac{dB}{dm}$ and $\frac{d^2B}{dm^2}$ such that, for all $0 \leq t \leq T$, $x \in \mathcal{C}([0,T];\mathbb{R}^d)$, $P, Q \in \mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d))$,

$$B(t,x,Q)-B(t,x,P) = \int_0^1 \int_{\mathcal{C}([0,T];\mathbb{R}^d)} \frac{dB}{dm}(t,x,(1-\alpha)P+\alpha Q;\omega)\,(Q(d\omega)-P(d\omega))\;d\alpha,$$

and, for all $0 \leq t \leq T$, $x \in \mathcal{C}([0,T];\mathbb{R}^d)$, $P, Q \in \mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d))$, $\omega \in \mathcal{C}([0,T];\mathbb{R}^d)$,

$$\begin{aligned}&\frac{dB}{dm}(t,x,Q;\omega) - \frac{dB}{dm}(t,x,P;\omega)\\ &= \int_0^1 \int_{\mathcal{C}([0,T];\mathbb{R}^d)} \frac{d^2B}{dm^2}(t,x,(1-\alpha)P+\alpha Q;\omega,\tilde{\omega})\,(Q(d\tilde{\omega})-P(d\tilde{\omega}))\;d\alpha.\end{aligned}$$

where $(1-\alpha)P+\alpha Q$, $0 \leq \alpha \leq 1$ is the set of probability measures given by the convex interpolations between P and Q.

Proposition 8

Assume that (H_1), (H_2) and (H_3) hold and that for all $0 \leq t \leq T$, $x \in \mathcal{C}([0,T];\mathbb{R}^d)$, $\mu \in \mathcal{P}(\mathcal{C}[0,T];\mathbb{R}^d) \mapsto B(t,x,\mu)$ admits a bounded first and second order derivative. Then the condition (H_4) holds.

• Rate of convergence for reflected McKean-Vlasov dynamics on a domain $\mathcal{D} \subset \mathbb{R}^d$ (Sznitman 1984) and Langevin dynamics with specular boundary condition :

$$dY_t = V_t\, dt,\ (Y_0, V_0) \sim \mu_0$$
$$dV_t = \Big(\int b(t, (Y_t, V_t), (y, v))\, \mu(t, dy, dv) \Big)\, dt + \sigma\, dW_t - 2 \sum_{s \le t} (V_{s^-} \cdot n_{\mathcal{D}}(Y_s)) n_{\mathcal{D}}(Y_s).$$

• On the parabolic-parabolic Keller-Segel :

$$dX_t = \chi \Big(\int_0^t \int \nabla_x g(t-s, X_t - y)\, \mu(s, dx)\, ds \Big)\, dt + dW_t,\ \mu(t) = \mathcal{L}(X_t),$$

$$dX_t^{i,N} = \chi \Big(\frac{1}{N} \sum_{j=1, j \neq i}^{N} \int_0^t \int \nabla_x g(t-s, X_t^{i,N} - X_s^{j,N})\, ds \Big)\, dt + dW_t^i.$$

No apriori estimate on the interaction kernel enable to deduce a quantitative propagation of chaos. Yet, it is known that for $d = 1$, and $\mu_0 \in L^\infty \cap L^1$, X admits a bounded smooth density.

• Weak propagation of chaos for the "Markovian" case

$$X_t = X_0 + \int_0^t c(s, X_s)\, ds + \int_0^t a(s, X_s) \Big(b(s, X_s; \mathcal{L}(X_s))\, ds + dW_s \Big),\ T_0 \le t \le T;$$