NOTES ON MEAN FIELD GAMES //1

Guangyu Hou

Contents

1 Background **3**

1.1 Weak convergence of empirical measures 3

1.2 Wasserstein metrics 3

1.3 Kantorovich duality 4

1.4 Interaction functions 5

2 Interacting Particle Systems **5**

2.1 McKean-Vlasov limit 5

2.1.1 Nice Coefficients 5

2.2 McKean-Vlasov equation 7

2.3 From SDE to PDE 8

2.4 Common Noise 10

2.5 Long-time behavior 11

1 Background

1.1 Weak convergence of empirical measures

empirical measures
Suppose (X_i) are i.i.d. $\mathcal{X}$-valued random variables. Define the $\mathcal{P}(\mathcal{X})$ -valued random variable

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}.$$

thus the integral of a test function takes the following form:

$$\int_{\mathcal{X}} f d\mu_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i).$$

Theorem.
If $(\mathcal{X}, d)$ is separable, then it holds with probability 1 that $\mu_n \to \mu$ weakly.

1.2 Wasserstein metrics

The p**-Wasserstein metric** on $\mathcal{P}^p(\mathcal{X})$ is defined by

$$\mathcal{W}_{\mathcal{X},p}(\mu,\nu) = \left(\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \pi(dx,dy)\right)^{1/p},$$

where $\mathcal{P}^p(\mathcal{X})$ is the set of probability measures $\mu \in \mathcal{P}(\mathcal{X})$ satisfying

$$\int_{\mathcal{X},\forall x_0\in\mathcal{X}} d(x,x_0)^p \mu(dx) < \infty$$

If the space $\mathcal{X}$ is understood, we write simply $\mathcal{W}_p$ instead of $\mathcal{W}_{\mathcal{X},p}$. An equivalent and more probabilistic definition reads

$$\mathcal{W}_p(\mu,\nu) = \left(\inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\left[d(X,Y)^p\right]\right)^{1/p},$$

where the infimum is over all pairs of $\mathcal{X}$-valued random variables X and Y with given marginals μ and ν. The Wasserstein metric is very convenient in that it involves an infimum, which makes it quite easy to bound.

It's a metric on $\mathcal{P}(\mathcal{X})$ which is compatible with weak convergence, based on the idea of a coupling. For $\mu, \nu \in \mathcal{P}(\mathcal{X})$, we write $\Pi(\mu,\nu)$ to denote the set of Borel probability measures π on $\mathcal{X} \times \mathcal{X}$ with two marginals. The theorem below can clearly describe this convergence.

Theorem.
Let $\mu, \mu_n \in \mathcal{P}^p(\mathcal{X})$ for some $p \geq 1$. The following are equivalent:
(i) $\mathcal{W}_p(\mu_n, \mu) \to 0$.
(ii) For every continuous function $f : \mathcal{X} \to \mathbb{R}$ with the property that there exist $x_0 \in \mathcal{X}$ and $c > 0$ such that $|f(x)| \leq c(1 + d(x, x_0)^p)$ for all $x \in \mathcal{X}$, we have

$$\int_{\mathcal{X}} f d\mu_n \to \int_{\mathcal{X}} f d\mu.$$

(iii) $\mu_n \to \mu$ weakly and $\int_{\mathcal{X}} d(x, x_0)^p \mu_n(dx) \to \int_{\mathcal{X}} d(x, x_0)^p \mu(dx)$ for some $x_0 \in \mathcal{X}$.
(iv) $\mu_n \to \mu$ weakly and

$$\lim_{r \to \infty} \sup_n \int_{\{d(,x_0) \geq r\}} d(x, x_0)^p \mu_n(dx) = 0.$$

While Wasserstein metrics do not precisely metrize weak convergence, it is occasionally useful to note that they can be forced to by replacing the metric on $\mathcal{X}$ with an equivalent bounded metric. That is, if $(\mathcal{X}, d)$ is a separable metric space, then we can define a new metric

$$\bar{d}(x, y) := 1 \wedge d(x, y) = \min\{1, d(x, y)\}, \quad \text{for } x, y \in \mathcal{X}.$$

which can ensure the equivalence.

Corollary.
Suppose $\mathcal{X}$ is separable. Suppose (X_i) are i.i.d. $\mathcal{X}$-valued r.v.s with law μ, for empirical measure μ_n, let $p \geq 1$. If $\mu \in \mathcal{P}^p(\mathcal{X})$, then $\mathcal{W}_p(\mu_n, \mu) \to 0$ almost surely, and also

$$\mathbb{E}[\mathcal{W}_p^p(\mu_n, \mu)] \to 0.$$

1.3 Kantorovich duality

Theorem.
Suppose $(\mathcal{X}, d)$ is a complete and separable metric space, and define $\mathcal{W}_p$ for $p \geq 1$ as in Definition 2.10. Then, for any $\mu, \nu \in \mathcal{P}^p(\mathcal{X})$,

$$\mathcal{W}_p^p(\mu, \nu) = \sup\left\{\int f d\mu + \int g d\nu : f, g \in C_b(E), f(x) + g(y) \leq d(x, y)^p \forall x, y \in \mathcal{X}\right\}.$$

Moreover, for $p = 1$ and $\mu, \nu \in \mathcal{P}^1(\mathcal{X})$, we have

$$\mathcal{W}_1(\mu, \nu) = \sup_f \left(\int f d\mu - \int f d\nu\right),$$

where the supremum is over all functions $f : \mathcal{X} \to \mathbb{R}$ which are 1-Lipschitz in the sense that $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in \mathcal{X}$.

It takes its simplest form for the 1-Wasserstein metric, worth singling out. The proof ultimately boils down to the Fenchel-Rockafellar theorem.

1.4 Interaction functions

When we turn to our study of mean field games and interacting particle systems, our models will involve functions defined on $\mathcal{X} \times \mathcal{P}^p(\mathcal{X})$. We will think of such a function $F = F(x, \mu)$ as determining an interaction of a particle x with a distribution of particles μ.

2 Interacting Particle Systems

We will discuss a problem set up on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the usual conditions. This space should support (at the very least) an i.i.d. sequence $\left(\xi^i\right)$ of $\mathbb{R}^d$-valued $\mathcal{F}_0$-measurable random variables as well as a sequence $\left(W^i\right)$ of independent $\mathbb{F}$-Brownian motions. The main object of study will be a system of n interacting particles $\left(X_t^{n,1}, \ldots, X_t^{n,n}\right)$, driven by stochastic differential equations (SDEs) of the form

$$
\begin{aligned}
dX_t^{n,i} &= b\left(X_t^{n,i}, \mu_t^n\right) dt + \sigma\left(X_t^{n,i}, \mu_t^n\right) dW_t^i, \quad X_0^{n,i} = \xi^i \\
\mu_t^n &= \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^{n,k}}
\end{aligned}
$$

Driving this SDE system are n independent Brownian motions, $W^1, \ldots, W^n$, and we typically assume the initial states $\xi^1, \ldots, \xi^n$ are i.i.d. We think of $X_t^{n,i}$ as the position of particle i at time t, in a Euclidean space $\mathbb{R}^d$.

- We think of the number n of particles as very large, and ultimately we will send it to infinity.
- Key structural feature that makes this system amenable to mean field analysis: The coefficients b and σ are the **same** for each particle, and the only dependence of particle i on the rest of the particles $k \neq i$ is through the empirical measure μ_t^n

2.1 McKean-Vlasov limit

In most cases, we cannot do explicit computation to solve the SDEs above, then we need to work on a general understanding of it and also find out the $n \to \infty$ behavior of the systems, but under certain assumptions.

2.1.1 Nice Coefficients

Assumption.
Assume the initial states $\left(\xi^i\right)_{i \in \mathbb{N}}$ are i.i.d. with $\mathbb{E}\left[\left|\xi^1\right|^2\right] < \infty$. The Brownian motions $\left(W^i\right)_{i \in \mathbb{N}}$ are independent and m-dimensional. Assume $b : \mathbb{R}^d \times \mathcal{P}^2\left(\mathbb{R}^d\right) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}^2\left(\mathbb{R}^d\right) \to \mathbb{R}^{d \times m}$ are **Lipschitz**, in the sense that there exists a constant $L > 0$ such that

$$
|b(x, m) - b\left(x', m'\right)| + |\sigma(x, m) - \sigma\left(x', m'\right)| \leq L\left(|x - x'| + \mathcal{W}_2\left(m, m'\right)\right).
$$

This assumption (**Lipschitz**) immediately lets us check that the SDE system is well-posed. We make heavy use of the following well-known well-posedness result: **the n-particle SDE system admits a unique strong solution, for each n.**

Proof. Define the $\mathbb{R}^{nd}$-valued process $\boldsymbol{X}_t = \left(X_t^{n,1}, \ldots, X_t^{n,n}\right)$, and similar define the nm-dimensional Brownian motion $\boldsymbol{W}_t = \left(W_t^1, \ldots, W_t^n\right)$. We may write

$$d\boldsymbol{X}_t = B\left(\boldsymbol{X}_t\right) dt + \Sigma\left(\boldsymbol{X}_t\right) d\boldsymbol{W}_t,$$

if we make the following definitions: $L_n : \left(\mathbb{R}^d\right)^n \rightarrow \mathcal{P}\left(\mathbb{R}^d\right)$ denotes the empirical measure map,

$$L_n\left(x_1, \ldots, x_n\right) = \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k},$$

where we notice that the range of L_n actually lies in $\mathcal{P}^p\left(\mathbb{R}^d\right)$ for any exponent p. Define also $B : \mathbb{R}^{\mathrm{nd}} \rightarrow \mathbb{R}^{\mathrm{nd}}$ and $\Sigma : \mathbb{R}^{\mathrm{nd}} \rightarrow \mathbb{R}^{\mathrm{nd}} \times \mathbb{R}^{nm}$, for $\boldsymbol{x} = \left(x_1, \ldots, x_n\right) \in \mathbb{R}^{\mathrm{nd}}$, by

$$B(\boldsymbol{x}) = \begin{pmatrix} b\left(x_1, L_n(\boldsymbol{x})\right) \\ b\left(x_2, L_n(\boldsymbol{x})\right) \\ \vdots \\ b\left(x_n, L_n(\boldsymbol{x})\right) \end{pmatrix},$$

$$\Sigma(\boldsymbol{x}) = \begin{pmatrix} \sigma\left(x_1, L_n(\boldsymbol{x})\right) & & & \\ & \sigma\left(x_2, L_n(\boldsymbol{x})\right) & & \\ & & \ddots & \\ & & & \sigma\left(x_n, L_n(\boldsymbol{x})\right) \end{pmatrix},$$

where Σ contains all zeros except for $d \times m$ blocks on the main diagonal. Therefore, we have transform it to a classic SDE to discuss the uniqueness. Finally, we need to point out that while deriving the upper bound of the Wasserstein distance, we use empirical measure $\pi = \frac{1}{n} \sum_{i=1}^{n} \delta_{\left(x_i, y_i\right)}$ to describe the coupling of $L_n(x)$ and $L_n(y)$. $\square$

In fact, as $n \rightarrow \infty$, we can expect the interaction to vanish. If particles do not affect the measure flow, then the particles should be i.i.d., as all of their coefficients b and σ are driven by independent Brownian motions. Thus the dynamics of any particle will look like

$$dY_t^i = b\left(Y_t^i, \mu_t\right) dt + \sigma\left(Y_t^i, \mu_t\right) dW_t^i,$$

where μ_t is a deterministic measure flow and it should still somehow represent the distribution of all of these particles though they are i.i.d..

We know that the empirical measure of i.i.d. samples converges to the true distribution (Theorem above), so we should expect that μ_t is actually the law of Y_t^i , for any i. In other words, the law of the solution shows up in the coefficients of the SDE. We call this a **McKean-Vlasov equation**

2.2 McKean-Vlasov equation

Lift the discussion to the path space and we will fix a time horizon $T > 0$ for convenience. Let $\mathcal{C}^d = C([0,T]);\mathbb{R}^d)$ denote the set of continuous $\mathbb{R}^d$-valued functions of time, equipped with the supremum norm $\|x\| = sup_{t\in[0,T]}|x_t|$ and the corresponding Borel σ-field.
Actually, we will work with probability measures on $\mathcal{C}^d$. There is a natural surjection

$$\mathcal{P}^2\left(\mathcal{C}^d\right) \ni \mu \to (\mu_t)_{t\in[0,T]} \in C\left([0,T];\mathcal{P}^2\left(\mathbb{R}^d\right)\right),$$

where μ_t is defined as the image of the measure μ (ensure square-integrability) through the map $\mathcal{C}^d \ni x \mapsto x_t \in \mathbb{R}^d$.

The McKean-Vlasov equation is **defined** precisely as follows:

$$\begin{aligned} &dY_t = b\left(Y_t,\mu_t\right)dt + \sigma\left(Y_t,\mu_t\right)dW_t, \quad t\in[0,T], \quad Y_0 = \xi, \\ &\mu = \mathcal{L}(Y), \quad \forall t \geq 0. \end{aligned}$$

Here we write $\mathcal{L}(Z)$ for the law or distribution of a random variable Z. Here W is a Brownian motion, ξ is an $\mathbb{R}^d$-valued random variable with the same law as ξ^i, and both are (say) supported on the same filtered probability space $(\Omega,\mathcal{F},\mathcal{F},\mathbb{P})$.

A strong solution of the McKean-Vlasov equation is a pair (Y,μ), where Y is a $C\left([0,T];\mathbb{R}^d\right)$-valued random variable, and μ is a probability measure on $C\left([0,T];\mathbb{R}^d\right)$.
Under the Lipschitz assumption, we can always uniquely solve for Y if we know μ, and so we sometimes refer to μ itself (instead of the pair (Y,μ)) as the solution of the McKean-Vlasov equation.

Theorem.
Suppose Assumption holds. There exists a unique strong solution of the McKean-Vlasov equation. Moreover, the n-particle system converges in the following two senses. First,

$$\lim_{n\to\infty} \mathbb{E}\left[w^2_{c^d,2}\left(\mu^n,\mu\right)\right] = 0.$$

Second, for a fixed $k\in\mathbb{N}$, we have the weak convergence

$$\left(X^{n,1},\ldots,X^{n,k}\right) \Rightarrow \left(Y^1,\ldots,Y^k\right),$$

where $Y^1,\ldots,Y^k$ are independent copies of the solution of the McKean-Vlasov equation.

- particles $X^{n,i}$ become asymptotically i.i.d. as $n\to\infty$.
- the "first k " particles here is inconsequential for any fixed k

Remark. *propagation of chaos*: For any $m_0 \in \mathcal{P}(\mathbb{R}^d)$ and any choice of deterministic initial states $(X_0^{n,i}$) satisfying "m_0-chaotic"

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{X_0^{n,i}} \to m_0,$$

we have the weak limit $\frac{1}{n}\sum_{i=1}^{n}\delta_{X_t^{n,i}} \to \mu_t$ in probability in $\mathcal{P}(\mathbb{R}^d)$, where (Y,μ) is the solution of the McKean-Vlasov equation with initial state $\xi \sim m_0$.

Proof. "existence and uniqueness" + proof of the limit
The first part lies mostly on classical theory from the usual Picard iteration argument after transforming the SDE to a problem searching for its fixed point.
As for the original particle system, define Y_i as the solution of the SDE

$$dY_t^i = b\left(Y_t^i, \mu_t\right)dt + \sigma\left(Y_t^i, \mu_t\right)dW_t, \quad t \in [0,T], \quad Y_0^i = \xi^i,$$

which is i.i.d. Then we estimate the difference $|X_t^{n,i} - Y_t^i|$, combining with the empirical measure. Repeatedly apply Gronwall's inequality we prove the convergence. □

2.3 From SDE to PDE

Suppose (Y,μ) solves the McKean-Vlasov equation. Apply Itô's formula to $\varphi\left(Y_t\right)$, where φ is a smooth function with compact support, to get

$$\begin{aligned} d\varphi\left(Y_t\right) = &\left(b\left(Y_t,\mu_t\right)\cdot\nabla\varphi\left(Y_t\right) + \frac{1}{2}\operatorname{Tr}\left[\sigma\sigma^{\top}\left(Y_t,\mu_t\right)\nabla^2\left(Y_t\right)\right]\right)dt \\ &+ \nabla\varphi\left(Y_t\right)\cdot\sigma\left(Y_t,\mu_t\right)dW_t \end{aligned}$$

Integrating this equation, taking expectations to kill the martingale term, and then differentiating, we find

$$\frac{d}{dt}\mathbb{E}\left[\varphi\left(Y_t\right)\right] = \mathbb{E}\left[b\left(Y_t,\mu_t\right)\cdot\nabla\varphi\left(Y_t\right) + \frac{1}{2}\operatorname{Tr}\left[\sigma\sigma^{\top}\left(Y_t,\mu_t\right)\nabla^2\left(Y_t\right)\right]\right].$$

Now, we know that $Y_t \sim \mu_t$. Suppose in addition that μ_t has a density (with respect to Lebesgue measure), which we write as $\mu(t,x)$.

$$\begin{aligned} &\frac{d}{dt}\int_{\mathbb{R}^d}\varphi(x)\mu(t,x)dx \\ &= \int_{\mathbb{R}^d}\left(b\left(x,\mu_t\right)\cdot\nabla\varphi(x) + \frac{1}{2}\operatorname{Tr}\left[\sigma\sigma^{\top}\left(x,\mu_t\right)\nabla^2\varphi(x)\right]\right)\mu(t,x)dx \\ &= \int_{\mathbb{R}^d}\left(-\operatorname{div}_x\left(b\left(x,\mu_t\right)\mu(t,x)\right) + \frac{1}{2}\operatorname{Tr}\left[\nabla^2\left(\sigma\sigma^{\top}\left(x,\mu_t\right)\mu(t,x)\right)\right]\right)\varphi(x)dx. \end{aligned}$$

Because this must hold for every test function φ, we conclude (formally) that $\mu(t,x)$ solves the PDE

$$\partial_t\mu(t,x) = -\operatorname{div}_x\left(b\left(x,\mu_t\right)\mu(t,x)\right) + \frac{1}{2}\operatorname{Tr}\left[\nabla^2\left(\sigma\sigma^{\top}\left(x,\mu_t\right)\mu(t,x)\right)\right].$$

Remark.
The solution has nonlinear dependence on μ_t and also typically *nonlinear.*
We can point out that if μ is a solution of the Mckean-Vlasov equation, then it is also a weak solution of the PDE above. The converse can often be shown as well.
The PDE can be used as the basis for studying the $n \to \infty$ limit of the n-particle system.

Let's start with this behavior of the $\mathcal{P}\left(\mathbb{R}^d\right)$-valued stochastic process $(\mu_t^n)_{t\geq 0}$ through the integrals of test functions. Shorthand notation

$$\langle \nu, \varphi \rangle := \int_{R^d} \varphi d\nu,$$

for a measure ν on $\mathbb{R}^d$ and a ν-integrable function φ.
Fix a smooth function φ on $\mathbb{R}^d$ with compact support. To identify the behavior of

$$\langle \mu_t^n, \varphi \rangle = \frac{1}{n}\sum_{i=1}^{n} \varphi(X_t^{n,i}),$$

use Ito's formula to write, for each $i = 1, \ldots, n$ and define the infinitesimal generator by setting

$$L_{\mathrm{m}}\varphi(x) = b(x,m)\cdot \nabla_{\varphi}(x) + \frac{1}{2}\operatorname{Tr}\left[\sigma\sigma^{\top}(x,m)\nabla^2\varphi(x)\right].$$

for each $m \in \mathcal{P}(R^d)$ and $x \in \mathbb{R}^d$. Average over $i = 1, \ldots, n$ to get

$$\begin{aligned} d\left\langle \mu_t^n, \varphi \right\rangle &= \frac{1}{n}\sum_{i=1}^{n} d\varphi(X_t^{n,i}) \\ &= \left\langle \mu_t^n, L_{\mu_t^n}\varphi \right\rangle dt + \frac{1}{n}\sum_{i=1}^{n} \nabla\varphi(X_t^{n,i})\cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i \\ &=: \left\langle \mu_t^n, L_{\mu_t^n}\varphi \right\rangle dt + dM_t^n, \end{aligned}$$

where the last line defines the local martingale M^n. In the simplest case, the fanction σ is uniformly bounded, and so there exists a constant C such that $|\sigma^{\top}\nabla\varphi| \leq C$ uniformly. Then, M^n is a martingale with quadratic variation

$$[M^n]_t = \frac{1}{n^2}\sum_{i=1}^{n}\int_0^t |\sigma^{\top}(X_s^{n,i}, \mu_s^n)\nabla\varphi(X_s^{n,i})|^2 ds \leq \frac{tC^2}{n}.$$

In particular, this implies $\mathbb{E}[(M_t^n)^2] \leq tC^2/n$.
$(\mu_t^n)_{t\geq 0}$ is tight as a sequence of random elements. By Prokhorov's theorem, it admits a subsequential limit point. As $n \to \infty$, we can derive $d\left\langle \mu_t, \varphi \right\rangle = \left\langle \mu_t, L_{\mu_t}\varphi \right\rangle dt$. Right now we know the identity holds a.s. for all smooth φ of compact support while with the work on switching the order of **quantifiers** one can derive the equation holds for all φ and all $t \geq 0$ with probability 1. This shows that the limit point μ is almost surely a weak solution of the PDE. **Loss of the Markov property.**
The classic Markov property holds as the coefficients only depend on (t, Y_t).
As for standard SDEs. Suppose we are given coefficients b and σ, and suppose the SDE

$$dX_t = b\left(X_t\right) dt + \sigma\left(X_t\right) dB_t, \quad X_0 = x,$$

has a unique solution for any initial state x. In fact, the SDE above has the even stronger property of defining what is often called a ***Markov family***. letting $\mathcal{C}^d = C\left([0,\infty); \mathbb{R}^d\right)$ denote the path space, suppose we define $P_x \in \mathcal{P}\left(\mathcal{C}^d\right)$ as the law of the solution starting from $x \in \mathbb{R}^d$. Now suppose we randomize the initial state. Consider an initial distribution

$m \in \mathcal{P}\left(\mathbb{R}^d\right)$. Then the SDE

$$dX_t = b\left(X_t\right)dt + \sigma\left(X_t\right)dB_t, \quad X_0 \sim m,$$

has a unique (in law) solution, and we may write $P_m \in \mathcal{P}\left(\mathcal{C}^d\right)$ for the law of the solution. An instance of the "Markov family" property is the statement that for any Borel set $A \subset \mathcal{C}^d$ we have

$$P_m(A) = \int_{\mathbb{R}^d} P_x(A)m(dx).$$

In other words, if we solve the SDE from every deterministic initial state, this is enough to determine the behavior of the SDE started from a random initial state, simply by integrating over the distribution of the initial state. Another way to say this: under P_m, the conditional law of the process X given $X_0 = x$ is equal to P_x.
With McKean-Vlasov equations, the situation is different. If we solve the McKean-Vlasov SDE with a random initial state, and then condition on $X_0 = x$ for some $x \in \mathbb{R}^d$, the resulting distribution is NOT the same as the law P_x obtained by solving the SDE with deterministic initial state $X_0 = x$.

2.4 Common Noise

We consider an important extension of the main model to allow some correlations between the driving Brownian motions. By adding the independent "common noise" term B, called "aggregate shocks" by economists, which effect the system as a whole as opposed to a single particle.

$$dX_t^{n,i} = b\left(X_t^{n,i}, \mu_t^n\right)dt + \sigma\left(X_t^{n,i}, \mu_t^n\right)dW_t^i + \sigma_0\left(X_t^{n,i}, \mu_t^n\right)dB_t, \quad X_0^{n,i} = \xi^i$$

Conditional Mckean-Vlasov equation.

$$\begin{aligned} &dY_t^i = b\left(Y_t, \mu_t\right)dt + \sigma\left(Y_t, \mu_t\right)dW_t + \sigma_0\left(Y_t, \mu_t\right)dB_t, t \in [0,T], Y_0 = \xi, \\ &\mu = \mathcal{L}(Y|B), \quad \forall t \geq 0. \end{aligned}$$

If Y is $\mathbb{F}$-adapted and B is an $\mathbb{F}$-Brownian motion, and if $\mu = \mathcal{L}(Y|B)$, then it holds automatically that

$$\mathcal{L}(Y_t|B) = \mathcal{L}(Y_t|\mathcal{F}_t^B),\ a.s., \forall t \in [0,T],$$

where $\mathcal{F}_t^B = \sigma(B_s : s \leq t)$ is the filtration generated by B.(future increments are independent of $\mathcal{F}_t$) Similarly, we can derive a theorem indicating uniqueness and existence of the solution.

Theorem.
Suppose Assumption holds, with the new coefficient σ_0 satisfying the same Lipschitz assumption. There exists a unique strong solution of the conditional McKean-Vlasov equation (3.8). Moreover, the n particle system converges in the following two senses. First,

$$\lim_{n \to \infty} \mathbb{E}\left[\mathcal{W}_{\mathcal{C}^d,2}^2\left(\mu^n, \mu\right)\right] = 0.$$

Second, for a fixed $k \in \mathbb{N}$, we have the weak convergence

$$\mathcal{L}\left(\left(X^{n,1}, \ldots, X^{n,k}\right) \mid B\right) \to \mu^{\otimes k}, \text{ weakly in probability}$$

where $\mu^{\otimes k}$ denotes the k-fold product measure of μ with itself.

The key idea of the limit theorem is now to construct not i.i.d. but rather conditionally i.i.d. copies of the solution of the McKean-Vlasov equation, driven by the same Brownian motions as the n-particle systems. That is, let

$$dY_t^i = b\left(Y_t^i, \mu_t\right) dt + \sigma\left(Y_t^i, \mu_t\right) dW_t^i + \sigma_0\left(Y_t^i, \mu_t\right) dB_t, \quad t \in [0,T], \quad Y_0 = \xi^i,$$

where μ solves the McKean-Vlasov equation. The two most pertinent details to check are as follows.

- estimate the distance between two conditional measure of the form $\mathcal{L}\left(Y^1 \mid B\right)$ and $\mathcal{L}\left(Y^2 \mid B\right)$ and $\mathcal{L}\left(\left(Y^1, Y^2\right) \mid B\right)$ defines a coupling of the two, almost surely.
- now we need to invoke a *conditional law* of large numbers.

Also, we can derive an alternative derivation by setting the infinitesimal generator,

$$L_{\mathrm{m}}\varphi(x) = b(x,m) \cdot \nabla\varphi(x) + \frac{1}{2} \operatorname{Tr}\left[(\sigma\sigma^\top(x,m) + \sigma_0\sigma_0^\top(x,m)\nabla^2\varphi(x)\right],$$

and the martingale

$$M_s^n = \int_0^s \frac{1}{n} \sum_{i=1}^n \nabla\varphi(X_t^{n,i}) \cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i.$$

2.5 Long-time behavior

Consider a smooth function V on $\mathbb{R}^k$, and consider the k-dimensional SDE

$$dX_t = \nabla V\left(X_t\right) dt + \sigma dW_t,$$

where $\sigma > 0$ is scalar and W a k-dimensional Brownian motion. Under suitable conditions on V, it is well known that X ergodic with invariant distribution

$$\rho(dx) = \frac{1}{Z} e^{\frac{2}{\sigma^2} V(x)} dx,$$

To keep things concrete, we focus on the following specific model of n one-dimensional

($d = 1$) particles:

$$dX_t^{n,i} = \left[a\left(\bar{X}_t^n - X_t^{n,i} \right) + \left(X_t^{n,i} - \left| X_t^{n,i} \right|^3 \right) \right] dt + \sigma dW_t^i,$$
$$\bar{X}_t^n = \frac{1}{n} \sum_{k=1}^{n} X_t^{n,k}.$$

References

Professor Daniel Lacker, IEOR at Columbia University.Rough lecture notes from the Spring 2018 PhD course (IEOR E8100) on mean field games and interacting diffusion models.