Weak particle expansions for McKean-Vlasov SDEs

Lukasz Szpruch

School of Mathematics
University of Edinburgh

joint work with Jean-Francois Chassagneux, Alvin Tse

BSDEs, Information and McKean-Vlasov equations, Leeds 2018

McKV-SDEs

Let $(\Omega, \mathcal{F}, \mathbb{P})$ and $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^q$-valued Wiener process. For $t \geq 0$, the McKean–Vlasov equation (McKV-SDE)

$$X_t^{s,\xi} = \xi + \int_s^t b(X_r^{s,\xi}, \mathscr{L}(X_r^{s,\xi}))\,dr + \int_s^t \sigma(X_r^{s,\xi}, \mathscr{L}(X_r^{s,\xi}))\,dW_r, \quad t \in [s, T],$$

For $\mu_t := \mathscr{L}(X_t)$ we have, for $t \in I$,

$$\langle \mu_t, \varphi \rangle = \langle \mu_0, \varphi \rangle + \int_0^t \left\langle \mu_s, b(s, \cdot, \mu_s)\partial_x \varphi + \tfrac{1}{2}\mathrm{tr}(\sigma\sigma^*)(s, \cdot, \mu_s)\partial_x^2 \varphi \right\rangle \, ds \quad \forall \varphi \in C_0^2(\mathbb{R}^d)\,.$$

McKV-SDEs

Let $(\Omega, \mathcal{F}, \mathbb{P})$ and $(W_t)_{t\in[0,T]}$ be an $\mathbb{R}^q$-valued Wiener process. For $t \geq 0$, the McKean–Vlasov equation (McKV-SDE)

$$X_t^{s,\xi} = \xi + \int_s^t b(X_r^{s,\xi}, \mathscr{L}(X_r^{s,\xi}))\, dr + \int_s^t \sigma(X_r^{s,\xi}, \mathscr{L}(X_r^{s,\xi}))\, dW_r, \quad t \in [s, T],$$

For $\mu_t := \mathscr{L}(X_t)$ we have, for $t \in I$,

$$\langle \mu_t, \varphi \rangle = \langle \mu_0, \varphi \rangle + \int_0^t \left\langle \mu_s, b(s, \cdot, \mu_s)\partial_x \varphi + \tfrac{1}{2}\mathrm{tr}(\sigma\sigma^*)(s, \cdot, \mu_s)\partial_x^2 \varphi \right\rangle ds \quad \forall \varphi \in C_0^2(\mathbb{R}^d)\,.$$

Particle approximation

McKean-Vlasov SDE can be derived as a limit of interacting diffusions:

$$\begin{cases} Y_T^{i,N} & = \xi_i + \int_0^t b\Big(Y_r^{i,N}, \mu_r^N\Big)\,dr + \int_0^t \sigma\Big(Y_r^{i,N}, \mu_r^N\Big)\,dW_r^i, \quad 1 \le i \le N, \quad t \in [0,T], \\ \mu_t^N & := \frac{1}{N}\sum_{j=1}^N \delta_{Y_t^{j,N}} \end{cases}$$

where W^i and ξ_i,, $1 \le i \le N$, are iid.

- Propagation of chaos says: for any $k \le N$, $\mathscr{L}(Y^{1,N}, \dots, Y^{k,N}) \Rightarrow \prod_{i=1}^k \mathscr{L}(X^i)$ when $N \to \infty$.
- "Strong" propagation of chaos $\|Y_T^{i,N} - X_T^i\|_{L^2} \lesssim N^{-1/2}$, where

$$X_t^i = \xi + \int_s^t b(X_r^i, \mathscr{L}(X_r^i))\,dr + \int_s^t \sigma(X_r^i, \mathscr{L}(X_r^i))\,dW_r^i$$

- Not trivial result (no dependence on the dimension) with general measure dependence.
- [Sznitman, 1991],[Méléard, 1996], [Bossy et al., 1996],[Antonelli et al., 2002], [Jourdain and Méléard, 1998]...

Particle approximation

McKean-Vlasov SDE can be derived as a limit of interacting diffusions:

$$\begin{cases} Y_T^{i,N} &= \xi_i + \int_0^t b\left(Y_r^{i,N}, \mu_r^N\right) dr + \int_0^t \sigma\left(Y_r^{i,N}, \mu_r^N\right) dW_r^i, \quad 1 \leq i \leq N, \quad t \in [0,T], \\ \mu_t^N &:= \frac{1}{N} \sum_{j=1}^N \delta_{Y_t^{j,N}} \end{cases}$$

where W^i and ξ_i,, $1 \leq i \leq N$, are iid.

- Propagation of chaos says: for any $k \leq N$, $\mathscr{L}(Y^{1,N}, \ldots, Y^{k,N}) \Rightarrow \prod_{i=1}^k \mathscr{L}(X^i)$ when $N \to \infty$.
- "Strong" propagation of chaos $\|Y_T^{i,N} - X_T^i\|_{L^2} \lesssim N^{-1/2}$, where

$$X_t^i = \xi + \int_s^t b(X_r^i, \mathscr{L}(X_r^i))\, dr + \int_s^t \sigma(X_r^i, \mathscr{L}(X_r^i))\, dW_r^i$$

- Not trivial result (no dependence on the dimension) with general measure dependence.
- [Sznitman, 1991],[Méléard, 1996], [Bossy et al., 1996],[Antonelli et al., 2002], [Jourdain and Méléard, 1998]...

Particle approximation

McKean-Vlasov SDE can be derived as a limit of interacting diffusions:

$$\begin{cases} Y_T^{i,N} &= \xi_i + \int_0^t b\left(Y_r^{i,N}, \mu_r^N\right) dr + \int_0^t \sigma\left(Y_r^{i,N}, \mu_r^N\right) dW_r^i, \quad 1 \le i \le N, \quad t \in [0,T], \\ \mu_t^N &:= \frac{1}{N} \sum_{j=1}^N \delta_{Y_t^{j,N}} \end{cases}$$

where W^i and ξ_i,, $1 \le i \le N$, are iid.

- Propagation of chaos says: for any $k \le N$, $\mathscr{L}(Y^{1,N}, \dots, Y^{k,N}) \Rightarrow \prod_{i=1}^k \mathscr{L}(X^i)$ when $N \to \infty$.
- "Strong" propagation of chaos $\|Y_T^{i,N} - X_T^i\|_{L^2} \lesssim N^{-1/2}$, where

$$X_t^i = \xi + \int_s^t b(X_r^i, \mathscr{L}(X_r^i))\, dr + \int_s^t \sigma(X_r^i, \mathscr{L}(X_r^i))\, dW_r^i$$

- Not trivial result (no dependence on the dimension) with general measure dependence.
- [Sznitman, 1991],[Méléard, 1996], [Bossy et al., 1996],[Antonelli et al., 2002], [Jourdain and Méléard, 1998]...

Towards Weak expansion

Let $\Phi : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$. Consider weak error

$$\Big|\Phi(\mathscr{L}(X_T^{0,\xi})) - \mathbb{E}[\Phi(\mu_T^N)]\Big|,$$

- It is well understood that study of the weak error relies on PDE theory

Define

$$\mathcal{V}(s,\mu) := \Phi(\mathscr{L}(X_T^{s,\mu})), \quad \text{for } (s,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d),$$

One can show that [Chassagneux et al., 2014][Buckdahn et al., 2017]

$$\begin{cases} \partial_s \mathcal{V}(s,\mu) + \int_{\mathbb{R}^d} \left[\partial_\mu \mathcal{V}(s,\mu)(y) b(y,\mu) + \frac{1}{2}\mathrm{tr}\big(\partial_v \partial_\mu \mathcal{V}(s,\mu)(y) a(y,\mu)\big)\right] \mu(dy) = 0, \\ \mathcal{V}(T,\mu) = \Phi(\mu), \end{cases}$$

where $a = (a_{i,k})_{1 \le i,k \le d} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$ denotes the diffusion operator

$$a_{i,k}(x,\mu) := \sum_{j=1}^{q} \sigma_{i,j}(x,\mu)\sigma_{k,j}(x,\mu), \qquad \forall x \in \mathbb{R}^d, \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Towards Weak expansion

Let $\Phi : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$. Consider weak error

$$\left|\Phi(\mathscr{L}(X_T^{0,\xi})) - \mathbb{E}[\Phi(\mu_T^N)]\right|,$$

- It is well understood that study of the weak error relies on PDE theory

Define

$$\mathcal{V}(s,\mu) := \Phi(\mathscr{L}(X_T^{s,\mu})), \quad \text{for } (s,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d),$$

One can show that [Chassagneux et al., 2014][Buckdahn et al., 2017]

$$\begin{cases} \partial_s \mathcal{V}(s,\mu) + \int_{\mathbb{R}^d} \left[\partial_\mu \mathcal{V}(s,\mu)(y) b(y,\mu) + \frac{1}{2}\mathrm{tr}\big(\partial_v \partial_\mu \mathcal{V}(s,\mu)(y) a(y,\mu)\big)\right] \mu(dy) = 0, \\ \mathcal{V}(T,\mu) = \Phi(\mu), \end{cases}$$

where $a = (a_{i,k})_{1 \le i,k \le d} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$ denotes the diffusion operator

$$a_{i,k}(x,\mu) := \sum_{j=1}^{q} \sigma_{i,j}(x,\mu)\sigma_{k,j}(x,\mu), \qquad \forall x \in \mathbb{R}^d, \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Towards Weak expansion

Let $\Phi : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$. Consider weak error

$$\left|\Phi(\mathscr{L}(X_T^{0,\xi})) - \mathbb{E}[\Phi(\mu_T^N)]\right|,$$

- It is well understood that study of the weak error relies on PDE theory

Define

$$\mathcal{V}(s,\mu) := \Phi(\mathscr{L}(X_T^{s,\mu})), \quad \text{for } (s,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)\,,$$

One can show that [Chassagneux et al., 2014][Buckdahn et al., 2017]

$$\begin{cases} \partial_s \mathcal{V}(s,\mu) + \int_{\mathbb{R}^d} \left[\partial_\mu \mathcal{V}(s,\mu)(y) b(y,\mu) + \frac{1}{2}\mathrm{tr}\big(\partial_v \partial_\mu \mathcal{V}(s,\mu)(y) a(y,\mu)\big)\right] \mu(dy) = 0, \\ \mathcal{V}(T,\mu) = \Phi(\mu), \end{cases}$$

where $a = (a_{i,k})_{1 \le i,k \le d} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$ denotes the diffusion operator

$$a_{i,k}(x,\mu) := \sum_{j=1}^{q} \sigma_{i,j}(x,\mu)\sigma_{k,j}(x,\mu), \qquad \forall x \in \mathbb{R}^d, \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Measure derivative, due to Lions [Cardaliaguet, 2013]

Given $u : \mathcal{P} \to \mathbb{R}$ consider

i) an atomless, Polish probability space $(\Omega, \mathcal{F}, \mathbb{P})$,

ii) $X \in L^2(\Omega)$ s.t. $\mu = \mathscr{L}(X)$,

iii) $U : L^2(\Omega) \to \mathbb{R}$ given by $U(X) := u(\mathscr{L}(X))$ - "lift of u".

Definition (L-Derivative)

The function u is L-differentiable at μ if its lift U is Frechet differentiable at X. Recall that this means that there is a bounded linear operator $A : L^2(\Omega) \to \mathbb{R}$ s.t.

$$\lim_{|Y|_2 \to 0} \left| \frac{U(X+Y) - U(X)}{|Y|_2} - \frac{AY}{|Y|_2} \right| = 0 .$$

The bounded linear operator defines $DU(X) \in L^2(\Omega)^*$ which is identified with $L^2(\Omega)$ via $(X, Y) = \mathbb{E}[XY]$.

- It can be shown that Lions derivative does not depend on the choice of the lift nor probability space on which lift is defined.
- Close links to Otto calculus in the theory of gradient flows [Ambrosio et al., 2008]
- See also Semigroup on $\mathcal{P}$ [Kolokoltsov, 2010], [Mischler et al., 2015] and [Del Moral et al., 2011]

Measure derivative, due to Lions [Cardaliaguet, 2013]

Given $u : \mathcal{P} \to \mathbb{R}$ consider

i) an atomless, Polish probability space $(\Omega, \mathcal{F}, \mathbb{P})$,

ii) $X \in L^2(\Omega)$ s.t. $\mu = \mathscr{L}(X)$,

iii) $U : L^2(\Omega) \to \mathbb{R}$ given by $U(X) := u(\mathscr{L}(X))$ - "lift of u".

Definition (L-Derivative)

The function u is L-differentiable at μ if its lift U is Frechet differentiable at X. Recall that this means that there is a bounded linear operator $A : L^2(\Omega) \to \mathbb{R}$ s.t.

$$\lim_{|Y|_2 \to 0} \left| \frac{U(X+Y) - U(X)}{|Y|_2} - \frac{A\,Y}{|Y|_2} \right| = 0 .$$

The bounded linear operator defines $DU(X) \in L^2(\Omega)^*$ which is identified with $L^2(\Omega)$ via $(X, Y) = \mathbb{E}[XY]$.

- It can be shown that Lions derivative does not depend on the choice of the lift nor probability space on which lift is defined.
- Close links to Otto calculus in the theory of gradient flows [Ambrosio et al., 2008]
- See also Semigroup on $\mathcal{P}$ [Kolokoltsov, 2010], [Mischler et al., 2015] and [Del Moral et al., 2011]

Measure derivative, due to Lions [Cardaliaguet, 2013]

Given $u : \mathcal{P} \to \mathbb{R}$ consider

i) an atomless, Polish probability space $(\Omega, \mathcal{F}, \mathbb{P})$,

ii) $X \in L^2(\Omega)$ s.t. $\mu = \mathscr{L}(X)$,

iii) $U : L^2(\Omega) \to \mathbb{R}$ given by $U(X) := u(\mathscr{L}(X))$ - "lift of u".

Definition (L-Derivative)

The function u is L-differentiable at μ if its lift U is Frechet differentiable at X. Recall that this means that there is a bounded linear operator $A : L^2(\Omega) \to \mathbb{R}$ s.t.

$$\lim_{|Y|_2 \to 0} \left| \frac{U(X+Y) - U(X)}{|Y|_2} - \frac{A\,Y}{|Y|_2} \right| = 0\,.$$

The bounded linear operator defines $DU(X) \in L^2(\Omega)^*$ which is identified with $L^2(\Omega)$ via $(X, Y) = \mathbb{E}[XY]$.

- It can be shown that Lions derivative does not depend on the choice of the lift nor probability space on which lift is defined.
- Close links to Otto calculus in the theory of gradient flows [Ambrosio et al., 2008]
- See also Semigroup on $\mathcal{P}$ [Kolokoltsov, 2010], [Mischler et al., 2015] and [Del Moral et al., 2011]

Measure derivative, due to Lions [Cardaliaguet, 2013]

Given $u : \mathcal{P} \to \mathbb{R}$ consider

i) an atomless, Polish probability space $(\Omega, \mathcal{F}, \mathbb{P})$,

ii) $X \in L^2(\Omega)$ s.t. $\mu = \mathscr{L}(X)$,

iii) $U : L^2(\Omega) \to \mathbb{R}$ given by $U(X) := u(\mathscr{L}(X))$ - "lift of u".

Definition (L-Derivative)

The function u is L-differentiable at μ if its lift U is Frechet differentiable at X. Recall that this means that there is a bounded linear operator $A : L^2(\Omega) \to \mathbb{R}$ s.t.

$$\lim_{|Y|_2 \to 0} \left| \frac{U(X+Y) - U(X)}{|Y|_2} - \frac{A\,Y}{|Y|_2} \right| = 0\,.$$

The bounded linear operator defines $DU(X) \in L^2(\Omega)^*$ which is identified with $L^2(\Omega)$ via $(X, Y) = \mathbb{E}[XY]$.

- It can be shown that Lions derivative does not depend on the choice of the lift nor probability space on which lift is defined.
- Close links to Otto calculus in the theory of gradient flows [Ambrosio et al., 2008]
- See also Semigroup on $\mathcal{P}$ [Kolokoltsov, 2010], [Mischler et al., 2015] and [Del Moral et al., 2011]

Example: $u(\mu) = \int_{\mathbb{R}^d} f(x)\,\mu(dx) = \langle \mu, f \rangle$, with $f \in C^1$ and $|\nabla_x f(x)| \leq C(1 + |x|)$. The lift $U(X) = \mathbb{E}[f(X)]$. We can check that $DU(X) = \nabla_x f(X)$:

$$U(X+Y) = U(X) + \int_0^1 \mathbb{E}[\nabla_x f(X + \lambda Y) Y]\, d\lambda\,.$$

Then

$$\frac{1}{|Y|_2}\Big| U(X+Y) - U(X) - \mathbb{E}\big[\nabla_x f(X) Y\big]\Big|$$
$$\leq \text{standard estimates}$$
$$\leq C\left(|Y|_2 + |Y|_2^{1/2} + \sup_{\mathbb{P}(A) \leq |Y|_2} \mathbb{E}[X^2 \mathbb{1}_A]\right) \to 0 \text{ as } |Y|_2 \to 0.$$

So $\partial_\mu v(\mu) = \nabla_x f$.

Example: $u(\mu) = \int_{\mathbb{R}^d} f(x)\,\mu(dx) = \langle \mu, f \rangle$, with $f \in C^1$ and $|\nabla_x f(x)| \leq C(1+|x|)$. The lift $U(X) = \mathbb{E}[f(X)]$. We can check that $DU(X) = \nabla_x f(X)$:

$$U(X+Y) = U(X) + \int_0^1 \mathbb{E}[\nabla_x f(X+\lambda Y)Y]\,d\lambda\,.$$

Then

$$\begin{aligned}
&\frac{1}{|Y|_2}\left|U(X+Y) - U(X) - \mathbb{E}\Big[\nabla_x f(X)Y\Big]\right| \\
&\leq \text{standard estimates} \\
&\leq C\left(|Y|_2 + |Y|_2^{1/2} + \sup_{\mathbb{P}(A)\leq|Y|_2} \mathbb{E}[X^2 \mathbb{1}_A]\right) \to 0 \ \text{ as } |Y|_2 \to 0.
\end{aligned}$$

So $\partial_\mu v(\mu) = \nabla_x f$.

Higher order L-derivatives

First derivative of $v : \mathcal{P}_2 \to \mathbb{R}$ is a mapping $\partial_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d$

Fix $\mu \in \mathcal{P}_2$. If $\partial_\mu v(\mu, \cdot) : \mathbb{R}^d \to \mathbb{R}$ is differentiable at some $y \in \mathbb{R}^d$ we write $\partial_y \partial_\mu v(\mu, y)$. So

$$\partial_y \partial_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d .$$

Fix $y \in \mathbb{R}^d$. If $\partial_\mu v(\cdot, y) : \mathcal{P}_2 \to \mathbb{R}$ is L-differentiable at $\mu \in \mathcal{P}_2$ then we get (above proposition) a $\psi : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d$ and we have

$$\partial^2_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d .$$

Higher order L-derivatives

First derivative of $v : \mathcal{P}_2 \to \mathbb{R}$ is a mapping $\partial_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d$
Fix $\mu \in \mathcal{P}_2$. If $\partial_\mu v(\mu, \cdot) : \mathbb{R}^d \to \mathbb{R}$ is differentiable at some $y \in \mathbb{R}^d$ we write $\partial_y \partial_\mu v(\mu, y)$.
So

$$\partial_y \partial_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d \,.$$

Fix $y \in \mathbb{R}^d$. If $\partial_\mu v(\cdot, y) : \mathcal{P}_2 \to \mathbb{R}$ is L-differentiable at $\mu \in \mathcal{P}_2$ then we get (above proposition) a $\psi : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d$ and we have

$$\partial_\mu^2 v : \mathcal{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d \,.$$

Higher order L-derivatives

First derivative of $v : \mathcal{P}_2 \to \mathbb{R}$ is a mapping $\partial_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d$
Fix $\mu \in \mathcal{P}_2$. If $\partial_\mu v(\mu, \cdot) : \mathbb{R}^d \to \mathbb{R}$ is differentiable at some $y \in \mathbb{R}^d$ we write $\partial_y \partial_\mu v(\mu, y)$. So

$$\partial_y \partial_\mu v : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d .$$

Fix $y \in \mathbb{R}^d$. If $\partial_\mu v(\cdot, y) : \mathcal{P}_2 \to \mathbb{R}$ is L-differentiable at $\mu \in \mathcal{P}_2$ then we get (above proposition) a $\psi : \mathcal{P}_2 \times \mathbb{R}^d \to \mathbb{R}^d$ and we have

$$\partial_\mu^2 v : \mathcal{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d .$$

L-derivatives for empirical measures

Let $u : \mathcal{P}_2 \to \mathbb{R}$. Let $\mu \in \mathcal{P}_2$ be given.
Define

$$u^N(x^1, \dots, x^N) := u\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right).$$

Proposition

If v is L-differentiable at μ then for any $N \in \mathbb{N}$ this v^N is differentiable and

$$\partial_{x^k} u^N(x^1, \dots, x^N) = \frac{1}{N}(\partial_\mu u)\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right)(x^k).$$

If u is twice L-differentiable at μ then for any $N \in \mathbb{N}$ this u^N is twice differentiable and

$$\partial_{x^k x^m} v^N(x^1, \dots, x^N) = \frac{1}{N}\partial_y\left[(\partial_\mu u)\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right)\right](x^k)\delta_{k,m} + \frac{1}{N^2}(\partial_\mu^2 u)\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right)(x^k, x^m).$$

L-derivatives for empirical measures

Let $u : \mathcal{P}_2 \to \mathbb{R}$. Let $\mu \in \mathcal{P}_2$ be given.
Define

$$u^N(x^1, \ldots, x^N) := u\left(\frac{1}{N}\sum_{\ell=1}^N \delta_{x^\ell}\right).$$

Proposition

If v is L-differentiable at μ then for any $N \in \mathbb{N}$ this v^N is differentiable and

$$\partial_{x^k} u^N(x^1, \ldots, x^N) = \frac{1}{N}(\partial_\mu u)\left(\frac{1}{N}\sum_{\ell=1}^N \delta_{x^\ell}\right)(x^k).$$

If u is twice L-differentiable at μ then for any $N \in \mathbb{N}$ this u^N is twice differentiable and

$$\partial_{x^k x^m} v^N(x^1, \ldots, x^N)$$
$$= \frac{1}{N}\partial_y\left[(\partial_\mu u)\left(\frac{1}{N}\sum_{\ell=1}^N \delta_{x^\ell}\right)\right](x^k)\delta_{k,m} + \frac{1}{N^2}(\partial_\mu^2 u)\left(\frac{1}{N}\sum_{\ell=1}^N \delta_{x^\ell}\right)(x^k, x^m).$$

L-derivatives for empirical measures

Let $u : \mathcal{P}_2 \to \mathbb{R}$. Let $\mu \in \mathcal{P}_2$ be given.
Define

$$u^N(x^1, \dots, x^N) := u\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right).$$

Proposition

If v is L-differentiable at μ then for any $N \in \mathbb{N}$ this v^N is differentiable and

$$\partial_{x^k} u^N(x^1, \dots, x^N) = \frac{1}{N}(\partial_\mu u)\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right)(x^k).$$

If u is twice L-differentiable at μ then for any $N \in \mathbb{N}$ this u^N is twice differentiable and

$$\begin{aligned}
&\partial_{x^k x^m} v^N(x^1, \dots, x^N) \\
&= \frac{1}{N}\partial_y\left[(\partial_\mu u)\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right)\right](x^k)\delta_{k,m} + \frac{1}{N^2}(\partial_\mu^2 u)\left(\frac{1}{N}\sum_{\ell=1}^{N} \delta_{x^\ell}\right)(x^k, x^m).
\end{aligned}$$

Proof of the expansion

We observe that

- By the terminal condition for the PDE, we notice that $\Phi(\mu_T^N) = \mathcal{V}(T, \mu_T^N)$.
- Therefore, the weak error decomposes as

$$\begin{aligned}\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi})) &= \mathcal{V}(T, \mu_T^N) - \mathcal{V}(0, \mathscr{L}(\xi))\\ &= (\mathcal{V}(T, \mu_T^N) - \mathcal{V}(0, \mu_0^N)) + (\mathcal{V}(0, \mu_0^N) - \mathcal{V}(0, \mathscr{L}(\xi))).\end{aligned}$$

Proof of the expansion

We observe that

- By the terminal condition for the PDE, we notice that $\Phi(\mu^N_T) = \mathcal{V}(T, \mu^N_T)$.
- Therefore, the weak error decomposes as

$$\begin{aligned}\Phi(\mu^N_T) - \Phi(\mathscr{L}(X^{0,\xi}_T)) &= \mathcal{V}(T, \mu^N_T) - \mathcal{V}(0, \mathscr{L}(\xi))) \\ &= \big(\mathcal{V}(T, \mu^N_T) - \mathcal{V}(0, \mu^N_0)\big) + \big(\mathcal{V}(0, \mu^N_0) - \mathcal{V}(0, \mathscr{L}(\xi))\big).\end{aligned}$$

Proof of the expansion

$$\mathcal{V}(T,\mu^N_T) - \mathcal{V}(0,\mu^N_0) = [V(T,\mathbf{Y}^N_T) - V(0,\mathbf{Y}^N_0)]$$

$$= \int_0^T \frac{\partial V}{\partial s}(s,\mathbf{Y^N}_s) + \sum_{i=1}^N \frac{\partial V}{\partial x_i}(s,\mathbf{Y^N}_s) b(Y^{i,N}_s,\mu^N_s) + \frac{1}{2}\mathrm{tr}\left(a(Y^{i,N}_s,\mu^N_s)\sum_{i=1}^N \frac{\partial^2 V}{\partial x_i^2}(s,\mathbf{Y^N}_s)\right) ds\Big]$$

$$+ \int_0^T \sum_{i=1}^N \frac{\partial V}{\partial x_i}(s,\mathbf{Y^N}_s)\sigma(Y^{i,N}_s,\mu^N_s)dW^i_s$$

$$= \int_0^T \partial_s \mathcal{V}(s,\mu^N_s) + \int_{\mathbb{R}^d} \partial_\mu \mathcal{V}(s,\mu^N_s)(y) b(y,\mu^N_s) + \frac{1}{2}\mathrm{tr}\left(a(y,\mu^N_s)\partial_v\partial_\mu \mathcal{V}(s,\mu^N_s)(y)\right)\mu^N_s(dy)\, ds$$

$$+ \frac{1}{2}\int_0^T \frac{1}{N}\int_{\mathbb{R}^d} \mathrm{tr}\left(a(y,\mu^N_s)\partial^2_\mu \mathcal{V}(s,\mu^N_s)(y,y)\right)\mu^N_s(dy)ds + \sum_{i=1}^N \frac{1}{N}\partial_\mu \mathcal{V}(s,\mu^N_s)(Y^{i,N}_s)\sigma(Y^{i,N}_s,\mu^N_s)dW^i_s$$

But recall that

$$\partial_s \mathcal{V}(s,\mu) + \int_{\mathbb{R}^d} \left[\partial_\mu \mathcal{V}(s,\mu)(y) b(y,\mu) + \frac{1}{2}\mathrm{tr}(\partial_v\partial_\mu \mathcal{V}(s,\mu)(y) a(y,\mu))\right]\mu(dy) = 0.$$

Proof of the expansion

$$
\begin{aligned}
&\mathcal{V}(T,\mu^N_T)-\mathcal{V}(0,\mu^N_0)=[V(T,\mathbf{Y}^N_T)-V(0,\mathbf{Y}^N_0)]\\
&=\int_0^T \frac{\partial V}{\partial s}(s,\mathbf{Y^N}_s)+\sum_{i=1}^N \frac{\partial V}{\partial x_i}(s,\mathbf{Y^N}_s)b(Y^{i,N}_s,\mu^N_s)+\frac{1}{2}\mathrm{tr}\Big(a(Y^{i,N}_s,\mu^N_s)\sum_{i=1}^N\frac{\partial^2 V}{\partial x_i^2}(s,\mathbf{Y^N}_s)\Big)\,ds\Big]\\
&+\int_0^T\sum_{i=1}^N\frac{\partial V}{\partial x_i}(s,\mathbf{Y^N}_s)\sigma(Y^{i,N}_s,\mu^N_s)dW^i_s\\
&=\int_0^T \partial_s\mathcal{V}(s,\mu^N_s)+\int_{\mathbb{R}^d}\partial_\mu\mathcal{V}(s,\mu^N_s)(y)b(y,\mu^N_s)+\frac{1}{2}\mathrm{tr}\Big(a(y,\mu^N_s)\partial_v\partial_\mu\mathcal{V}(s,\mu^N_s)(y)\Big)\mu^N_s(dy)\,ds\\
&+\frac{1}{2}\int_0^T\frac{1}{N}\int_{\mathbb{R}^d}\mathrm{tr}\Big(a(y,\mu^N_s)\partial^2_\mu\mathcal{V}(s,\mu^N_s)(y,y)\Big)\mu^N_s(dy)ds+\sum_{i=1}^N\frac{1}{N}\partial_\mu\mathcal{V}(s,\mu^N_s)(Y^{i,N}_s)\sigma(Y^{i,N}_s,\mu^N_s)dW^i_s
\end{aligned}
$$

But recall that

$$
\partial_s\mathcal{V}(s,\mu)+\int_{\mathbb{R}^d}\big[\partial_\mu\mathcal{V}(s,\mu)(y)b(y,\mu)+\frac{1}{2}\mathrm{tr}(\partial_v\partial_\mu\mathcal{V}(s,\mu)(y)a(y,\mu))\big]\mu(dy)=0.
$$

Proof of the expansion

Hence

$$\mathbb{E}\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi})) = \mathbb{E}(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mu_0^N)) + \frac{1}{2N}\int_0^T \mathbb{E}\int_{\mathbb{R}^d} \mathrm{tr}\Big(a(y,\mu_s^N)\partial_\mu^2\mathcal{V}(s,\mu_s^N)(y,y)\Big)\mu_s^N(dy)ds$$

- Strong variant of propagation of chaos also possible for $\mathbb{E}|\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi}))|$

This gives the decomposition

$$\begin{aligned}\mathbb{E}\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi})) &= \mathbb{E}(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mathscr{L}(\xi))) + \frac{C_1}{N} \\ &+ \int_0^T \frac{1}{2N}\Bigg(\mathbb{E}\Bigg[\frac{1}{N}\sum_{i=1}^N \mathrm{tr}\Big(a(Y_s^{i,N},\mu_s^N)\partial_\mu^2\mathcal{V}(s,\mu_{t_2}^N)(Y_s^{i,N},Y_s^{i,N})\Big)\Bigg] \\ &- \mathbb{E}\Big[\mathrm{tr}\Big(a(X_s^{0,\xi},\mathscr{L}(X_s^{0,\xi}))\partial_\mu^2\mathcal{V}(s,\mathscr{L}(X_s^{0,\xi}))(X_s^{0,\xi},X_s^{0,\xi})\Big)\Big]\Bigg)ds.\end{aligned}$$

- Constant C_1 does not depend on N!

Proof of the expansion

Hence

$$\mathbb{E}\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi})) = \mathbb{E}\big(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mu_0^N)\big) + \frac{1}{2N}\int_0^T \mathbb{E}\int_{\mathbb{R}^d} \mathrm{tr}\Big(a(y,\mu_s^N)\partial_\mu^2\mathcal{V}(s,\mu_s^N)(y,y)\Big)\mu_s^N(dy)ds$$

- Strong variant of propagation of chaos also possible for $\mathbb{E}|\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi}))|$

This gives the decomposition

$$\begin{aligned}\mathbb{E}\Phi(\mu_T^N) - \Phi(\mathscr{L}(X_T^{0,\xi})) &= \mathbb{E}\big(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mathscr{L}(\xi))\big) + \frac{C_1}{N} \\ &+ \int_0^T \frac{1}{2N}\Bigg(\mathbb{E}\Bigg[\frac{1}{N}\sum_{i=1}^N \mathrm{tr}\Big(a(Y_s^{i,N},\mu_s^N)\partial_\mu^2\mathcal{V}(s,\mu_{t_2}^N)(Y_s^{i,N},Y_s^{i,N})\Big)\Bigg] \\ &- \mathbb{E}\Big[\mathrm{tr}\Big(a(X_s^{0,\xi},\mathscr{L}(X_s^{0,\xi}))\partial_\mu^2\mathcal{V}(s,\mathscr{L}(X_s^{0,\xi}))(X_s^{0,\xi},X_s^{0,\xi})\Big)\Big]\Bigg)\,ds.\end{aligned}$$

- Constant C_1 does not depend on N!

Proof of the expansion

Hence

$$\mathbb{E}\Phi(\mu^N_T) - \Phi(\mathscr{L}(X^{0,\xi}_T)) = \mathbb{E}(\mathcal{V}(0,\mu^N_0) - \mathcal{V}(0,\mu^N_0)) + \frac{1}{2N}\int_0^T \mathbb{E}\int_{\mathbb{R}^d} \operatorname{tr}\Big(a(y,\mu^N_s)\partial^2_\mu \mathcal{V}(s,\mu^N_s)(y,y)\Big)\mu^N_s(dy)ds$$

- Strong variant of propagation of chaos also possible for $\mathbb{E}|\Phi(\mu^N_T) - \Phi(\mathscr{L}(X^{0,\xi}_T))|$

This gives the decomposition

$$\begin{aligned}\mathbb{E}\Phi(\mu^N_T) - \Phi(\mathscr{L}(X^{0,\xi}_T)) &= \mathbb{E}(\mathcal{V}(0,\mu^N_0) - \mathcal{V}(0,\mathscr{L}(\xi))) + \frac{C_1}{N} \\ &+ \int_0^T \frac{1}{2N}\Bigg(\mathbb{E}\Bigg[\frac{1}{N}\sum_{i=1}^N \operatorname{tr}\Big(a(Y^{i,N}_s,\mu^N_s)\partial^2_\mu \mathcal{V}(s,\mu^N_{t_2})(Y^{i,N}_s,Y^{i,N}_s)\Big)\Bigg] \\ &- \mathbb{E}\Big[\operatorname{tr}\Big(a(X^{0,\xi}_s,\mathscr{L}(X^{0,\xi}_s))\partial^2_\mu \mathcal{V}(s,\mathscr{L}(X^{0,\xi}_s))(X^{0,\xi}_s,X^{0,\xi}_s)\Big)\Big]\Bigg)\,ds.\end{aligned}$$

- Constant C_1 does not depend on N!

Higher order terms

- Let $\mathcal{V}^{(1)} = \mathcal{V}$
- For each $m \in \{2, \ldots, k\}$ and $0 < t_m < \ldots < t_2 < T$, we define $\mathcal{V}^{(m)}_{t_2,\ldots,t_m} : [0, t_m] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ by

$$\mathcal{V}^{(m)}_{t_2,\ldots,t_m}(s, \mu) := \Phi^{(m)}_{t_2,\ldots,t_m}(\mathscr{L}(X^{s,\mu}_{t_m})),$$

where $\Phi^{(m)}_{t_2,\ldots,t_m} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is given by

$$\Phi^{(m)}_{t_2,\ldots,t_m}(\mu) := \int_{\mathbb{R}^d} \operatorname{tr}\left(\partial^2_\mu \mathcal{V}^{(m-1)}_{t_2,\ldots,t_{m-1}}(t_m, \mu)(x, x) a(x, \mu) \right) \mu(dx).$$

Theorem (Weak error expansion)

Assume coefficients of the SDE are $\mathcal{C}^{2k}$. Furthermore, suppose that, for each $m \in \{1, \dots, k\}$ and $0 < t_m < \dots < t_2 < T$, $\mathcal{V}^{(m)}_{t_2,\dots,t_m}$ is a unique classical solution to PDE

$$\begin{cases} \partial_s \mathcal{V}^{(m)}_{t_2,\dots,t_m}(s,\mu) + \int_{\mathbb{R}^d} \left[\partial_\mu \mathcal{V}^{(m)}_{t_2,\dots,t_m}(s,\mu)(x) b(x,\mu) + \frac{1}{2} tr\big(\partial_v \partial_\mu \mathcal{V}^{(m)}_{t_2,\dots,t_m}(s,\mu)(x) a(x,\mu)\big)\right] \mu(dx) = 0, \\ \mathcal{V}^{(m)}_{t_2,\dots,t_m}(t_m,\mu) = \Phi^{(m)}_{t_2,\dots,t_m}(\mu). \end{cases}$$

Then the weak error in the particle approximation can be expressed as

$$\mathbb{E}[\Phi(\mu^N_T)] - \Phi(\mathscr{L}(X^{0,\xi}_T)) = \mathbb{E}\big(\mathcal{V}(0,\mu^N_0) - \mathcal{V}(0,\mathscr{L}(\xi))\big) + \frac{C_1}{N} + \dots + \frac{C_{k-1}}{N^{k-1}} + O\left(\frac{1}{N^k}\right) + \mathscr{R}^0,$$

where $C_1, \dots, C_{k-1}$ are constants that depend on Φ, b, σ and T, but are independent of N and

$$\mathscr{R}^0 := \sum_{p=1}^{k-1} \int_0^T \int_0^{t_2} \dots \int_0^{t_p} \frac{1}{(2N)^p} \mathbb{E}\Big(\mathcal{V}^{(p+1)}_{t_2,\dots,t_{p+1}}(0,\mu^N_0) - \mathcal{V}^{(p+1)}_{t_2,\dots,t_{p+1}}(0,\mathscr{L}(\xi))\Big)\, dt_{p+1}\dots dt_2$$

Romberg Extrapolation

$$\mathbb{E}f(Y_T^{i,N}) - \mathbb{E}f(X_T) = \frac{C}{N} + O\left(\frac{1}{N^2}\right)$$

and

$$\mathbb{E}f(Y_T^{i,2N}) - \mathbb{E}f(X_T) = \frac{C}{2N} + O\left(\frac{1}{N^2}\right).$$

Hence,

$$\left|\left(2\mathbb{E}f(Y_T^{i,2N}) - \mathbb{E}f(Y_T^{i,N})\right) - \mathbb{E}f(X_T)\right| = O\left(\frac{1}{N^2}\right).$$

For general m, we can use a similar method to show that

$$\left|\sum_{k=1}^{m} \alpha_k \mathbb{E}f(Y_T^{i,kN}) - \mathbb{E}f(X_T)\right| = O\left(\frac{1}{N^m}\right),$$

where

$$\alpha_k = (-1)^{m-k}\frac{k^m}{k!(m-k)!}, \quad 1 \leq k \leq m.$$

Weak Error and ensemble particle system

One should study

$$|\mathbb{E}[f(X_t)] - \sum_{k=1}^{m} \alpha_k \mathbb{E} f(Y_t^{i,kN})|$$

Motivation: M-ensambles

$$\mathbb{E}\Big[\Big(\mathbb{E}[f(X_t)] - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_T^{i,j,kN})\Big)^2\Big]$$

$$\leq \Big[\Big(\mathbb{E}[f(X_t)] - \sum_{k=1}^{m}\alpha_k \mathbb{E} f(Y_t^{i,kN})\Big)^2\Big]$$

$$+ \mathbb{E}\Big[\Big(\mathbb{E}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_t^{i,kN}) - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_T^{i,j,kN})\Big)^2\Big]$$

- The first term is of order N^{-2m} the second of order $M^{-1}\sum_k \alpha_k^2 (kN)^{-1}$
- Cost $\mathcal{C} = MN^2$ for the mean-square-error ϵ^2 is $\mathcal{C} = \mathcal{O}(\epsilon^{-2-1/m})$

Weak Error and ensemble particle system

One should study

$$|\mathbb{E}[f(X_t)] - \sum_{k=1}^{m} \alpha_k \mathbb{E} f(Y_t^{i,kN})|$$

Motivation: M-ensambles

$$\mathbb{E}\Bigg[\bigg(\mathbb{E}[f(X_t)] - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_T^{i,j,kN})\bigg)^2\Bigg]$$

$$\leq \Bigg[\bigg(\mathbb{E}[f(X_t)] - \sum_{k=1}^{m}\alpha_k \mathbb{E} f(Y_t^{i,kN})\bigg)^2\Bigg]$$

$$+\, \mathbb{E}\Bigg[\bigg(\mathbb{E}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_t^{i,kN}) - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_T^{i,j,kN})\bigg)^2\Bigg]$$

- The first term is of order N^{-2m} the second of order $M^{-1}\sum_k \alpha_k^2 (kN)^{-1}$
- Cost $\mathcal{C} = MN^2$ for the mean-square-error ϵ^2 is $\mathcal{C} = \mathcal{O}(\epsilon^{-2-1/m})$

Weak Error and ensemble particle system

One should study

$$|\mathbb{E}[f(X_t)] - \sum_{k=1}^{m} \alpha_k \mathbb{E} f(Y_t^{i,kN})|$$

Motivation: M-ensambles

$$\mathbb{E}\left[\left(\mathbb{E}[f(X_t)] - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_T^{i,j,kN})\right)^2\right]$$

$$\leq \left[\left(\mathbb{E}[f(X_t)] - \sum_{k=1}^{m}\alpha_k \mathbb{E} f(Y_t^{i,kN})\right)^2\right]$$

$$+ \mathbb{E}\left[\left(\mathbb{E}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_t^{i,kN}) - \frac{1}{M}\sum_{j=1}^{M}\sum_{k=1}^{m}\alpha_k \frac{1}{kN}\sum_{i=1}^{kN} f(Y_T^{i,j,kN})\right)^2\right]$$

- The first term is of order N^{-2m} the second of order $M^{-1}\sum_k \alpha_k^2 (kN)^{-1}$
- Cost $\mathcal{C} = MN^2$ for the mean-square-error ϵ^2 is $\mathcal{C} = \mathcal{O}(\epsilon^{-2-1/m})$

Antithetic trick

Take $2N$ independent samples from μ_0 and $2N$ independent Brownian motions $(\xi^i, W^i)_{i=1\dots 2N}$. Next consider a system of particles

$$Y_t^i = \xi_i + \int_0^t b\left(Y_r^i, \mu_r^{2N}\right) dr + \int_0^t \sigma\left(Y_r^i, \mu_r^{2N}\right) dw_r^i, \quad 1 \leq i \leq 2N.$$

Next take two non-intersecting subsets of $(\xi^i, w^i)_{i=1\dots 2N}$ consisting of N pairs each i.e take $(\xi^i, w^i)_{i=1\dots N}$ and $(\xi^i, W^i)_{i=N+1\dots 2N}$. Define

$$Y_t^{i,(1)} = \xi_i + \int_0^t b\left(Y_r^{i,(1)}, \mu_r^{N,(1)}\right) dr + \int_0^t \sigma\left(Y_r^{i,(1)}, \mu_r^{N,(1)}\right) dw_r^i, \quad 1 \leq i \leq N,$$
$$Y_t^{i,(2)} = \xi_i + \int_0^t b\left(Y_r^{i,(2)}, \mu_r^{N,(2)}\right) dr + \int_0^t \sigma\left(Y_r^{i,(2)}, \mu_r^{N,(2)}\right) dw_r^i, \quad N+1 \leq i \leq 2N,$$

Study

$$\mathbb{V}ar[\Phi(\mu_t^{2N}) - \frac{1}{2}(\Phi(\mu_t^{N,(1)}) + \Phi(\mu_t^{N,(2)}))].$$

Antithetic trick

Take $2N$ independent samples from μ_0 and $2N$ independent Brownian motions $(\xi^i, W^i)_{i=1\ldots 2N}$. Next consider a system of particles

$$Y_t^i = \xi_i + \int_0^t b\left(Y_r^i, \mu_r^{2N}\right) dr + \int_0^t \sigma\left(Y_r^i, \mu_r^{2N}\right) dw_r^i, \quad 1 \le i \le 2N.$$

Next take two non-intersecting subsets of $(\xi^i, w^i)_{i=1\ldots 2N}$ consisting of N pairs each i.e take $(\xi^i, w^i)_{i=1\ldots N}$ and $(\xi^i, W^i)_{i=N+1\ldots 2N}$. Define

$$Y_t^{i,(1)} = \xi_i + \int_0^t b\left(Y_r^{i,(1)}, \mu_r^{N,(1)}\right) dr + \int_0^t \sigma\left(Y_r^{i,(1)}, \mu_r^{N,(1)}\right) dw_r^i, \quad 1 \le i \le N,$$
$$Y_t^{i,(2)} = \xi_i + \int_0^t b\left(Y_r^{i,(2)}, \mu_r^{N,(2)}\right) dr + \int_0^t \sigma\left(Y_r^{i,(2)}, \mu_r^{N,(2)}\right) dw_r^i, \quad N+1 \le i \le 2N,$$

Study

$$\mathbb{V}ar[\Phi(\mu_t^{2N}) - \frac{1}{2}(\Phi(\mu_t^{N,(1)}) + \Phi(\mu_t^{N,(2)}))].$$

Antithetic trick

Take $2N$ independent samples from μ_0 and $2N$ independent Brownian motions $(\xi^i, W^i)_{i=1\dots 2N}$. Next consider a system of particles

$$Y_t^i = \xi_i + \int_0^t b\left(Y_r^i, \mu_r^{2N}\right) dr + \int_0^t \sigma\left(Y_r^i, \mu_r^{2N}\right) dw_r^i, \quad 1 \le i \le 2N.$$

Next take two non-intersecting subsets of $(\xi^i, w^i)_{i=1\dots 2N}$ consisting of N pairs each i.e take $(\xi^i, w^i)_{i=1\dots N}$ and $(\xi^i, W^i)_{i=N+1\dots 2N}$. Define

$$Y_t^{i,(1)} = \xi_i + \int_0^t b\left(Y_r^{i,(1)}, \mu_r^{N,(1)}\right) dr + \int_0^t \sigma\left(Y_r^{i,(1)}, \mu_r^{N,(1)}\right) dw_r^i, \quad 1 \le i \le N,$$

$$Y_t^{i,(2)} = \xi_i + \int_0^t b\left(Y_r^{i,(2)}, \mu_r^{N,(2)}\right) dr + \int_0^t \sigma\left(Y_r^{i,(2)}, \mu_r^{N,(2)}\right) dw_r^i, \quad N+1 \le i \le 2N,$$

Study

$$\mathbb{V}ar[\Phi(\mu_t^{2N}) - \frac{1}{2}(\Phi(\mu_t^{N,(1)}) + \Phi(\mu_t^{N,(2)}))].$$

Antithetic trick

Recall representation

$$
\begin{aligned}
\Phi(\mu_T^N) &= \Phi(\mathscr{L}(X_T^{0,\xi})) + \big(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mathscr{L}(\xi))\big) \\
&+ \int_0^T \frac{1}{2}\left[\frac{1}{N^2}\sum_{i=1}^N \mathrm{tr}\Big(a(Y_s^{i,N},\mu_s^N)\partial_\mu^2\mathcal{V}(s,\mu_s^N)(Y_s^{i,N},Y_s^{i,N})\Big)\right] ds \\
&+ \int_0^T \frac{1}{N}\sum_{i=1}^N \partial_\mu\mathcal{V}(s,\mu_s^N)(Y_s^{i,N})\sigma(Y_s^{i,N},\mu_s^N)dW_s^i .
\end{aligned}
$$

Hence

$$
\Phi(\mu_t^{2N}) - \frac{1}{2}(\Phi(\mu_t^{N,(1)}) + \Phi(\mu_t^{N,(2)})) = \Phi(\mu_0^{2N}) - \frac{1}{2}(\Phi(\mu_0^{N,(1)}) + \Phi(\mu_0^{N,(2)})) + \mathscr{D} + \mathscr{S},
$$

Easy to show

$$
\mathbb{E}[\mathscr{D}] \lesssim 1/N \quad \text{and} \quad \mathbb{E}[\mathscr{D}^2] \lesssim 1/N^2.
$$

Antithetic trick

Recall representation

$$\Phi(\mu_T^N) = \Phi(\mathscr{L}(X_T^{0,\xi})) + \big(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mathscr{L}(\xi))\big)$$
$$+ \int_0^T \frac{1}{2}\left[\frac{1}{N^2}\sum_{i=1}^N \mathrm{tr}\Big(a\big(Y_s^{i,N},\mu_s^N\big)\partial_\mu^2\mathcal{V}\big(s,\mu_s^N\big)\big(Y_s^{i,N},Y_s^{i,N}\big)\Big)\right] ds$$
$$+ \int_0^T \frac{1}{N}\sum_{i=1}^N \partial_\mu\mathcal{V}\big(s,\mu_s^N\big)\big(Y_s^{i,N}\big)\sigma\big(Y_s^{i,N},\mu_s^N\big)dW_s^i.$$

Hence

$$\Phi(\mu_t^{2N}) - \frac{1}{2}(\Phi(\mu_t^{N,(1)}) + \Phi(\mu_t^{N,(2)})) = \Phi(\mu_0^{2N}) - \frac{1}{2}(\Phi(\mu_0^{N,(1)}) + \Phi(\mu_0^{N,(2)})) + \mathscr{D} + \mathscr{S},$$

Easy to show

$$\mathbb{E}[\mathscr{D}] \lesssim 1/N \quad \text{and} \quad \mathbb{E}[\mathscr{D}^2] \lesssim 1/N^2.$$

Antithetic trick

Recall representation

$$\Phi(\mu_T^N) = \Phi(\mathscr{L}(X_T^{0,\xi})) + \big(\mathcal{V}(0,\mu_0^N) - \mathcal{V}(0,\mathscr{L}(\xi))\big)$$
$$+ \int_0^T \frac{1}{2}\left[\frac{1}{N^2}\sum_{i=1}^N \operatorname{tr}\Big(a(Y_s^{i,N},\mu_s^N)\partial_\mu^2\mathcal{V}(s,\mu_s^N)(Y_s^{i,N},Y_s^{i,N})\Big)\right] ds$$
$$+ \int_0^T \frac{1}{N}\sum_{i=1}^N \partial_\mu \mathcal{V}(s,\mu_s^N)(Y_s^{i,N})\sigma(Y_s^{i,N},\mu_s^N) dW_s^i.$$

Hence

$$\Phi(\mu_t^{2N}) - \frac{1}{2}(\Phi(\mu_t^{N,(1)}) + \Phi(\mu_t^{N,(2)})) = \Phi(\mu_0^{2N}) - \frac{1}{2}(\Phi(\mu_0^{N,(1)}) + \Phi(\mu_0^{N,(2)})) + \mathscr{D} + \mathscr{S},$$

Easy to show

$$\mathbb{E}[\mathscr{D}] \lessapprox 1/N \quad \text{and} \quad \mathbb{E}[\mathscr{D}^2] \lessapprox 1/N^2.$$

Antithetic trick

$$\mathscr{S} := \int_0^T \frac{1}{2N} \sum_{i=1}^{2N} \partial_\mu \mathcal{V}(s, \mu_s^{2N})(Y_s^i)\sigma(Y_s^i, \mu_s^{2N}) dw_s^i$$
$$- \frac{1}{2N}\Big(\sum_{i=1}^{N} \partial_\mu \mathcal{V}(\mu_s^{N,(1)})(Y_s^{i,(1)})\sigma(Y_s^{i,(1)}, \mu_s^{N,(1)}) dw_s^i + \sum_{i=N+1}^{2N} \partial_\mu \mathcal{V}(s, \mu_s^{N,(2)})(Y_s^{i,(2)})\sigma(Y_s^{i,(2)}, \mu_s^{N,(2)}) dw_s^i \Big).$$

$$\mathbb{E}[\mathscr{S}^2] = \mathbb{E}[\Big(\int_0^T \frac{1}{2N} \sum_{i=1}^{N} \Sigma(s, Y_s^i, \mu_s^{2N}) - \Sigma(s, Y_s^{i,(1)}, \mu_s^{N,(1)}) dw_s^i \Big)^2]$$
$$+ \mathbb{E}[\Big(\int_0^T \frac{1}{2N} \sum_{i=N+1}^{2N} \Sigma(s, Y_s^i, \mu_s^{2N}) - \Sigma(s, Y_s^{i,(2)}, \mu_s^{N,(2)}) dw_s^i \Big)^2].$$

$$\mathbb{E}[\Big(\frac{1}{2N} \sum_{i=1}^{N} \int_0^T \Sigma(s, Y_s^i, \mu_s^{2N}) - \Sigma(s, Y_s^{i,(1)}, \mu_s^{N,(1)}) dW_s^i \Big)^2]$$
$$= \frac{1}{4N^2} \sum_{i=1}^{N} \int_0^T \mathbb{E}[(\Sigma(s, Y_s^i, \mu_s^{2N}) - \Sigma(s, Y_s^{i,(1)}, \mu_s^{N,(1)})^2] ds]$$

Hence

$$\mathbb{E}(\mathscr{S}^2) \lesssim 1/N^2$$

Antithetic trick

$$\mathscr{S} := \int_0^T \frac{1}{2N}\sum_{i=1}^{2N} \partial_\mu \mathcal{V}(s,\mu_s^{2N})(Y_s^i)\sigma(Y_s^i,\mu_s^{2N})dw_s^i$$
$$-\frac{1}{2N}\Big(\sum_{i=1}^{N} \partial_\mu \mathcal{V}(\mu_s^{N,(1)})(Y_s^{i,(1)})\sigma(Y_s^{i,(1)},\mu_s^{N,(1)})dw_s^i + \sum_{i=N+1}^{2N} \partial_\mu \mathcal{V}(s,\mu_s^{N,(2)})(Y_s^{i,(2)})\sigma(Y_s^{i,(2)},\mu_s^{N,(2)})dw_s^i\Big).$$

$$\mathbb{E}[\mathscr{S}^2] = \mathbb{E}[\Big(\int_0^T \frac{1}{2N}\sum_{i=1}^{N}\Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(1)},\mu_s^{N,(1)})dw_s^i\Big)^2]$$
$$+\,\mathbb{E}[\Big(\int_0^T \frac{1}{2N}\sum_{i=N+1}^{2N}\Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(2)},\mu_s^{N,(2)})dw_s^i\Big)^2].$$

$$\mathbb{E}[\Big(\frac{1}{2N}\sum_{i=1}^{N}\int_0^T \Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(1)},\mu_s^{N,(1)})dW_s^i\Big)^2]$$
$$= \frac{1}{4N^2}\sum_{i=1}^{N}\int_0^T \mathbb{E}[(\Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(1)},\mu_s^{N,(1)})^2]ds]$$

Hence

$$\mathbb{E}(\mathscr{S}^2) \lesssim 1/N^2$$

Antithetic trick

$$\mathscr{S} := \int_0^T \frac{1}{2N}\sum_{i=1}^{2N} \partial_\mu \mathcal{V}(s,\mu_s^{2N})(Y_s^i)\sigma(Y_s^i,\mu_s^{2N})dw_s^i$$

$$-\frac{1}{2N}\Big(\sum_{i=1}^{N} \partial_\mu \mathcal{V}(\mu_s^{N,(1)})(Y_s^{i,(1)})\sigma(Y_s^{i,(1)},\mu_s^{N,(1)})dw_s^i + \sum_{i=N+1}^{2N} \partial_\mu \mathcal{V}(s,\mu_s^{N,(2)})(Y_s^{i,(2)})\sigma(Y_s^{i,(2)},\mu_s^{N,(2)})dw_s^i\Big).$$

$$\mathbb{E}[\mathscr{S}^2] = \mathbb{E}[\Big(\int_0^T \frac{1}{2N}\sum_{i=1}^{N} \Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(1)},\mu_s^{N,(1)})dw_s^i\Big)^2]$$

$$+\,\mathbb{E}[\Big(\int_0^T \frac{1}{2N}\sum_{i=N+1}^{2N} \Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(2)},\mu_s^{N,(2)})dw_s^i\Big)^2].$$

$$\mathbb{E}[\Big(\frac{1}{2N}\sum_{i=1}^{N}\int_0^T \Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(1)},\mu_s^{N,(1)})dW_s^i\Big)^2]$$

$$= \frac{1}{4N^2}\sum_{i=1}^{N}\int_0^T \mathbb{E}[(\Sigma(s,Y_s^i,\mu_s^{2N}) - \Sigma(s,Y_s^{i,(1)},\mu_s^{N,(1)})^2]ds]$$

Hence

$$\mathbb{E}(\mathscr{S}^2) \lesssim 1/N^2$$

Return to ICMS front page

...us and related topics

1st Moscow-UK workshop on stochastic analysis

ICMS, Bayes Centre, 47 Potterrow, Edinburgh EH8 9BT

19 - 23 November 2018

Organisers

Mireille Bossy, INRIA, Sophia Antipolis
Valentin Konakov, HSE, Moscow
David Siska, University of Edinburgh
Lukasz Szpruch, University of Edinburgh

This workshop will review and consolidate knowledge on the emerging field of stochastic calculus on Wasserstein space and asses its current impact on several branches of mathematics. The workshop will bring together:

- key developers of calculus on Wasserstein spaces
- researchers on mean-field games
- mathematicians working on kinetic theory and optimal transport
- applied mathematicians working on a mean-field approach to power grid applications and crowd behaviour.

A key benefit of the workshop will be to strengthen the collaboration between UK-based researchers and the Laboratory of Stochastic Analysis and its Applications, at the Higher School of Economics (HSE) in Moscow, led by Valentin Konakov. The consolidation of knowledge and exchange of ideas in the emerging field of Wasserstein calculus with its links to optimal transport, nonlinear PDEs, particle representations and mean-field games. The workshop will also develop research strategies emphasising applications in energy markets, crowd dynamics and cybersecurity.

Invited speakers

Pierre Cardaliaguet Université Paris-Dauphine
Dan Crisan, Imperial College London
Paul Eric Chaudru de Raynal Université Savoie Mont Blanc
Franco Fandoli Scuola Normale Superiore di Pisa
Alexander Gushchin, HSE Moscow
Elena Issoglio, Univeristy of Leeds
Jean-Francois Jabir, HSE, Moscow
Benjamin Jourdain, CERMICS
Mark Kelbert, HSE Moscow
Anna Kozhina, HSE Moscow
Pierre-Louise Lions, College de France
Mario Maurelli, University of York
Stanislav Molchanov, HSE Moscow
Ekaterina Palamarchuk, HSE Moscow
Huyen Pham, Université Paris Diderot
Denis Talay, INRIA, Sophia Antipolis
Alexander Veretennikov, Univeristy of Leeds

Bibliography

Ambrosio, L., Gigli, N., and Savaré, G. (2008).
Gradient flows: in metric spaces and in the space of probability measures.
Springer Science & Business Media.

Antonelli, F., Kohatsu-Higa, A., et al. (2002).
Rate of convergence of a particle method to the solution of the McKean–Vlasov equation.
The Annals of Applied Probability, 12(2):423–476.

Bossy, M., Talay, D., et al. (1996).
Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation.
The Annals of Applied Probability, 6(3):818–861.

Buckdahn, R., Li, J., Peng, S., and Rainer, C. (2017).
Mean-field stochastic differential equations and associated PDEs.
The Annals of Probability, 45(2):824–878.

Cardaliaguet, P. (2013).
Notes on mean field games (from P.-L. Lions' lectures at Collège de France).
Online at `https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf`.

Chassagneux, J.-F., Crisan, D., and Delarue, F. (2014).
Classical solutions to the master equation for large population equilibria.
arXiv:1411.3009.

Del Moral, P., Patras, F., and Rubenthaler, S. (2011).
Convergence of u-statistics for interacting particle systems.
Journal of Theoretical Probability, 24(4):1002.