OXFORD SPA 2015

François Delarue
University of Nice (France)

38th Conference on Stochastic Processes and their Applications
Spa2015@oxford-man.ox.ac.uk

Forward-Backward SDEs of McKean-Vlasov type and PDEs on $\mathcal{P}_2(\mathbb{R}^d)$

Oxford - SPA Meeting

François Delarue (Nice – J.-A. Dieudonné)

July 14th, 2015

Joint work with J.-F. Chassagneux and D. Crisan

Basic purpose

- Forward-backward system of McKean-Vlasov type

$$dX_t = b(X_t, Y_t, Z_t, \mathcal{L}(X_t, Y_t))dt + \sigma(X_t, Y_t, \mathcal{L}(X_t, Y_t))dW_t$$
$$dY_t = -f(X_t, Y_t, Z_t, \mathcal{L}(X_t, Y_t))dt + Z_t dW_t, \quad Y_T = g(X_T, \mathcal{L}(X_T))$$

  - on $(\Omega, (\mathcal{F}_t)_t, \mathbb{P})$, $X_t, W_t : \Omega \to \mathbb{R}^d$, $Y_t : \Omega \to \mathbb{R}^m$, $Z_t : \Omega \to \mathbb{R}^{m\times d}$
  - example 1: $b(x,\mu) = b(x, \int_{\mathbb{R}} \varphi d\mu)$, $\varphi = \mathrm{Id}$, $\varphi \approx \delta_x$...
  - example 2: $b(x,\mu) = \int_{\mathbb{R}} b(x-v)d\mu(v) \quad b(x,v) = |x-v|$...
- Main motivation $\rightsquigarrow$ probabilistic interpretation of some stochastic control problem set over large population
  - mean-field games $\rightsquigarrow$ Nash for games with ∞ players
  - optimization over McKean-Vlasov diffusion processes
- General questions
  - solvability $\rightsquigarrow$ difficulty due to the forward-backward structure
  - notion of decoupling field $\rightsquigarrow$ associated PDE? $(m = 1)$

Nash equilibrium within infinite population

- Mean-field games Lasry-Lions, Huang-Caines-Malhamé (2006) …
- Nash equilibrium for an ∞ population with mean-field interaction

(1) one representative player

$$dX_t = \alpha_t dt + dW_t,$$

◦ with cost $J(\alpha) = \mathbb{E}\Big[g(X_T, \mu_T) + \int_0^T \Big(f(X_t, \mu_t) + \frac{1}{2}|\alpha_t|^2\Big)dt\Big]$

◦ in an environment formed by a flow of probability measures $(\mu_t)_{0 \le t \le T}$ (in $\mathcal{P}_2(\mathbb{R}^d)$) describing the state of the population

(2) let $(X_t^{\star,\mu})_{0 \le t \le T}$ be the unique optimizer (under nice assumptions)

$$dX_t^{\star,\mu} = -Z_t^{\star,\mu} dt + dW_t$$
$$dY_t^{\star,\mu} = \Big(f(X_t^{\star,\mu}, \mu_t) + \frac{1}{2}|Z_t^{\star,\mu}|^2\Big)dt + Z_t^{\star,\mu} dW_t, \quad Y_T^{\star,\mu} = g(X_T^{\star,\mu}, \mu_T)$$

$$(\mu_t)_{0 \le t \le T} \text{ is Nash if } \quad \mu_t = \mathcal{L}(X_t^{\star,\mu}), \quad t \in [0, T]$$

- FBSDE of MKV type!
- Tomorrow's talk (P. Tarrès session): justification...

1. Small Time Analysis

Implementing Picard theorem

- General difficulty

  ◦ no Cauchy-Lipschitz theory for forward-backward systems in arbitrary time $\rightsquigarrow$ well-known counter examples (with or without noise and with or without McKean-Vlasov)

  ◦ Cauchy-Lipschitz theory in small time only!

- What does Lipschitz mean? Need distance on $\mathcal{P}_2(\mathbb{R}^\ell)$ (probability measures with a second order moment)

$$\mu, \eta \in \mathcal{P}_2(\mathbb{R}^\ell), \quad W_2(\mu,\eta) = \left(\inf_\pi \int_{\mathbb{R}^\ell \times \mathbb{R}^\ell} |x-y|^2 d\pi(x,y) \right)^{1/2},$$

where π has μ and η as marginals on $\mathbb{R}^\ell \times \mathbb{R}^\ell$

  ◦ X and X' two r.v.'s $\Rightarrow W_2(\mathcal{L}(X), \mathcal{L}(X')) \le \mathbb{E}[|X - X'|^2]^{1/2}$

- Theorem: If K-Lipschitz coefficients $\Rightarrow \exists!$ for $T \le c(K)$

  ◦ for any initial condition $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$

  ◦ strong ! $\Rightarrow$ weak !

Decoupling field ($T \leq c(K)$)

- Recall non MKV case $\rightsquigarrow \exists U : [0,T] \times \mathbb{R}^d \to \mathbb{R}^m$ such that

$$Y_t = U(t, X_t) \quad \Leftrightarrow \quad U(t,x) = Y_t^{t,x} \text{ (with } X_t^{t,x} = x)$$

  - reminiscent of Markov property of $(X_t)_{t \in [0,T]}$ on $\mathbb{R}^d$

- MKV setting $\rightsquigarrow$ Markov property must be on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$
- Construction for $t \in [0,T]$, $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$
  - 1st MKV FBSDE with $X_t \sim \mu$

$$dX_s = b(X_s, Y_s, Z_s, \mathcal{L}(X_s, Y_s))ds + \sigma(X_s, Y_s, \mathcal{L}(X_s, Y_s))dW_s$$
$$dY_s = -f(X_s, Y_s, Z_s, \mathcal{L}(X_s, Y_s))ds + Z_s dW_s, \quad Y_T = g(X_T, \mathcal{L}(X_T))$$

  - 2nd non-MKV FBSDE with $\tilde{X}_t = x$ and input

$$d\tilde{X}_s = b(\tilde{X}_s, \tilde{Y}_s, \tilde{Z}_s, \mathcal{L}(X_s, Y_s))ds + \sigma(\tilde{X}_s, \tilde{Y}_s, \mathcal{L}(X_s, Y_s))dW_s$$
$$d\tilde{Y}_s = -f(\tilde{X}_s, \tilde{Y}_s, \tilde{Z}_s, \mathcal{L}(X_s, Y_s))ds + \tilde{Z}_s dW_s, \quad \tilde{Y}_T = g(\tilde{X}_T, \mathcal{L}(X_T))$$

  - let $U(t,x,\mu) = \tilde{Y}_t \Rightarrow Y_t = U(t, X_t, \mu) = U(t, X_t, \mathcal{L}(X_t))$

Road for a PDE

- Question 1: smoothness of U?
- Question 2: dynamics of U? (non MKV case $\rightsquigarrow$ PDE)

$$dX_t = b(X_t, Y_t, Z_t, \mathcal{L}(X_t, Y_t))dt + \sigma(X_t, Y_t, \mathcal{L}(X_t, Y_t))dW_t$$

$$dY_t = -f(X_t, Y_t, Z_t, \mathcal{L}(X_t, Y_t))dt + Z_t dW_t, \quad Y_T = g(X_T, \mathcal{L}(X_T))$$

$\circ$ $Y_t = U(t, X_t, \mathcal{L}(X_t)) \rightsquigarrow$ need chain rule to expand the right-hand side and compare the dt-terms

$$d_t Y_t = [d_t \underbrace{U(t, X_t, \mathcal{L}(X_s))]_{|s=t}}_{\text{standard Itô}} + \underbrace{d_t[U(s, X_s, \mathcal{L}(X_s))]_{|s=t}}_{\text{chain rule?}}$$

$\circ$ global PDE contains classical PDE

$$\big(\partial_t U + L_U U\big)(t, x, \mu) + f\Big(x, U(t, x, \mu), \nabla_x U(t, x, \mu)\sigma(x, U(t, x, \mu), \eta), \eta\Big)$$

$$+ \text{ op. diff. in } \mu \text{ acting on } U = 0$$

$\circ$ $\eta = \mu \circ (Id, U(t, \cdot, \mu))^{-1}$

$$L_U = b\Big(x, U(t, x, \mu), \nabla_x U(t, x, \mu)\sigma(x, U(t, x, \mu), \eta), \eta\Big) \cdot \nabla$$

$$+ \tfrac{1}{2}\sigma\sigma^\top\Big(x, U(t, x, \mu), \eta\Big) \cdot \nabla^2$$

2. Implementing differential calculus on $\mathcal{P}_2({}^d)$

Derivative on the Wasserstein space

- Differentiation on $\mathcal{P}_2(\mathbb{R}^d)$ taken from Lions
- Consider $\mathcal{U} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$
- Lifted-version of $\mathcal{U}$

$$\hat{\mathcal{U}} : L^2(\Omega, \mathbb{P}) \ni X \mapsto \mathcal{U}(\mathrm{Law}(X))$$

  - $\mathcal{U}$ differentiable if $\hat{\mathcal{U}}$ Fréchet differentiable

- Differential of $\mathcal{U}$
  - Fréchet derivative of $\hat{\mathcal{U}}$

$$D\hat{\mathcal{U}}(X) = \partial_\mu \mathcal{U}(\mu)(X), \quad \partial_\mu \mathcal{U}(\mu) : \mathbb{R}^d \ni v \mapsto \partial_\mu \mathcal{U}(\mu)(v) \quad \mu = \mathcal{L}(X)$$

  - derivative of $\mathcal{U}$ at $\mu \rightsquigarrow \partial_\mu \mathcal{U}(\mu) \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$

- Finite dimensional projection

$$\partial_{x_i}\Big[\mathcal{U}\Big(\frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)\Big] = \frac{1}{N}\partial_\mu \mathcal{U}\Big(\frac{1}{N}\sum_{j=1}^N \delta_{x_j}\Big)(x_i), \quad x_1, \ldots, x_N \in \mathbb{R}^d$$

Chain rule on $\mathcal{P}_2(\mathbb{R}^d)$

- Itô process $dX_t = b_t dt + \sigma_t dW_t, \quad \int_0^T \mathbb{E}[|b_t|^2 + |\sigma_t|^4]dt < \infty$
  - μ_t = law of X_t
- $\hat{\mathcal{U}}$ twice Fréchet differentiable
  - chain rule for $(\mathcal{U}(\mu_t))_{t \geq 0}$?
- Approximate μ_t by particle system

$$\mu_t \sim \frac{1}{N}\sum_{j=1}^N \delta_{X_t^j} \quad \text{and} \quad d_t\Big[\mathcal{U}\Big(\frac{1}{N}\sum_{j=1}^N \delta_{X_t^j}\Big)\Big]$$

  - expand the right-hand side and pass to the limit
- Chain rule
  - need $\mathbb{R}^d \ni v \mapsto \partial_\mu \mathcal{U}(\mu)(v) \in \mathbb{R}^d$ differentiable

$$\frac{d}{dt}\mathcal{U}(\mu_t) = \mathbb{E}[\langle b_t, \partial_\mu \mathcal{U}(\mu_t)(X_t)\rangle] + \frac{1}{2}\mathbb{E}[\mathrm{Trace}(\sigma_t \sigma_t^\top \partial_v(\partial_\mu \mathcal{U}(\mu_t))(X_t))]$$

Smoothness of the decoupling field

- Assume that coefficients are C^2 in suitable sense

  ∘ example for $g : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \rightsquigarrow \partial_x g$, $\partial^2_{xx} g$, $\partial_\mu g$, $\partial_x \partial_\mu g$ and $\partial_v \partial_\mu g$ exist + suitable Lipschitz type conditions

  ∘ claim that $U(t, \cdot)$ satisfy (almost) the same type of regularity

- Identify the differential operator in the direction of the measure in

$$\Big(\partial_t U + L_U U\Big)(t, x, \mu) + f\Big(x, U(t, x, \mu), \nabla_x U(t, x, \mu)\sigma(x, U(t, x, \mu), \eta), \eta\Big)$$
$$+ \text{ op. diff. in } \mu \text{ acting on } U = 0$$

  ∘ chain rule yields

$$\int_{\mathbb{R}^d} b\Big(v, U(t, v), \nabla_x U(t, v)\sigma(t, v, U(t, v, \mu), \eta)\Big) \cdot \underbrace{\partial_\mu U(t, x, \mu)(v)}_{\to \mathbb{R}^d} d\mu(v)$$
$$+ \frac{1}{2} \int_{\mathbb{R}^d} (\sigma\sigma^\top)(v, U(t, v, \mu), \eta) \cdot \underbrace{\partial_v \partial_\mu U(t, x, \mu)(v)}_{\to \mathbb{R}^{d\times d}} d\mu(v)$$

- Theorem: $T \leq c(K) \Rightarrow U$ classical solution

3. Getting the smoothness

Implementing a flow method

- Strategy $\rightsquigarrow$ difficult part in direction of the measure only
  - investigate the flow $L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d) \ni \xi \mapsto (X_s^{t,\xi}, Y_s^{t,\xi}, Z_s^{t,\xi})_{t \le s \le T}$

$$dX_s = b(X_s, Y_s, Z_s, \mathcal{L}(X_s, Y_s))ds + \sigma(X_s, Y_s, \mathcal{L}(X_s, Y_s))dW_s, \quad X_t = \xi$$

$$dY_s = -f(X_s, Y_s, Z_s, \mathcal{L}(X_s, Y_s))ds + Z_s dW_s, \quad Y_T = g(X_T, \mathcal{L}(X_T))$$

  - investigate linearization in L^2! for $\xi,\ \chi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$

$$(\partial_\chi X_s^{t,\xi}, \partial_\chi Y_s^{t,\xi}, \partial_\chi Z_s^{t,\xi}) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big(\underbrace{X_s^{t,\xi+\varepsilon\chi} - X_s^{t,\xi}, Y_s^{t,\xi+\varepsilon\chi} - Y_s^{t,\xi}}_{\text{in } \mathbb{E}[\sup_{t \le s \le T} |\cdot_s|^2]}, \ \underbrace{Z_s^{t,\xi+\varepsilon\chi} - Z_s^{t,\xi}}_{\text{in } \mathbb{E}\int_t^T |\cdot_s|^2 ds} \Big)$$

- Get formal representation

$$\partial_\chi Y_t^{t,\xi} = \partial_x U(t, \xi, \mathcal{L}(\xi)) \cdot \chi + \underbrace{\hat{\mathbb{E}}[\partial_\mu U(t, \xi, \mathcal{L}(\xi))(\hat{\xi}) \cdot \hat{\chi}]}_{\hat{\Omega} = \text{copy space}}$$

  - access $\partial_\mu U$ through second term!

Linearization in the uncoupled case

- Forget forward-backward and consider the decoupled case

$$dX_t = b(X_t, \mathcal{L}(X_t))dt + dW_t, \quad X_0 = \xi$$

  - choose $\sigma = \mathrm{Id}$ for simplicity

- Analogue with above $\rightsquigarrow$ choose $f \equiv 0$ and $g(x,\mu) = g(\mu)$

$$U(0,\mu) = \mathbb{E}[g(\mathcal{L}(X_T^{0,\xi}))], \quad \xi \sim \mu$$

- Perturbation of X_0 in direction $\chi \in L^2$
  - $X_0^\varepsilon = \xi + \varepsilon\chi \rightsquigarrow (X_t^\varepsilon)_{0\le t\le T} \Rightarrow \partial_\chi X_t = \frac{dX_t^\varepsilon}{d\varepsilon}|_{\varepsilon=0}$

- Dynamics of $\partial_\chi X \rightsquigarrow$ new MKV equation

$$\begin{aligned} d\partial_\chi X_t &= \partial_x b(X_t, \mathcal{L}(X_t)) \cdot \partial_\chi X_t dt \\ &\quad + \hat{\mathbb{E}}[\partial_\mu b(X_t, \mathcal{L}(X_t))(\hat{X}_t) \cdot \partial_\chi \hat{X}_t]dt, \quad \partial_\chi X_0 = \chi \end{aligned}$$

- Derivative of $U(0,\cdot)$ reads

$$\hat{\mathbb{E}}\Big[\partial_\mu U(0, \mathcal{L}(\xi))(\hat{\xi}) \cdot \hat{\chi}\Big] = \hat{\mathbb{E}}\Big[\partial_\mu g(\mathcal{L}(X_T))(\hat{X}_T) \cdot \partial_\chi \hat{X}_T\Big]$$

Estimating the solution

- Real challenge is quantitative estimate! Key fact for FBSDEs is to control Lipschitz constant of $U(t,\cdot)$ to iterate the small time result

- Theorem: If smooth coefficients and *a priori* bound for Lip U in (x,μ) on $[0,T] \Rightarrow$ classical solution on $[0,T]$

- In linear regime

$$\mathbb{E}\Big[\big|\partial_\mu U(0,\mathcal{L}(X_0))(X_0)\big|^2\Big]^{1/2}$$
$$\leq \mathbb{E}\Big[\big|\partial_\mu g(\mathcal{L}(X_T))(X_T)\big|^2\Big]^{1/2} \sup_{\chi:\mathbb{E}[|\chi|^2]\leq 1} \mathbb{E}\Big[\big|\partial_\chi X_t\big|^2\Big]^{1/2}$$

- L^2 estimate of $\mathbb{E}[|\partial_\chi X_t|^2]$

$$d\mathbb{E}[|\partial_\chi X_t|^2] = 2\mathbb{E}[\langle \partial_\chi X_t, \partial_x b(X_t,\mathcal{L}(X_t))\partial_\chi X_t\rangle]dt$$
$$+ \mathbb{E}\hat{\mathbb{E}}[\langle \partial_\chi X_t, \partial_\mu b(X_t,\mathcal{L}(X_t))(\hat{X}_t)\partial_\chi \hat{X}_t\rangle]dt$$

  - ∘ deduce $\mathbb{E}[|\partial_\chi X_t|^2] \leq C\mathbb{E}[|\chi|^2]$ with

$$C = C\Big(T, \sup_{x,\mu}\big|\partial_x b(x,\mu)\big|^2, \sup_{x,\mu}\int |\partial_\mu b(x,\mu)(v)|^2 d\mu(v)\Big)$$

Example in coupled case ($m = 1$)

- $b \equiv -z, \quad f \equiv f(x,\mu) + \frac{1}{2}|z|^2, \quad g, f$ bded, smooth and monotonous

$$\int_{\mathbb{R}^d} [g(x,\mu) - g(x,\mu')] d(\mu - \mu')(x) \geq 0$$

$$\mathbb{E}\hat{\mathbb{E}}[\partial_x(\partial_\mu g)(X, \mathcal{L}(X))(\hat{X}) \cdot \hat{\chi} \otimes \chi] \geq 0$$

- Dynamics of $(X_t)_t$ and $(\partial_\chi X_t)_t$

$$\begin{aligned} dX_t &= -\partial_x U(t, X_t, \mathcal{L}(X_t))dt + dW_t \\ d\partial_\chi X_t &= -\partial^2_{xx} U(t, X_t, \mathcal{L}(X_t))\partial_\chi X_t dt \\ &\quad - \hat{\mathbb{E}}[\partial_\mu(\partial_x U)(t, X_t, \mathcal{L}(X_t))(\hat{X}_t)\partial_\chi \hat{X}_t]dt \end{aligned}$$

  - $\partial^2_{xx} U$ already estimated! (thanks to Laplace)

- Propagation of monotonicity

$\mathbb{E}\hat{\mathbb{E}}[\partial_x(\partial_\mu \mathcal{U})(t, X_t, \mathcal{L}(X_t))(\hat{X}_t)\partial_\chi \hat{X}_t \partial_\chi X_t] \geq 0 \rightsquigarrow \mathbb{E}[|\partial_\chi X_t|^2] \leq C\mathbb{E}[|\zeta|^2]$

  - gives a way to control derivative in $\mu \rightsquigarrow$ avoid any blow-up

- Theorem: Under monotonicity like conditions and smoothness of coefficients, there exists a classical solution over any interval.

Higher-order derivatives

- Master equation $\rightsquigarrow$ differentiate once again w.r.t. v

$$(\mu, v) \mapsto \partial_\mu P_t\phi(\mu)(v)$$

- Derivatives in the direction v/X_0
  - freeze ζ and consider new perturbation $X_0 \rightsquigarrow X_0^\varepsilon$

$$\mathcal{L}(X_0^\varepsilon) \quad \text{independent of} \quad \varepsilon \Rightarrow \mathcal{L}(X_t^{\star,\varepsilon}) \quad \text{independent of} \quad \varepsilon$$

- differentiate the formula for the derivative

$$\mathbb{E}\Big[\Big\langle \partial_v \partial_\mu \Big(P_t\phi(\mathcal{L}(X_0^0))\Big)(X_0^0), \zeta \otimes \frac{dX_0^\varepsilon}{d\varepsilon}_{|\varepsilon=0}\Big\rangle\Big]$$

$$= \mathbb{E}\Big[\Big\langle \partial_v \partial_\mu \phi(\mathcal{L}(X_t^{0,\star}))(X_t^{0,\star}), \partial_\zeta X_t^{0,\star} \otimes \frac{d}{d\varepsilon}_{|\varepsilon=0} X_t^{\varepsilon,\star}\Big\rangle\Big]$$

$$+ \mathbb{E}\Big[\Big\langle \partial_\mu \phi(\mathcal{L}(X_t^{0,\star}))(X_t^{0,\star}), \frac{d}{d\varepsilon}_{|\varepsilon=0} \partial_\zeta X_t^{\varepsilon,\star}\Big\rangle\Big]$$

  - example $X_0^\varepsilon = X_0 + \delta\Big(\cos(\varepsilon)Z + \sin(\varepsilon)Z'\Big)$
  - with $(Z, Z') \sim \mathcal{N}(0,1)^{\otimes 2}$ and (Z, Z') independent of X_0

Checking the monotonicity condition

- Lasry-Lions monotonicity condition (choose $d = 1$)

$$\int_{\mathbb{R}} (h(x,\mu') - h(x,\mu)) d(\mu' - \mu)(x) \geq 0$$

  - $X \sim \mu$ and $X' \sim \mu'$

$$\mathbb{E}\Big[h(X', \mathcal{L}(X')) - h(X', \mathcal{L}(X)) - \big(h(X, \mathcal{L}(X')) - h(X, \mathcal{L}(X))\big)\Big] \geq 0$$

- Make a perturbation $X' = X + \varepsilon Y$
  - first step

$$\mathbb{E}\hat{\mathbb{E}}\Big[\partial_\mu h(X', \mathcal{L}(X))(\hat{X})\hat{Y} - \partial_\mu h(X, \mathcal{L}(X))(\hat{X})\hat{Y}\Big] + o(\varepsilon) \geq 0$$

  - need copies $\hat{X}$ and $\hat{Y}$ on another space
  - second step

$$\mathbb{E}\hat{\mathbb{E}}\Big[\partial_x \partial_\mu h(X, \mathcal{L}(X))(\hat{X})\hat{Y} Y\Big] \geq 0$$