Probabilistic Approach for Mean Field Games with a Common Noise

Workshop MFG
Università di Padova

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

September 2013, 5th

(Joint work with R. Carmona)

Common noise

- Modelling (in $\mathbb{R}$ to simplify)
  - Dynamics of player number $1 \leq i \leq N$

$$dX_t^i = b(X_t^i, \bar{\mu}_t^N, \alpha_t^i)dt + \sigma dW_t^i + \varsigma dB_t$$
$$X_0^i = x_0$$

  - independent noises $B, W^1, \ldots, W^N$
  - $\bar{\mu}_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i} \quad 0 \leq t \leq T$
  - α_t^i prog. meas. w.r.t. $\sigma(W^1, \ldots, W^N, B)$
- Nash equilibrium w.r.t.
  - $J^i = \mathbb{E}\left[g(X_T^i, \bar{\mu}_T^N) + \int_0^T f(X_t^i, \bar{\mu}_t^N, \alpha_t^i) dt \right]$

Conditional law of large numbers

- Exchangeable equilibria $\rightsquigarrow$ conditional LLN

$$\bar{\mu}_t^N \sim_{N\uparrow+\infty} \mathcal{L}(X_t^1|B) \quad (= \mathcal{L}(X_t^1|(B_s)_{0\le s\le t}))$$

- Scheme
  - Fix the flow of random measures $(\mu_t)_{0\le t\le T}$
  - μ_t depends on $(B_s)_{0\le s\le t}$ only!
  - Optimize

$$dX_t^1 = b(X_t^1, \mu_t, \alpha_t^1)dt + \sigma W_t + \varsigma dB_t$$

$$J^1 = \mathbb{E}\left[g(X_T^1, \mu_T) + \int_0^T f(X_t^1, \mu_t, \alpha_t^1)dt\right]$$

  - Solve the matching problem $\mu_t = \mathcal{L}(X_t|B)$

Strong vs. weak equilibra

- Strong sense
  - Probability space is given
  - Canonical space: $\underbrace{\mathcal{C}([0,T],\mathbb{R})}_{\text{for } B} \times \underbrace{\mathcal{C}([0,T],\mathbb{R})}_{\text{for } W}$
  - $(\mu_t)_{0\le t\le T}$ is prog. meas. w.r.t. $\sigma(B)$ (function of the 1st coordinate)
- Weak sense: probability space is not given
  - $\exists$ 2 filtered probability spaces $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i)$, $i = 1, 2$
  - $(B_t, \mu_t)_{0\le t\le T}$ is carried on Ω^1, $(W_t)_{0\le t\le T}$ on Ω^2
  - $\mu_t = \mathcal{L}(X_t|\mathcal{F}_t^1)$ (conditioning is enlarged but independent of W)
- Yamada-Watanabe: strong ! + weak $\exists \Rightarrow$ strong $\exists$
  - reconstruct solutions on the same space

Stochastic maximum principle

- extended Hamiltonian
  - $H(x, y, \mu, \alpha) = b(x, \mu, \alpha)y + f(x, \mu, \alpha)$
  - $\hat{\alpha}(x, y, \mu) = \text{argmin}_{\alpha} H(x, y, \mu, \alpha)$
- adjoint equations for a nonanticipative frozen $(\mu_t)_{0 \le t \le T}$

$$dX_t = b\big(X_t, \mu_t, \hat{\alpha}(X_t, Y_t, \mu_t)\big)dt + \sigma dW_t + \varsigma dB_t$$
$$dY_t = -\partial_x H\big(X_t, Y_t, \mu_t, \hat{\alpha}(X_t, Y_t, \mu_t)\big)dt + d \underbrace{\text{martingale}_t}_{\text{measurability constraint}}$$

$$Y_T = \partial_x g(X_T, \mu_T)$$

- Solve eq. with the constraint $\mu_t = \mathcal{L}(X_t|B)$: MKV FBSDE
  - H and g convex w.r.t. $(x, \alpha) \Rightarrow X$ equilibrium
  - numerics: discretization of the adjoint equations

Dynamics of X

- Decoupling random field $u : [0, T] \times \mathbb{R} \times \Omega \to \mathbb{R}$
  - $(u(t,\cdot))_{0\le t\le T}$ is $\sigma(B)$ prog. meas. and $Y_t = u(t, X_t)$
- Dynamics of X at equilibrium

$$dX_t = \underbrace{b\big(X_t, \mu_t, \hat{\alpha}(X_t, u(t, X_t), \mu_t)\big)}_{\hat{b}(t, X_t)} dt + \sigma dW_t + \varsigma dB_t$$

  - Convex Hamiltonian $\Rightarrow$ u Lipschitz in x
- Conditional path of X given $\mathcal{F}^1$ $\rightsquigarrow$ Freeze B
  - $\mathcal{L}(X_t|B = \beta) = \mathcal{L}(X_t^\beta)$

$$X_t^\beta = x_0 + \int_0^t \hat{b}\big(s, X_s^\beta\big) ds + \sigma W_t + \varsigma \beta_t$$

PDE point of view: stochastic HJB

- Common noise $\Rightarrow$ Value function = random field

$$U(t,x) = \inf_{\alpha, X_t = x} \mathbb{E}\Big[g(X_T,\mu_T) + \int_t^T L(X_s,\mu_s,\alpha_s)ds \Big| \mathcal{F}_t^B\Big]$$

- U adapted $\Rightarrow$ Backward Stochastic HJB

$$\begin{aligned} &d_t U(t,x) \\ &\quad + \Big(\underbrace{\mathcal{L}U(t,x)}_{\text{generator}} + \underbrace{\inf_\alpha\big[b(x,\mu_t,\alpha)\partial_x U(t,x) + L(x,\mu_t,\alpha)\big]}_{\text{standard Hamiltonian in HJB}} \\ &\qquad + \underbrace{\zeta\partial_x V(t,x)}_{\text{Ito Wentzell cross term}} \Big)dt - \underbrace{V(t,x)dB_t}_{\text{backward term}} = 0 \end{aligned}$$

- Maximum principle $\rightsquigarrow$ $Y_t = \partial_x U(t,X_t)$ i.e. $u = \partial_x U$

Stochastic Kolmogorov

- Dynamics of the conditional law of X given B
- Replace $(B_t)_{0\le t\le T}$ by a piecewise affine curve $(\beta_t)_{0\le t\le T}$
  - Kolmogorov equation

$$d_t\mu_t = -\mathrm{div}\big(b(x,\mu_t,\hat{\alpha}(x,\mu_t,u(t,x))\big)dt + \frac{\sigma^2}{2}\partial^2_{xx}\mu_t dt - \mathrm{div}\big(\mu_t\varsigma\big)\dot{\beta}_t dt$$

- Use $\beta = \beta^N =$ affine interpolation of B with N nodes

$$d_t\mu_t = -\mathrm{div}\big(b(x,\mu_t,\hat{\alpha}(x,\mu_t,u(t,x))\big)dt + \frac{\sigma^2}{2}\partial^2_{xx}\mu_t dt - \mathrm{div}\big(\mu_t\varsigma\big)\circ dB_t$$

- HJB + Kolmogorov $\Rightarrow$ forward backward SDE in ∞ dim

Lifted value function

- Representation of the value random function

$$U(t,x,\omega) = \mathcal{U}\big(t,x,\mu_t(\omega)\big),$$

  - $\mathcal{U} : [0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$
  - $\exists$ if strong uniqueness
- Write second-order (master?) PDE in infinite dimension
  - Derivatives on $\mathcal{P}_2(\mathbb{R}) \rightsquigarrow$ r.v. $\partial_\mu \mathcal{U}(t,x,\mu_t)(X_t)$
  - Connection with V in stochastic HJB equation

$$V(t,x,\omega) = \int \partial_\mu \mathcal{U}(t,x,\mu_t(\omega))(y)\varsigma\mu_t(\omega,dy)$$

- Used in parametric models $\mu_t = \mu(q_t)$

$$\mathcal{U}(t,x,\mu_t) \rightsquigarrow \mathcal{U}(t,x,q_t)$$

Weak solvability conditions

- Weakening notion of solutions $\Rightarrow$ same strategy
- Convexity of the Hamiltonian (very much demanding!)
  - $b(x, \mu, \alpha) = b_0(\mu) + b_1 x + b_2 \alpha$
  - f convex in (x, α) (and strictly convex in α)
- Local Lipschitz bound (linear growth) in (x, μ, α)
  - use Wasserstein W_2 for the measure
- Mean-reverting
  - $\langle x, \partial_x f(0, \delta_x, 0) \rangle, \langle x, \partial_x g(0, \delta_x) \rangle \geq -c(1 + |x|)$
  - example: bounded derivatives
- Smoothness ($\mathcal{C}^{1,1}$ in (x, α))...
- Strong ! is welcome

Strategy of proof for solvability

- Forget strong vs. weak! Freeze the conditional measure

$$
\begin{aligned}
dX_t &= b\big(X_t, \mu_t, \hat{\alpha}(X_t, Y_t, \mu_t)\big)dt + \sigma W_t + \varsigma dB_t \\
dY_t &= -\partial_x H\big(X_t, Y_t, \mu_t, \hat{\alpha}(X_t, Y_t, \mu_t)\big)dt + dmartingale_t \\
Y_T &= \partial_x g(X_T, \mu_T)
\end{aligned}
$$

- Find a fixed point $\Phi : (\mu_t)_{0 \le t \le T} \mapsto (\mathcal{L}(X_t^{\mu}|B))_{0 \le t \le T}$
- If no common noise
  - Fixed point in $\mathcal{C}\big([0,T], \underbrace{\mathcal{P}(\mathbb{R})}_{\text{set of prob. meas.}}\big)$
  - Use Schauder's th: Φ continuous and range of Φ compact
- If common noise
  - Fixed point in subset of $\big(\mathcal{C}([0,T], \mathcal{P}(\mathbb{R}))\big)^{\Omega}$
  - Compactness?

Discretization of the conditioning

- Discretization: $\mathcal{L}(X_t|B) \rightsquigarrow \mathcal{L}(X_t|\text{finitely supported process})$
  - Π projection mapping onto space grid $\{x_1, \dots, x_M\} \subset \mathbb{R}$
  - $t_1, \dots, t_N$ a finite time grid $\subset [0, T]$
  - $\hat{B}_{t_i} = \Pi(B_{t_i})$
- Forward-backward system with

$$\mathcal{L}(X_t|\hat{B}_{t_1}, \dots, \hat{B}_{t_i}), \quad t_i \le t < t_{i+1}$$

  - $(\hat{B}_{t_1}, \dots, \hat{B}_{t_N})$ has finite support of size MN
  - Fixed point in $\left(\mathcal{C}([0, T], \mathcal{P}(\mathbb{R}))\right)^{MN}$
  - $\exists \underbrace{\hat{X}^{M,N}}_{\text{optimum}}, \underbrace{\hat{\mu}^{M,N}}_{\text{equilibrium}}, \underbrace{\hat{u}^{M,N}}_{\text{decoupling field}}$ s.t.

$$\hat{\mu}_t^{M,N} = \mathcal{L}(\hat{X}_t^{M,N}|\hat{B}_{t_1}, \dots, \hat{B}_{t_i}), \quad \hat{Y}_t^{M,N} = \hat{u}^{M,N}(t, \hat{X}_t^{M,N})$$

Extraction of converging subsequence

- Conditional measure $\hat{\mu}^{M,N} \sim_{M,N\uparrow\infty} \mathcal{L}(\hat{X}_t^{M,N}|B)$
- Tightness $\hat{X}^{M,N}$ in $\mathcal{C}([0,T]\times\mathbb{R},\mathbb{R})$
  - Standard Kolmogorov criterion
- Tightness $\hat{u}^{M,N}$ in $\mathcal{C}([0,T]\times\mathbb{R},\mathbb{R})$
  - Convexity of Hamiltonian $\Rightarrow$ regularity of $\hat{u}^{M,N}$
- Tightness $\mathcal{L}(\hat{X}_t^{M,N}|B)$ in $\mathcal{C}([0,T],\mathcal{P}(\mathbb{R}))$
  - Given $B=\beta$, $\mathcal{L}(\hat{X}_t^{M,N}|B)$ is the law of

$$d\hat{X}_t^{\beta,M,N} = \hat{b}^{M,N}(t,\hat{X}_t^{\beta,M,N})dt + \sigma dW_t + \varsigma d\beta_t$$

  - Law of $\hat{X}^{\beta,M,N}$ is explicitly controlled by β
- Converging subsequence in weak sense (convergence of distributions)

Uniqueness

- Standard criterion
  - b doesn't depend on μ
  - $f(x, \mu, \alpha) = f_0(x, \mu) + f_1(x, \alpha)$ (μ and α are separated)
  - monotonicity property for f_0 and g w.r.t. μ
- Application: weak $\rightsquigarrow$ strong
- Other strategy: small time $\rightsquigarrow$ long time?
  - a priori Lipschitz estimates of u?
  - degeneracy because of ∞ dim setting

Smoothing effect

- Example: $dX_t = \big(b(\mu_t) + X_t + \alpha_t\big)dt + dW_t + dB_t$

$$J(\alpha) = \mathbb{E}\left[\frac{1}{2}\big(X_T + g(\mu_T)\big)^2 + \int_0^T \frac{1}{2}\big[\big(X_s + f(\mu_s)\big)^2 + \alpha_s^2\big]ds\right]$$

  - b, f, g bounded and smooth
- Adjoint system

$$\begin{aligned} dX_t &= \big(b(\mu_t) + X_t - Y_t\big)dt + dW_t + dB_t \\ dY_t &= -\big(X_t + Y_t + f(\mu_t)\big)dt + dmartingale_t. \\ Y_T &= X_T + g(\mu_T) \end{aligned}$$

  - $Y_t = \eta_t X_t + v(t, \mu_t)$
  - $(\eta_t)_{0 \le t \le T}$ bounded
- Uniqueness related with $\partial_\mu v$?

Integration by parts

- Dynamics of X

$$dX_t = \big((1-\eta_t)X_t + \underbrace{b(\mu_t) - v(t,\mu_t)}_{\beta(t,\mu_t)}\big)dt + dW_t + dB_t.$$

- Sensitivity w.r.t. initial condition $\nabla X = \frac{d}{d\varepsilon}|_{\varepsilon=0} X^{\xi+\varepsilon}$

$$\frac{d}{dt}\nabla X_t = (1-\eta_t)\nabla X_t + \mathbb{E}\big[\big(\partial_\mu \beta(t,\mu_t)(X_t)\nabla X_t|\mathcal{F}_t^B\big]$$

  - $\nabla X_0 = 1 \Rightarrow \nabla X_t$ is $\mathcal{F}_t^B$-meas $\Rightarrow$ linear!

- Malliavin derivative $\Rightarrow \nabla X_t = D_s X_t \nabla X_s$

$$\nabla \mathbb{E}\Big[V(X_T,\mu_T)|\mathcal{F}_0\Big] = \frac{1}{T}\mathbb{E}\Big[V(X_T,\mu_T)\int_0^T \nabla X_s dB_s|\mathcal{F}_0\Big]$$