MFG MASTER EQUATIONS INVOLVING NON-SEPARABLE HAMILTONIANS UNDER DISPLACEMENT MONOTONICITY

Alpár R. MÉSZÁROS
Durham University, UK

[based on works with M. Bansil, W. Gangbo, C. Mou, J. Zhang]

NYUAD SITE SEMINAR

January 2024

- MFG initiated around the same time in 2006, by LASRY-LIONS & HUANG-MALHAMÉ-CAINES

- Aim : study & characterize limits of Nash equilibria of N-player stochastic games as $N \to +\infty$.

- Two main approaches : PDE & probabilistic.

Our setting : • $T > 0$ (arbitrary long time horizon)

• $\mathcal{P}_2(\mathbb{R}^d)$ = space of Borel prob. measures with finite second moment.

• $G : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ terminal cost

• $L : \underset{x}{\mathbb{R}^d} \times \underset{v}{\mathbb{R}^d} \times \underset{\mu}{\mathcal{P}_2(\mathbb{R}^d)} \to \mathbb{R}$ Lagrangian

$$H(x,\cdot,\mu) = L(x,\cdot,\mu)^*$$

• $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$: initial agent population distribution.

• $0 < t_0 < T$: initial time

- A typical agent predicts $(\rho_t)_{t\in[t_0,T]}$, with $\rho_{t_0}=\mu_0$ and solves

$$\inf_{\alpha} \mathbb{E}\left\{ \int_{t_0}^{T} L(X_s,\alpha_s,\rho_s)\,ds + G(X_T;\rho_T)\right\}$$

s.t. $\quad dX_s = \alpha_s\,ds + \beta\,dB_s + \beta_0\,dB^0$

idiosyncratic noise $\qquad$ common noise

- $(B_s)_{s\in[t_0,T]}$ & $(B^0_s)_{s\in[t_0,T]}$ are given independent Brownian motions in $\mathbb{R}^d$.

- Suppose that $\exists$ an optimal process $(X_t^*)_{t\in[t_0,T]}$.
- $(\rho_t)_{t\in[t_0,T]}$ is a MFE if $\rho_t = \mathcal{L}aw\left(X_t^* \mid \mathcal{F}_t^{B^0}\right)$

  $\mathbb{F}^{B^0} = \left(\mathcal{F}_t^{B^0}\right)_{t\in[t_0,T]}$: filtration generated by B^0.

Rmk:

- if $\beta_0 = 0$, $(\rho_t)_{t\in[t_0,T]}$ is deterministic, otherwise this is stochastic, adapted to $\mathbb{F}^{B^0}$.

One can find MFE by solving the MFG system (MFG)

$$\begin{cases} du_t = -\left\{\left(\frac{\beta^2}{2}+\frac{\beta_0^2}{2}\right)\Delta u_t + H(x, D_x u_t; \rho_t) - \beta_0 \operatorname{div}(v_t)\right\} dt + v_t \cdot dB_t^0 \\ \qquad \text{in } (t_0,T)\times\mathbb{R}^d \\ d\rho_t = \left\{\left(\frac{\beta^2}{2}+\frac{\beta_0^2}{2}\right)\Delta\rho_t + \operatorname{div}\left(\rho_t D_p H(x, D_x u_t, \rho_t)\right)\right\} dt - \beta_0 \operatorname{div}(\rho_t dB_t^0) \\ \qquad \text{in } (t_0,T)\times\mathbb{R}^d \\ \rho(t_0,\cdot) = \mu_0; \quad u(T,\cdot) = G(\cdot, \rho_T) \quad \text{in } \mathbb{R}^d. \end{cases}$$

- Equivalently, one can study a system of FBSDEs.

- Central object: master equation introduced by LIONS in his lectures.

- This is set on $(0,T)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ and reads as

(Master)

$$\begin{cases} -\partial_t V(t,x,\mu) - \left(\frac{\beta^2}{2}+\frac{\beta_0^2}{2}\right)\Delta_x V(t,x,\mu) + H(x, D_x V, \mu) + (\mathcal{N} V)(t,x,\mu) = 0 \\ \qquad \text{in } (0,T)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \\ V(T,x,\mu) = G(x,\mu) \quad \text{in } \mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d), \end{cases}$$

where for $\mathcal{U}:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ sufficiently regular

$$(\mathcal{N}\mathcal{U})(x,\mu) := -\left(\frac{\beta^2}{2}+\frac{\beta_0^2}{2}\right)\int_{\mathbb{R}^d}\mathrm{div}_y\left[\partial_\mu \mathcal{U}\right]d\mu(y) + \int_{\mathbb{R}^d}\partial_\mu\mathcal{U}\cdot D_pH(y, D_x\mathcal{U}(y,y))\,d\mu(y)$$

$$-\beta_0^2\int_{\mathbb{R}^d}\mathrm{div}_x\left[\partial_\mu\mathcal{U}\right]d\mu(y) - \frac{\beta_0^2}{2}\iint_{\mathbb{R}^{2d}}\mathrm{tr}\left[\partial^2_{\mu\mu}\mathcal{U}\right]d\mu(y)\,d\mu(y').$$

Here $\partial_\mu \mathcal{U}$ stands for the so-called **W_2-Wasserstein gradient** of $\mathcal{U}$.

[i.e. if $\mathcal{U}(x,\cdot)$ is diff. at μ, then

$$\mathcal{U}(x,\nu) = \mathcal{U}(x,\mu) + \int_{\mathbb{R}^d\times\mathbb{R}^d} \partial_\mu \mathcal{U}(x,\mu)(x')\cdot(y'-x')\, d\gamma(x',y') + o\big(W_2(\mu,\nu)\big) \quad \forall\, \gamma \in \Pi_0(\mu,\nu). \ ]$$

[see Ambrosio–Gigli–Savaré, Birkhäuser, 2008]

[essentially equivalent to the "L-derivative", c.f. Carmona–Delarue]

(set of optimal plans in the def. of $W_2(\mu,\nu)$)

- $W_2^2(\mu;\nu) := \inf\left\{ \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\, d\gamma(x,y) : \gamma \in \Pi(\mu,\nu) \right\}$
- $\Pi(\mu,\nu) := \left\{ \gamma \in \mathcal{P}_2(\mathbb{R}^d\times\mathbb{R}^d) : (\pi^x)_\# \gamma = \mu;\ (\pi^y)_\# \gamma = \nu \right\}$

Derivation of (Master)

$$\mathcal{V}(t_0, x, \mu_0) := u(t_0; x) \quad \& \quad V(t, x, \rho_t) = u(t, x)$$

sol. to MFG HJB with $(t_0; \mu_0)$

$t \in (t_0, T)$. DPP

Formally: $\bullet\ D_x V = D_x u$

$\bullet\ \partial_t u(t,x) = \partial_t \big(V(t, x, \rho_t) \big) = \partial_t V + \langle \partial_\mu V, v_t \rangle_{L^2_{\rho_t}}$,

if $\partial_t \rho + \operatorname{div}(\rho_t v_t) = 0$

$\bullet$ In our case, if $\beta_0 = 0$, $v_t := -\frac{\nabla \rho_t}{\rho_t} - D_p H(x, D_x V, \rho_t)$

so $\langle \partial_\mu V, v_t \rangle_{L^2_{\rho_t}} := \int_{\mathbb{R}^d} \operatorname{div}_y [\partial_\mu V(t, x, \rho_t)(y)] \, d\rho_t(y)$

$$- \int_{\mathbb{R}^d} \partial_\mu V(t, x, \rho_t)(y) \cdot D_p H(y, D_x V(t, x, \rho_t), \rho_t) \, d\rho_t(y)$$

$\bullet$ In general, if $\beta_0 > 0$, we need an Itô lemma.

- Our goal : present a well-posedness result (classical solutions) for (Master) for arbitrary long $T > 0$.

  $\rightsquigarrow$ clearly: one needs to impose assumptions on the data H & G.

  $\rightsquigarrow$ in our setting $\beta > 0$ (so for simplicity we set it $\beta = 1$) and $\beta_0 \geq 0$.

- Most results on classical solutions involve so-called separable Hamiltonians, i.e.

(SEP) $$H(x, p, \mu) = H_0(x, p) - F(x, \mu),$$

for some H_0 & F

Literature review :

Short time well-posedness

- Gangbo – Swiech [JDE, 2015] : potential game

  ($\beta = \beta_0 = 0$; (SEP): $H_0(x,p) = \frac{1}{2}|p|^2$)

  → Extended by Mayorga [JDE, 2020], (SEP), general H_0.
  (OT techniques)

- Bensoussan – Yam [ESAIM: COCV, 2019] ; (SEP),
  $H_0(x,p) = \frac{1}{2}|p|^2$ (Hilbert space techniques)
  ($\beta = \beta_0 = 0$) potential game

- Carmona – Delarue [Book, Vol 2 ; 2018]
  ($\beta > 0$; $\beta_0 \geq 0$) no structural assump.
  on H or G (just smoothness)

Global in time well-posedness

(A.) (SEP) and Lasry–Lions monotonicity of F & G.

(LL) : $\int_{\mathbb{R}^d} \big(F(x;\mu) - F(x;\nu)\big)\, d(\mu-\nu)(x) \geq 0 \quad \forall_{\mu,\nu}$

Potential game : $\exists\ \mathcal{F}, \mathcal{G} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, s.t.

$$\begin{cases} \partial_\mu \mathcal{F}(\mu)(x) = D_x F(x,\mu) & \forall\ (x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \\ \partial_\mu \mathcal{G}(\mu)(x) = D_x G(x,\mu) \end{cases}$$

[i.e. $\frac{\delta}{\delta\mu}\mathcal{F}(\mu)(x) = F(x,\mu)$

and $\frac{\delta}{\delta\mu}\mathcal{G}(\mu)(x) = G(x,\mu)$]

Rmk : In the potential game case (LL) is equivalent to convexity of $\mathcal{F}$ & $\mathcal{G}$ along "flat interpolation" of probability measures,

i.e. $\mathcal{F}((1-\lambda)\mu + \lambda\nu) \leq (1-\lambda)\mathcal{F}(\mu) + \lambda\mathcal{F}(\nu) \quad \forall\ \mu,\nu \in \mathcal{P}_2(\mathbb{R}^d)$
$\forall\ \lambda \in [0,1]$

- Chassagneux – Crisan – Delarue [Mem. AMS, 2022]

$\rightarrow \begin{cases} \beta > 0 \\ \beta_0 = 0 \end{cases}$; (probabilistic techniques)

- Cardaliaguet – Delarue – Lasry – Lions [Ann. Math. Sci., 2019]

$\rightarrow \begin{cases} \beta > 0 \\ \beta_0 \geq 0 \end{cases}$ (PDE techniques)

B (SEP) & displacement monotonicity of the data

$\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is displacement convex (cf. McCann '96) if

(DC) $\mathcal{F}(\mu_t) \leq (1-t)\mathcal{F}(\mu_0) + t\mathcal{F}(\mu_1) \quad \forall \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$

$(\mu_t)_{t\in[0,1]}$ W_2-geod. connecting μ_0 to μ_1.

- Gangbo – M. [CPAM, 2022]
  → $\beta = \beta_0 = 0$; potential game case (DC) data.
  (OT techniques)
- Bensoussan – Graber – Yam [arXiv, 2020] (potential case)
  → $\beta > 0$; $\beta_0 = 0$; $H_0(x,p) = \frac{1}{2}|p|^2$ (Hilbertian tech.)

<u>Remarks</u> on (DM) in MFG related problems

- Ahuja [SICON, 2016] ; Ahuja – Ren – Yang [SPA, 2019]

- Carmona – Delarue [AoP, 2015]
- Chassagneux – Crisan – Delarue [M. of AMS, 2022]

[Various other authors had contributions to master :
→ Bertucci ; Cecchin ; Grant ; Lacker ; Ramanan ;
Porretta ; Souganidis ; etc, ...]

Observation: $\mathcal{F}$ satisfies (DC) and regular, then $\partial_\mu \mathcal{F}$ is monotone in the following sense

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \left(\partial_\mu \mathcal{F}(\mu)(x) - \partial_\mu \mathcal{F}(\nu)(y)\right)\cdot(x-y)\, d\gamma(x,y) \geq 0$$

$$\forall \mu,\nu \in \mathcal{P}_2(\mathbb{R}^d) \quad \forall \gamma \in \Pi(\mu,\nu).$$

- This inspires a definition of monotonicity for not nec. potential games.

Def: $F : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to be displ. mon. if: (DM)

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \left(D_x F(x,\mu) - D_x F(y,\nu)\right)\cdot(x-y)\, d\gamma(x,y) \geq 0.$$

$$\forall \mu,\nu \in \mathcal{P}_2(\mathbb{R}^d) \quad \forall \gamma \in \Pi(\mu,\nu).$$

Our setting : • G satisfies (DM)

• H non-(SEP), but appropriate monotonicity (to described)

[Gangbo-M. Mou-Zhang AoP, 2022]

• $\beta > 0$; $\beta_0 \geq 0$.

Rmk: non-(SEP) Hamiltonians are motivated by

• economical applications; see for instance

[Achdou-Buera-Lasry-Lions-Moll; 2014]

[Achdou-Han-Lasry-Lions-Moll; 2021]

⋮

• congestion modeling

[Achdou-Porretta; 2018]

[Gomes-Voskanyan; 2015]

⋮

Theorem [Gangbo – M. Mou – Zhang ; 2022]

Suppose that $\beta > 0$; $\beta_0 \geq 0$, G & H are sufficiently regular. Suppose that G is (DM), H is convex in p ; H satisfies the (HDM) monotonicity condition. Suppose that $\mu \mapsto G(x,\mu)$ and $\mu \longmapsto H(x,p,\mu)$ are W_1-Lipschitz continuous.

Then (Master) has a global in time classical solution for arbitrary long $T > 0$.

- Need to define (HDM) !

- First observations: $G : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \longrightarrow \mathbb{R}$ is regular enough, then

G is (LL) monotone

$$\Longleftrightarrow \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle \partial^2_{x\mu} G(x,\mu)(y) v(y), v(x) \rangle d\mu(x) d\mu(y) \geqslant 0$$

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d) \quad \forall v \in L^2_\mu(\mathbb{R}^d; \mathbb{R}^d).$$

$$\Longleftrightarrow \mathbb{E}\left[\langle \partial^2_{x\mu} G(\xi; \mathcal{L}_\xi)(\tilde{\xi}) \tilde{\eta}, \eta \rangle \right] \geqslant 0$$

$\forall \ \xi, \eta \ L^2$-R.V. $(\tilde{\xi}, \tilde{\eta})$ is an indep. copy of (ξ, η).

G is (DM)

$$\Longleftrightarrow \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle \partial^2_{x\mu} G(x,\mu)(y) v(y), v(x) \rangle d\mu(x) d\mu(y)$$

$$+ \int_{\mathbb{R}^d} \langle \partial^2_{xx} G(x,\mu) v(x), v(x) \rangle d\mu(x) \geq 0 \quad \forall \mu$$

$$\forall v \in L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$$

$$\Longleftrightarrow \mathbb{E}\left[\langle \partial^2_{x\mu} G(\xi, \mathcal{L}_\xi)(\tilde{\xi}) \tilde{\eta}, \eta \rangle + \langle \partial^2_{xx} G(\xi, \mathcal{L}_\xi) \eta, \eta \rangle \right] \geq 0$$

$$\forall \xi, \eta \ L^2\text{-R.V.}$$

$(\tilde{\xi}, \tilde{\eta})$ indep. copy of (ξ, η).

$$\parallel$$

$$(d_x d)_\xi G(\eta, \eta)$$

Example : $G(x,\mu) = (\phi * \mu)(x)$; $\phi : \mathbb{R}^d \to \mathbb{R}$

smooth, even;

G (LL) monotone $\Longleftrightarrow$ $\hat{\phi} \geq 0$ (Fourier tr.)

G (DM) monotone $\Longleftrightarrow$ ϕ is convex.

Def : $H : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, $\partial^2_{pp} H > 0$ and sufficiently regular is displacement monotone if

(HDM)

$$(\mathrm{displ}^{\varphi}_{\xi} H)(\eta, \eta) := (d_x d)_{\xi} H(\cdot, \varphi(\xi), \cdot)(\eta, \eta)$$

$$+ Q^{\varphi}_{\xi} H(\eta, \eta) \leq 0 \quad \forall\, \xi, \eta \in L^2 \text{ R.V.}$$

$\forall\, \varphi \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ bdd, Lip.,

where

$$Q^{\varphi}_{\xi} H(\eta, \eta) := \frac{1}{4} \mathbb{E}\left\{ \left| \left[\partial^2_{pp} H(\xi, \varphi(\xi), \mathcal{L}_{\xi})\right]^{-\frac{1}{2}} \tilde{\mathbb{E}}\left[\partial_{p\mu} H(\xi, \varphi(\xi), \mathcal{L}_{\xi}, \tilde{\xi})\tilde{\eta}\right] \right|^2 \right\}$$

Strategy of the proof of our main theorem

Step 0. Classical results [Carmona-Delarue, Vol 2] imply that under our assumptions (Master) has a smooth solution for $T>0$ short enough.

Heart of our analysis is

Step 1. (Propagation of (DM))

If G satisfies (DM) & H satisfies (HDM) and V is a regular enough classical solution to (Master), then

$V(t,\cdot,\cdot)$ satisfies (DM) $\forall\, t \in [0,T]$.

(in a quantified way)

$\hookrightarrow$ Proof of Step 1: get estimates on $(d_x d)V(t,\cdot,\cdot)$ by differentiating (Master).

(This is how we discovered the (HDM) condition)

Step 2 : ((DM) implies W_2-Lipschitz)

If $V(t,\cdot,\cdot)$ solves (Master) and satisfies (DM), then

V & $\partial_x V$ are W_2-Lipschitz in the μ-variable!

↳ We show that both $\partial_\mu V(t,x,\mu,\cdot)$ and $\partial^2_{\mu x} V(t,x,\mu,\cdot)$ have uniformly bdd L^2_μ norms

↳ The constants depend on the data and on $\|\partial_x V\|_{L^\infty}$ & $\|\partial^2_{xx} V\|_{L^\infty}$, which is a consequence of the assumptions on the data (in particular that $\partial_x G$ & $\partial^2_{xx} G$ are bdd)

Step 3 ([most subtle one!] W_2-Lipschitz + assumptions on the data $\Longrightarrow$ W_1-Lipschitz)

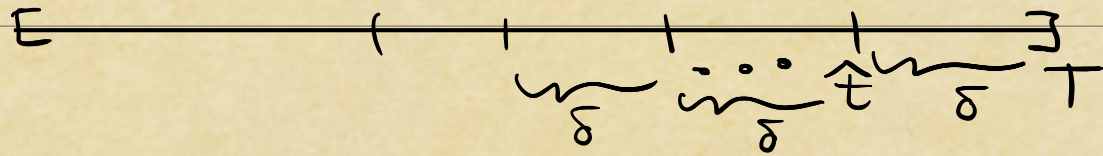
Crucial observation: the local existence result in [Carmona-Delarue, Vol 2] needs W_1-Lipschitz condition on the terminal data.

- Thus, Step 3 is really needed!

$\rightarrow$ To prove this, we show that $\partial_\mu V$ & $\partial_{\mu x} V$ are unif. bounded $\rightsquigarrow$ obtained via a representation formula using the corresponding FBSDE system [cf. Mou-Zhang 2020]

Step 4 (gluing step ; more or less standard)

- Solve (Master) on $[\hat{t}, T]$ of size $\delta > 0$.



- δ depends only on W_2-Lip. constant (and the data)

- Using the previous steps, iterate now with $V(\hat{t}, \cdot, \cdot)$ as final datum.

Some final remarks

- In the case when H satisfies (SEP) & F & G satisfy (LL), so does $V(t,\cdot,\cdot)$ $\forall\, t\in[0,T]$.
- $V(t,\cdot,\cdot)$ (LL) $\Rightarrow$ $V(t,x,\cdot)$ W_1-Lipschitz
- One can go "directly" to the "gluing step".
- However, it remains a challenge to find a condition on H NON-(SEP), that allows the propagation of (LL)-monotonicity!
- Our approach applies to problems with degenerate idiosyncratic noise (also common noise only) [Bansil-M.-Mou, 2023]

Some words about the improvements in [BMM, 2023]

- this approach allows us to handle degenerate idiosyncratic noise, i.e. $\beta = 0$ and $\beta_0 \geq 0$.
- discovered a first order condition to characterise (DM) Hamiltonians (see also [M.-Mau, '24])

$$\mathbb{E}\left\{\left[\partial_p H(X^1, P^1, \mathcal{L}_{X^1}) - \partial_p H(X^2, P^2, \mathcal{L}_{X^2})\right] \cdot (X^1 - X^2)\right\}$$
$$+ \mathbb{E}\left\{\left[-\partial_x H(X^1, P^1, \mathcal{L}_{X^1}) + \partial_x H(X^2, P^2, \mathcal{L}_{X^2})\right] \cdot (P^1 - P^2)\right\} \geq 0$$
$$\forall X^1, P^1, X^2, P^2 \in L^2(\Omega; \mathbb{R}^d)$$

- as a result a this, we relax the assumptions on the regularity for H & G.

THANK YOU FOR
YOUR ATTENTION!