Weak solutions to the master equation of potential mean field games

François Delarue
Université Côte d'Azur (Nice)

GDR MOA, October 11–14 2022

based on joint work with A. Cecchin from Padua

Mean Field Games (MFG) study

- Games = each agent controls her state in order to minimize a proper cost which depends on the other agents' positions
- with infinitely many agents = having individually a negligible influence on the global system
(Ref: Aumann ('64), Schmeidler ('73), Hildenbrand ('74), Mas-Colell ('84), ...)

Mean Field Games (MFG) study

- Games = each agent controls her state in order to minimize a proper cost which depends on the other agents' positions
- with infinitely many agents = having individually a negligible influence on the global system
(Ref: Aumann ('64), Schmeidler ('73), Hildenbrand ('74), Mas-Colell ('84), ...)

Early references:

– Heterogeneous agent models in economy
(Aiyagari ('94), Bewley ('86), Krusell-Smith ('98),...)

– Early works by Lasry-Lions (2005) and Huang-Caines-Malhamé (2005)

- Fix $N \in \mathbb{N}$, $N \geq 2$ the number of players: N large; fix $T > 0$, time horizon.

A class of $N-$player games

- Fix $N \in \mathbb{N}$, $N \geq 2$ the number of players: N large; fix $T > 0$, time horizon.

- Fix $i \in \{1, \ldots, N\}$. Player i wants to minimize over her control (α_t^i) the quantity

$$J^{N,i}(\alpha^i, (\alpha^j)_{j \neq i}) := \mathbb{E}\left[\int_0^T \left(\frac{1}{2}|\alpha_t^i|^2 + F(X_t^i, m_{\boldsymbol{X}_t}^{N,i}) \right) dt + G(X_T^i, m_{\boldsymbol{X}_T}^{N,i}) \right]$$

where $\boldsymbol{X}_t = (X_t^1, \ldots, X_t^N) \in (\mathbb{R}^d)^N$ are the states of the players

A class of $N-$player games

- Fix $N \in \mathbb{N}$, $N \geq 2$ the number of players: N large; fix $T > 0$, time horizon.

- Fix $i \in \{1, \ldots, N\}$. Player i wants to minimize over her control (α_t^i) the quantity

$$J^{N,i}(\alpha^i, (\alpha^j)_{j \neq i}) := \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t^i|^2 + F(X_t^i, m_{\boldsymbol{X}_t}^{N,i})\right) dt + G(X_T^i, m_{\boldsymbol{X}_T}^{N,i})\right]$$

where $\boldsymbol{X}_t = (X_t^1, \ldots, X_t^N) \in (\mathbb{R}^d)^N$ and, for any $j \in \{1, \ldots, N\}$,

$$dX_t^j = \alpha_t^j dt + \sqrt{2} dW_t^j, \quad X_0^j = x_0^j, \quad m_{\boldsymbol{X}_t}^{N,i} = \frac{1}{N-1} \sum_{j \neq i} \delta_{X_t^j},$$

the $(x_0^k)_{k \geq 0}$ being i.i.d. initial conditions and $(W^k)_{k \geq 0}$ being independent Brownian motions of dimension d: W^i is proper noise to i.

A class of $N-$player games

- Fix $N \in \mathbb{N}$, $N \geq 2$ the number of players: N large; fix $T > 0$, time horizon.

- Fix $i \in \{1, \dots, N\}$. Player i wants to minimize over her control (α^i_t) the quantity

$$J^{N,i}(\alpha^i, (\alpha^j)_{j \neq i}) := \mathbb{E}\left[\int_0^T \Big(\tfrac{1}{2}|\alpha^i_t|^2 + F(X^i_t, m^{N,i}_{\boldsymbol{X}_t})\Big) dt + G(X^i_T, m^{N,i}_{\boldsymbol{X}_T})\right]$$

where $\boldsymbol{X}_t = (X^1_t, \dots, X^N_t) \in (\mathbb{R}^d)^N$ and, for any $j \in \{1, \dots, N\}$,

$$dX^j_t = \alpha^j_t dt + \sqrt{2} dW^j_t, \quad X^j_0 = x^j_0, \quad m^{N,i}_{\boldsymbol{X}_t} = \frac{1}{N-1}\sum_{j \neq i} \delta_{X^j_t},$$

the $(x^k_0)_{k \geq 0}$ being i.i.d. initial conditions and $(W^k)_{k \geq 0}$ being independent Brownian motions of dimension d: W^i is proper noise to i.

A class of $N-$player games

- Fix $N \in \mathbb{N}$, $N \geq 2$ the number of players: N large; fix $T > 0$, time horizon.

- Fix $i \in \{1, \dots, N\}$. Player i wants to minimize over her control (α_t^i) the quantity

$$J^{N,i}(\alpha^i, (\alpha^j)_{j \neq i}) := \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t^i|^2 + F(X_t^i, m_{\boldsymbol{X}_t}^{N,i}) \right) dt + G(X_T^i, m_{\boldsymbol{X}_T}^{N,i}) \right]$$

where $\boldsymbol{X}_t = (X_t^1, \dots, X_t^N) \in (\mathbb{R}^d)^N$ and, for any $j \in \{1, \dots, N\}$,

$$dX_t^j = \alpha_t^j dt + \sqrt{2} dW_t^j, \quad X_0^j = x_0^j, \quad m_{\boldsymbol{X}_t}^{N,i} = \frac{1}{N-1} \sum_{j \neq i} \delta_{X_t^j},$$

the $(x_0^k)_{k \geq 0}$ being i.i.d. initial conditions and $(W^k)_{k \geq 0}$ being independent Brownian motions of dimension d: W^i is proper noise to i.

F, G are interaction costs.

A class of $N-$player games

- Fix $N \in \mathbb{N}$, $N \geq 2$ the number of players: N large; fix $T > 0$, time horizon.

- Fix $i \in \{1, \dots, N\}$. Player i wants to minimize over her control (α_t^i) the quantity

$$J^{N,i}(\alpha^i, (\alpha^j)_{j \neq i}) := \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t^i|^2 + F(X_t^i, m_{\boldsymbol{X}_t}^{N,i})\right) dt + G(X_T^i, m_{\boldsymbol{X}_T}^{N,i})\right]$$

where $\boldsymbol{X}_t = (X_t^1, \dots, X_t^N) \in (\mathbb{R}^d)^N$ and, for any $j \in \{1, \dots, N\}$,

$$dX_t^j = \alpha_t^j dt + \sqrt{2} dW_t^j, \quad X_0^j = x_0^j, \quad ,$$

the $(x_0^k)_{k \geq 0}$ being i.i.d. initial conditions and $(W^k)_{k \geq 0}$ being independent Brownian motions of dimension d: W^i is proper noise to i.

F, G are interaction costs.

- Nash equilibria (After Nash ('51)): We say that $\bar{\alpha} = (\bar{\alpha}^1, \dots, \bar{\alpha}^N)$ is a Nash Equilibrium if

$$J^{N,i}(\bar{\alpha}^i, (\bar{\alpha})_{j \neq i}) \leq J^{N,i}(\alpha^i, (\bar{\alpha})_{j \neq i}) \qquad \forall \alpha^i, \ \forall i.$$

- When players play closed-loop controls: $\bar{\alpha}^i_t = \bar{\alpha}^i(t, X^1_t, \dots, X^N_t)$, the value function

$$v^{N,i} : \text{initial states } (t, \boldsymbol{x}) = (t, x^1, \dots, x^N) \mapsto \text{equilibrium cost } v^{N,i}(t, \boldsymbol{x})$$

satisfies the Nash system:

$$(\text{Nash}) \begin{cases} -\partial_t v^{N,i}(t,\boldsymbol{x}) - \mathcal{L}^{N,i}\Big(\boldsymbol{x}, D^2_{\boldsymbol{x}\boldsymbol{x}} v^{N,i}(t,\boldsymbol{x}), \big(D_{\boldsymbol{x}} v^{N,j}(t,\boldsymbol{x})\big)_j\Big) = 0 \quad \text{in } [0,T] \times (\mathbb{R}^d)^N, \\ v^{N,i}(T,\boldsymbol{x}) = G(x_i, m^{N,i}_{\boldsymbol{x}}) \quad \text{in } (\mathbb{R}^d)^N. \end{cases}$$

From the Nash system to the equilibrium trajectories

- When players play closed-loop controls: $\bar{\alpha}^i_t = \bar{\alpha}^i(t, X^1_t, \ldots, X^N_t)$, the value function

$$v^{N,i} : \text{initial states } (t, \boldsymbol{x}) = (t, x^1, \ldots, x^N) \mapsto \text{equilibrium cost } v^{N,i}(t, \boldsymbol{x})$$

satisfies the Nash system:

$$(\text{Nash}) \begin{cases} -\partial_t v^{N,i}(t,\boldsymbol{x}) - \mathcal{L}^{N,i}\Big(\boldsymbol{x}, D^2_{\boldsymbol{xx}} v^{N,i}(t,\boldsymbol{x}), \big(D_{\boldsymbol{x}} v^{N,j}(t,\boldsymbol{x})\big)_j\Big) = 0 \quad \text{in } [0,T] \times (\mathbb{R}^d)^N, \\ v^{N,i}(T,\boldsymbol{x}) = G(x_i, m^{N,i}_{\boldsymbol{x}}) \quad \text{in } (\mathbb{R}^d)^N. \end{cases}$$

- We denote by $\boldsymbol{X}^N_t = (X^{N,1}_t, \ldots, X^{N,N}_t)$ the "optimal trajectories" of the $N-$player game: they solve the system of N coupled stochastic differential equations (SDE):

$$dX^{N,i}_t = -D_{x_i} v^{N,i}(t, \boldsymbol{X}^N_t)dt + \sqrt{2}dW^i_t, \quad t \in [0,T].$$

From the Nash system to the equilibrium trajectories

- When players play closed-loop controls: $\bar{\alpha}_t^i = \bar{\alpha}^i(t, X_t^1, \ldots, X_t^N)$, the value function

$$v^{N,i} : \text{initial states } (t, \boldsymbol{x}) = (t, x^1, \ldots, x^N) \mapsto \text{equilibrium cost } v^{N,i}(t, \boldsymbol{x})$$

satisfies the Nash system:

$$(\text{Nash}) \quad \begin{cases} -\partial_t v^{N,i}(t, \boldsymbol{x}) - \mathcal{L}^{N,i}\Big(\boldsymbol{x}, D^2_{\boldsymbol{xx}} v^{N,i}(t, \boldsymbol{x}), \big(D_{\boldsymbol{x}} v^{N,j}(t, \boldsymbol{x})\big)_j\Big) = 0 \quad \text{in } [0, T] \times (\mathbb{R}^d)^N, \\ v^{N,i}(T, \boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i}) \quad \text{in } (\mathbb{R}^d)^N. \end{cases}$$

- We denote by $\boldsymbol{X}_t^N = (X_t^{N,1}, \ldots, X_t^{N,N})$ the "optimal trajectories" of the $N-$player game: they solve the system of N coupled stochastic differential equations (SDE):

$$dX_t^{N,i} = -D_{x_i} v^{N,i}(t, \boldsymbol{X}_t^N)dt + \sqrt{2}dW_t^i, \quad t \in [0, T].$$

- **Aim:** $\boxed{\text{We are interested in the behavior, as } N \to \infty \text{, of the } (v^{N,i}) \text{ and of the } (X^{N,i}).}$

Expected limit: the Mean Field Game

- **"The master equation":**
  By symmetry property, the solution $(v^{N,i})_i$ of the Nash system can be written in the form
  $$v^{N,i}(t,\boldsymbol{x}) = V^N\big(t, x^i, m^{N,i}\big).$$
  - The (formal) limit U of V^N is expected to satisfy the Master equation.

The master equation is a (backward) nonlinear nonlocal transport PDE set on $[0, T] \times \mathcal{P}(\mathbb{R}^d)$.

- The unknown is $U = U(t, x, m) : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$.

The master equation is a (backward) nonlinear nonlocal transport PDE set on $[0,T] \times \mathcal{P}(\mathbb{R}^d)$.

- The unknown is $U = U(t,x,m) : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$.
- The equation reads

$$\begin{cases} -\partial_t U - \mathcal{L}\Big(x, m, D_{(x,m)} U(t,x,m), D^2_{(x,m)} U(t,x,m)\Big) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d), \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d). \end{cases}$$

where

- D_m, D^2_{mm}, ... are derivatives on the space $\mathcal{P}(\mathbb{R}^d)$: explained later,
- $\mathcal{L}$ is expected to be the asymptotic form of $\mathcal{L}^N$ in the Nash system.

The master equation is a (backward) nonlinear nonlocal transport PDE set on $[0,T] \times \mathcal{P}(\mathbb{R}^d)$.

- The unknown is $U = U(t,x,m) : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$.
- The equation reads

$$\begin{cases} -\partial_t U - \mathcal{L}\Big(x, m, D_{(x,m)} U(t,x,m), D^2_{(x,m)} U(t,x,m)\Big) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0,T) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d), \\ U(T,x,m) = G(x,m) \quad \text{in } \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d). \end{cases}$$

where

- D_m, D^2_{mm}, ... are derivatives on the space $\mathcal{P}(\mathbb{R}^d)$: explained later,
- $\mathcal{L}$ is expected to be the asymptotic form of $\mathcal{L}^N$ in the Nash system.

"Goal of the MFG theory"

- Analysis of the various MFG formulations and equilibria (including master equation).
- Derive these models from the Nash system as $N \to \infty$ (mean field limit). Classical solutions to the master equation give the rate (Cardliaguet, D., Lasry, Lions).

Derivatives in the space of measures

We consider 2 notions of derivatives of a map $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$:

- The directional derivative
  see, e.g., Dawson, Kolokoltsov, Mischler-Mouhot
- The intrinsic derivative
  see, e.g., Otto, Ambrosio-Gigli-Savaré, Albeverio-Kondratiev-Röckner, Lions

Derivatives in the space of measures

We consider 2 notions of derivatives of a map $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$:

- The directional derivative $\frac{\delta U}{\delta m}(m, y)$
  see, e.g., Dawson, Kolokoltsov, Mischler-Mouhot
- The intrinsic derivative $D_m U(m, y)$
  see, e.g., Otto, Ambrosio-Gigli-Savaré, Albeverio-Kondratiev-Röckner, Lions

Derivatives

A map $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is C^1 if the limit

$$\frac{\delta U}{\delta m}(m, y) = \lim_{h \to 0^+} \frac{U((1-h)m + h\delta_y) - U(m)}{h}$$

exists, is continuous and bounded. It satisfies $\displaystyle\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m, y) dm(y) = 0$.

Derivatives in the space of measures

We consider 2 notions of derivatives of a map $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$:

- The directional derivative $\frac{\delta U}{\delta m}(m, y)$
  see, e.g., Dawson, Kolokoltsov, Mischler-Mouhot
- The intrinsic derivative $D_m U(m, y)$
  see, e.g., Otto, Ambrosio-Gigli-Savaré, Albeverio-Kondratiev-Röckner, Lions

Derivatives

A map $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is C^1 if the limit

$$\frac{\delta U}{\delta m}(m, y) = \lim_{h \to 0^+} \frac{U((1-h)m + h\delta_y) - U(m)}{h}$$

exists, is continuous and bounded. We set $D_m U(m, y) := D_y \frac{\delta U}{\delta m}(m, y)$ (when the derivative exists)

Detour (1) : Optimal control of large particle systems

- Objective is to study master equation by using connection with control...

  ... requires a detour: **cooperative instead of competitive models.** In MFG theory, players are in competition. What about a cooperative version?

- Objective is to study master equation by using connection with control...

  ... requires a detour: **cooperative instead of competitive models.** In MFG theory, players are in competition. What about a cooperative version?

- We are interested in the optimal control of large particle systems of the form

$$\min_{(\alpha^{N,i})_{i=1,\dots,N}} \mathbb{E}\left[\int_0^T \left(\frac{1}{N}\sum_{i=1}^N \frac{1}{2}|\alpha_t^{N,i}|^2 + \mathcal{F}(m^N_{\boldsymbol{X}_t^N})\right)dt + \mathcal{G}(m^N_{\boldsymbol{X}_T^N})\right],$$

where, as before, $m^N_{\boldsymbol{X}_t^N} = \dfrac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}$ and, for $i = 1, \dots, N$,

$$dX_t^{N,i} = \alpha_t^{N,i}dt + \sqrt{2}dW_t^i,$$

Detour (1) : Optimal control of large particle systems

- Objective is to study master equation by using connection with control...

  ... requires a detour: **cooperative instead of competitive models.** In MFG theory, players are in competition. What about a cooperative version?

- We are interested in the optimal control of large particle systems of the form

$$\min_{(\alpha^{N,i})_{i=1,\dots,N}} \mathbb{E}\left[\int_0^T \left(\frac{1}{N}\sum_{i=1}^N \frac{1}{2}|\alpha_t^{N,i}|^2 + \mathcal{F}\big(m^N_{\boldsymbol{X}_t^N}\big)\right)dt + \mathcal{G}\big(m^N_{\boldsymbol{X}_T^N}\big)\right].$$

- main differences with MFG...:
  - cost to the society: we look for a true minimizer,
  - $\mathcal{F}, \mathcal{G} : \mathcal{P} \to \mathbb{R}$ are energy costs on the whole state of the population.

Detour (1) : Optimal control of large particle systems

- Objective is to study master equation by using connection with control...

  ... requires a detour: **cooperative instead of competitive models.** In MFG theory, players are in competition. What about a cooperative version?

- We are interested in the optimal control of large particle systems of the form

$$\min_{(\alpha^{N,i})_{i=1,\dots,N}} \mathbb{E}\left[\int_0^T \Big(\frac{1}{N}\sum_{i=1}^N \tfrac{1}{2}|\alpha_t^{N,i}|^2 + \mathcal{F}\big(m^N_{\boldsymbol{x}^N_t}\big)\Big)dt + \mathcal{G}\big(m^N_{\boldsymbol{x}^N_T}\big)\right].$$

- main differences with MFG...:
  - cost to the society: we look for a true minimizer,
  - $\mathcal{F}, \mathcal{G} : \mathcal{P} \to \mathbb{R}$ are energy costs on the whole state of the population.

- ... but similar questions: behavior of the value functions and of the optimal trajectories as $N \to +\infty$.

Detour (1) : Optimal control of large particle systems

- Objective is to study master equation by using connection with control...

  ... requires a detour: **cooperative instead of competitive models.** In MFG theory, players are in competition. What about a cooperative version?

- We are interested in the optimal control of large particle systems of the form

$$\min_{(\alpha^{N,i})_{i=1,\dots,N}} \mathbb{E}\left[\int_0^T \left(\frac{1}{N}\sum_{i=1}^N \frac{1}{2}|\alpha_t^{N,i}|^2 + \mathcal{F}(m^N_{\mathbf{X}^N_t})\right)dt + \mathcal{G}(m^N_{\mathbf{X}^N_T})\right].$$

- main differences with MFG...:
  - cost to the society: we look for a true minimizer,
  - $\mathcal{F}, \mathcal{G} : \mathcal{P} \to \mathbb{R}$ are energy costs on the whole state of the population.

- ... but similar questions: behavior of the value functions and of the optimal trajectories as $N \to +\infty$.

  Here, the value function $\mathcal{V}^N$ is defined by:

$$\mathcal{V}^N(0, \mathbf{x}_0^N) := \inf_{(\alpha^{N,i})_{i=1,\dots,N}} \mathbb{E}\left[\int_0^T \left(\frac{1}{N}\sum_{i=1}^N \frac{1}{2}|\alpha_t^{N,i}|^2 + \mathcal{F}(m^N_{\mathbf{X}^N_t})\right)dt + \mathcal{G}(m^N_{\mathbf{X}^N_T})\right],$$

  when $X_0^{N,i} = x_0^{N,i}$.

Detour (2) : Mean field control

- The mean field limit: Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\left[\int_0^T \Big(\frac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t|W^0)\big)\Big) dt + \mathcal{G}\big(\mathcal{L}(X_T|W^0)\big)\right]$$

where

$$dX_t = \alpha_t dt + \sqrt{2} dW_t,$$

a.k.a. mean field control.

See also Kolokoltsov ('12) and Cecchin ('21) for finite state space problems, Gangbo-Mayorga-Swiech ('21) for a HJ approach, Bayraktar et al. ('18) and Djete ('19) for various equivalent formulations

Detour (2) : Mean field control

- The mean field limit: Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t|W^0)\big)\right) dt + \mathcal{G}\big(\mathcal{L}(X_T|W^0)\big)\right]$$

  where

$$dX_t = \alpha_t dt + \sqrt{2} dW_t,$$

  a.k.a. mean field control.

  See also Kolokoltsov ('12) and Cecchin ('21) for finite state space problems, Gangbo-Mayorga-Swiech ('21) for a HJ approach, Bayraktar et al. ('18) and Djete ('19) for various equivalent formulations

- Value function (at time $t = 0$)

$$\mathcal{U}(0, m_0) = \inf_{\alpha} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t)\big)\right) dt + \mathcal{G}\big(\mathcal{L}(X_T)\big)\right]$$

  - If smooth, it is a classical solution of related Hamilton-Jacobi equation (see next) on the space of probability measures.

Detour (2) : Mean field control

- The mean field limit: Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t|W^0)\big)\right) dt + \mathcal{G}\big(\mathcal{L}(X_T|W^0)\big)\right]$$

  where

$$dX_t = \alpha_t dt + \sqrt{2} dW_t,$$

  a.k.a. mean field control.

  See also Kolokoltsov ('12) and Cecchin ('21) for finite state space problems, Gangbo-Mayorga-Swiech ('21) for a HJ approach, Bayraktar et al. ('18) and Djete ('19) for various equivalent formulations

- Value function (at time $t = 0$)

$$\mathcal{U}(0, m_0) = \inf_{\alpha} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t)\big)\right) dt + \mathcal{G}\big(\mathcal{L}(X_T)\big)\right]$$

  - If smooth, it is a classical solution of related Hamilton-Jacobi equation on the space of probability measures.
  - Example for smoothness: if $\mathcal{F}$ and $\mathcal{G}$ are convex on the space of measures. (Cardaliaguet-D.-Lasry-Lions ('19))

- Mean field control: Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\Big[\int_0^T \Big(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t)\big)\Big)dt + \mathcal{G}\big(\mathcal{L}(X_T)\big)\Big]$$

where

$$dX_t = \alpha_t dt + \sqrt{2}dW_t.$$

  - Value function (at time $t = 0$)

$$\mathcal{U}(0, m_0) = \inf_{\alpha} \mathbb{E}\Big[\int_0^T \Big(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t)\big)\Big)dt + \mathcal{G}\big(\mathcal{L}(X_T)\big)\Big]$$

Back to MFG: Potential games

- Mean field control: Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}(\mathcal{L}(X_t))\right) dt + \mathcal{G}(\mathcal{L}(X_T))\right]$$

where

$$dX_t = \alpha_t dt + \sqrt{2} dW_t.$$

  - Value function (at time $t = 0$)

$$\mathcal{U}(0, m_0) = \inf_{\alpha} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}(\mathcal{L}(X_t))\right) dt + \mathcal{G}(\mathcal{L}(X_T))\right]$$

- **Potential MFG.** Consider MFG with

$$D_m\mathcal{F}(m, y) = \partial_y F(y, m), \quad D_m\mathcal{G}(m, y) = \partial_y G(y, m).$$

- Mean field control: Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}(\mathcal{L}(X_t))\right) dt + \mathcal{G}(\mathcal{L}(X_T))\right]$$

where

$$dX_t = \alpha_t dt + \sqrt{2} dW_t.$$

  - Value function (at time $t = 0$)

$$\mathcal{U}(0, m_0) = \inf_{\alpha} \mathbb{E}\left[\int_0^T \left(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}(\mathcal{L}(X_t))\right) dt + \mathcal{G}(\mathcal{L}(X_T))\right]$$

- **Potential MFG.** Consider MFG with

$$D_m\mathcal{F}(m, y) = \partial_y F(y, m), \quad D_m\mathcal{G}(m, y) = \partial_y G(y, m).$$

  - Minimizers of the mean field control problem are equilibria of the MFG!

- **Mean field control:** Following Carmona, D., Lachapelle ('14), Lacker ('17) and Djete-Possamaï-Tan ('19), the limit problem as $N \to +\infty$ is expected to be (a weak solution of) an optimal control problem of a McKean-Vlasov type

$$\inf_{(\alpha_t)_t} \mathbb{E}\Big[\int_0^T \Big(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t)\big)\Big) dt + \mathcal{G}\big(\mathcal{L}(X_T)\big)\Big]$$

where

$$dX_t = \alpha_t dt + \sqrt{2} dW_t.$$

  - Value function (at time $t = 0$)

$$\mathcal{U}(0, m_0) = \inf_{\alpha} \mathbb{E}\Big[\int_0^T \Big(\tfrac{1}{2}|\alpha_t|^2 + \mathcal{F}\big(\mathcal{L}(X_t)\big)\Big) dt + \mathcal{G}\big(\mathcal{L}(X_T)\big)\Big]$$

- **Potential MFG.** Consider MFG with

$$D_m\mathcal{F}(m, y) = \partial_y F(y, m), \quad D_m\mathcal{G}(m, y) = \partial_y G(y, m).$$

  - Minimizers of the mean field control problem are equilibria of the MFG!
  - Link with the master equation: When $\mathcal{F}$ and $\mathcal{G}$ are convex, then F and G are monotone and the solution U of the master equation is given by ($\beta = 0$)

$$\partial_x U(t, x, m) = D_m\mathcal{U}(t, m, x) \quad \text{(Cardaliaguet-D.-Lasry-Lions ('19))}.$$

Outside convexity/monotonicity

- Without convexity, smoothness of $\mathcal{U}$ may be lost, even if the data of smooth: $\mathcal{U}$ may not be a classical sense, but is 'not so far'.

Theorem 6 (Cardaliaguet-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), the map $\mathcal{U}$ is globally Lipschitz continuous on $[0, T] \times \mathcal{P}$ and there exists an open and dense subset $\mathcal{O}$ on which $\mathcal{U}$ is of class C^1. Moreover $\mathcal{U}$ satisfies in a classical sense in $\mathcal{O}$ the Hamilton-Jacobi equation:

$$-\partial_t \mathcal{U}(t, m) - \int_{\mathbb{R}^d} \mathrm{div}(D_m \mathcal{U}(t, m, y)) m(dy) + \int_{\mathbb{R}^d} \frac{1}{2} |D_m \mathcal{U}(t, m, y)|^2 m(dy) = \mathcal{F}(m).$$

Outside convexity/monotonicity

Theorem 6 (Cardaliaguet-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), the map $\mathcal{U}$ is globally Lipschitz continuous on $[0, T] \times \mathcal{P}$ and there exists an open and dense subset $\mathcal{O}$ on which $\mathcal{U}$ is of class C^1. Moreover $\mathcal{U}$ satisfies in a classical sense in $\mathcal{O}$ the Hamilton-Jacobi equation:

$$-\partial_t \mathcal{U}(t,m) - \int_{\mathbb{R}^d} \mathrm{div}(D_m \mathcal{U}(t,m,y)) m(dy) + \int_{\mathbb{R}^d} \frac{1}{2} |D_m \mathcal{U}(t,m,y)|^2 m(dy) = \mathcal{F}(m).$$

- HJ at many many points: uniqueness can be forced by requiring more on the solution.

Outside convexity/monotonicity

Theorem 6 (Cardaliaguet-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), the map $\mathcal{U}$ is globally Lipschitz continuous on $[0, T] \times \mathcal{P}$ and there exists an open and dense subset $\mathcal{O}$ on which $\mathcal{U}$ is of class C^1. Moreover $\mathcal{U}$ satisfies in a classical sense in $\mathcal{O}$ the Hamilton-Jacobi equation:

$$-\partial_t \mathcal{U}(t, m) - \int_{\mathbb{R}^d} \mathrm{div}(D_m \mathcal{U}(t, m, y)) m(dy) + \int_{\mathbb{R}^d} \frac{1}{2} |D_m \mathcal{U}(t, m, y)|^2 m(dy) = \mathcal{F}(m).$$

Theorem 7 (Cecchin-D. ('22))

Under regularity assumptions on the data (but no convexity) and in the periodic setting, there exists a probability measure $\mathbb{P}$ on $\mathcal{P}_1$ with full support such that $\mathcal{U}$ is the unique globally Lipschitz and displacement semi-concave that 'solves' $\mathrm{Leb} \otimes \mathbb{P}$ almost everywhere the HJ equation.

Outside convexity/monotonicity

Theorem 6 (Cardaliaguet-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), the map $\mathcal{U}$ is globally Lipschitz continuous on $[0, T] \times \mathcal{P}$ and there exists an open and dense subset $\mathcal{O}$ on which $\mathcal{U}$ is of class C^1. Moreover $\mathcal{U}$ satisfies in a classical sense in $\mathcal{O}$ the Hamilton-Jacobi equation:

$$-\partial_t \mathcal{U}(t, m) - \int_{\mathbb{R}^d} \mathrm{div}(D_m \mathcal{U}(t, m, y)) m(dy) + \int_{\mathbb{R}^d} \frac{1}{2} |D_m \mathcal{U}(t, m, y)|^2 m(dy) = \mathcal{F}(m).$$

Theorem 7 (Cecchin-D. ('22))

Under regularity assumptions on the data (but no convexity) and in the periodic setting, there exists a probability measure $\mathbb{P}$ on $\mathcal{P}_1$ with full support such that $\mathcal{U}$ is the unique globally Lipschitz and displacement semi-concave that 'solves' $\mathrm{Leb} \otimes \mathbb{P}$ almost everywhere the HJ equation.

- Displacement convexity: convexity along geodescis in the space of probability measures.

  Use of semi-concavity: reminiscent of HJ in finite dimension, Kruzkov ('60), Douglis ('61)
- There is a form of Rademacher theorem that ensures that derivatives exist almost surely.
- The almost everywhere formulation of the HJ equation requires some care: the equation is formulated on finite dimensional slices obtained by truncating the Fourier expansion of m.

Theorem 6 (Cardaliaguet-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), the map $\mathcal{U}$ is globally Lipschitz continuous on $[0, T] \times \mathcal{P}$ and there exists an open and dense subset $\mathcal{O}$ on which $\mathcal{U}$ is of class C^1. Moreover $\mathcal{U}$ satisfies in a classical sense in $\mathcal{O}$ the Hamilton-Jacobi equation:

$$-\partial_t \mathcal{U}(t, m) - \int_{\mathbb{R}^d} \mathrm{div}(D_m \mathcal{U}(t, m, y)) m(dy) + \int_{\mathbb{R}^d} \frac{1}{2} |D_m \mathcal{U}(t, m, y)|^2 m(dy) = \mathcal{F}(m).$$

Theorem 7 (Cecchin-D. ('22))

Under regularity assumptions on the data (but no convexity) and in the periodic setting, there exists a probability measure $\mathbb{P}$ on $\mathcal{P}_1$ with full support such that $\mathcal{U}$ is the unique globally Lipschitz and displacement semi-concave that 'solves' $\mathrm{Leb} \otimes \mathbb{P}$ almost everywhere the HJ equation.

- Displacement convexity: convexity along geodescis in the space of probability measures.
- In finite dimension, connection between semi-concave solutions to HJ equations and viscosity solutions, see Lions ('82) .

Outside convexity/monotonicity

Theorem 6 (Cardaliaguet-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), the map $\mathcal{U}$ is globally Lipschitz continuous on $[0, T] \times \mathcal{P}$ and there exists an open and dense subset $\mathcal{O}$ on which $\mathcal{U}$ is of class C^1. Moreover $\mathcal{U}$ satisfies in a classical sense in $\mathcal{O}$ the Hamilton-Jacobi equation:

$$-\partial_t \mathcal{U}(t,m) - \int_{\mathbb{R}^d} \mathrm{div}(D_m \mathcal{U}(t,m,y)) m(dy) + \int_{\mathbb{R}^d} \frac{1}{2} |D_m \mathcal{U}(t,m,y)|^2 m(dy) = \mathcal{F}(m).$$

Theorem 7 (Cecchin-D. ('22))

Under regularity assumptions on the data (but no convexity) and in the periodic setting, there exists a probability measure $\mathbb{P}$ on $\mathcal{P}_1$ with full support such that $\mathcal{U}$ is the unique globally Lipschitz and displacement semi-concave that 'solves' $\mathrm{Leb} \otimes \mathbb{P}$ almost everywhere the HJ equation.

- Displacement convexity: convexity along geodescis in the space of probability measures.
- In finite dimension, connection between semi-concave solutions to HJ equations and viscosity solutions, see Lions ('82) .

  Analysis of viscosity solutions is not straightforward, see Burzoni et al' ('20), Conforti et al ('21), Cosso et al. ('22).

Theorem 8 (Cardaliaguet.-Daudin-Jackson-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), there exists $\gamma \in (0, 1]$ (depending only on d) and $C > 0$ (depending on the data) such that, for any $(t, \boldsymbol{x}) \in [0, T] \times (\mathbb{R}^d)^N$,

$$\left| \mathcal{V}^N(t, \boldsymbol{x}) - \mathcal{U}(t, m^N_{\boldsymbol{x}}) \right| \leq C N^{-\gamma} \Big(1 + \frac{1}{N} \sum_{i=1}^{N} |x_i|^2 \Big).$$

- The proof is based on semiconcavity estimates on $\mathcal{V}^N$ and on concentration inequalities.

Theorem 8 (Cardaliaguet.-Daudin-Jackson-Souganidis ('22))

Under regularity assumptions on the data (but no convexity), there exists $\gamma \in (0,1]$ (depending only on d) and $C > 0$ (depending on the data) such that, for any $(t, \boldsymbol{x}) \in [0,T] \times (\mathbb{R}^d)^N$,

$$\Big| \mathcal{V}^N(t, \boldsymbol{x}) - \mathcal{U}(t, m^N_{\boldsymbol{x}}) \Big| \leq C N^{-\gamma} \Big(1 + \frac{1}{N} \sum_{i=1}^N |x_i|^2 \Big).$$

- The proof is based on semiconcavity estimates on $\mathcal{V}^N$ and on concentration inequalities.
- Moreover, Theorem 6 yields to (quantitative) propagation of chaos for the optimal trajectories.

Theorem 9 (Cecchin-D. ('22))

Within the framework of Theorem 7, the function $\bar{U} : (t, x, m) \mapsto \frac{\delta \mathcal{U}}{\delta m}(t, m, x)$ is the unique solution, in a 'weak sense', of

$$(\bar{\boldsymbol{M}}) \left\{ \begin{array}{l} -\partial_t \bar{U} - \Delta \bar{U} + \frac{1}{2}|D_x \bar{U}|^2 + \int_{\mathbb{T}^d} D_m \bar{U}(\cdot, y) \cdot D_x \bar{U} m(dy) \\ \quad - \int_{\mathbb{T}^d} \mathrm{Tr}\big(D^2_{ym} \bar{U}\big)(\cdot, y) m(dy) = F(x, m) + C(t, m). \end{array} \right.$$

- Equation ($\bar{\boldsymbol{M}}$) is the centred version of the master equation: $C(t, m)$ guarantees that $\bar{U}(t, x, m, \cdot)$ has zero mean w.r.t. m.

  Equation ($\bar{\boldsymbol{M}}$) is also $\frac{\delta}{\delta m}$HJ

Theorem 9 (Cecchin-D. ('22))

Within the framework of Theorem 7, the function $\bar{U} : (t, x, m) \mapsto \dfrac{\delta \mathcal{U}}{\delta m}(t, m, x)$ is the unique solution, in a 'weak sense', of

$$(\bar{\boldsymbol{M}}) \left\{ \begin{aligned} & -\partial_t \bar{U} - \Delta \bar{U} + \tfrac{1}{2}|D_x \bar{U}|^2 + \int_{\mathbb{T}^d} D_m \bar{U}(\cdot, y) \cdot D_x \bar{U} m(dy) \\ & \quad - \int_{\mathbb{T}^d} \mathrm{Tr}\big(D^2_{ym} \bar{U}\big)(\cdot, y) m(dy) = F(x, m) + C(t, m). \end{aligned} \right.$$

- Equation ($\bar{\boldsymbol{M}}$) is the centred version of the master equation: $C(t, m)$ guarantees that $\bar{U}(t, x, m, \cdot)$ has zero mean w.r.t. m.

  Equation ($\bar{\boldsymbol{M}}$) is also $\dfrac{\delta}{\delta m}$HJ

- Any weak solution derives from a potential: uniqueness requires the potential to be semi-concave. This is a generalization of Kruzkov ('67) in finite dimension.

Theorem 9 (Cecchin-D. ('22))

Within the framework of Theorem 7, the function $\bar{U} : (t, x, m) \mapsto \frac{\delta \mathcal{U}}{\delta m}(t, m, x)$ is the unique solution, in a 'weak sense', of

$$(\bar{\boldsymbol{M}}) \left\{ \begin{aligned} &-\partial_t \bar{U} - \Delta \bar{U} + \tfrac{1}{2}|D_x \bar{U}|^2 + \int_{\mathbb{T}^d} D_m \bar{U}(\cdot, y) \cdot D_x \bar{U} m(dy) \\ &\quad - \int_{\mathbb{T}^d} \mathrm{Tr}\big(D^2_{ym} \bar{U}\big)(\cdot, y) m(dy) = F(x, m) + C(t, m). \end{aligned} \right.$$

- Equation $(\bar{\boldsymbol{M}})$ is the centred version of the master equation: $C(t, m)$ guarantees that $\bar{U}(t, x, m, \cdot)$ has zero mean w.r.t. m.

  Equation $(\bar{\boldsymbol{M}})$ is also $\frac{\delta}{\delta m}$HJ

- Any weak solution derives from a potential: uniqueness requires the potential to be semi-concave. This is a generalization of Kruzkov ('67) in finite dimension.

- Integration by parts for weak solution is understood on finite dimensional slices obtained by truncating the Fourier expansion of m and by using the fact that $\frac{\delta}{\delta m}$ identifies with the derivatives w.r.t. $(\widehat{m}^k)_k$:

$$\widehat{\frac{\delta}{\delta m}}^k = \partial_{\widehat{m}^k}.$$

Using Fourier analysis

- Use periodic setting by working on $\mathcal{P}(\mathbb{T}^d)$ and

$$\phi(m),\ m \in \mathcal{P}(\mathbb{T}^d) \ \Rightarrow\ \phi\left(\left(\widehat{m}^k\right)_k\right),\ \hat{m}^k = \int_{\mathbb{T}^d} e^{i2\pi k\cdot x} dm(x)$$

- Derivative $\delta\phi/\delta m$ is the same as derivative with respect to Fourier coefficients:

$$\partial_{\hat{m}^k/\overline{\hat{m}^k}}\phi = \left(\widehat{\delta\phi/\delta m}\right)^{-k}$$

- Good point because spaces generated by finite number of Fourier modes is stable by the heat equation which is the characteristic equation of the operator

$$\left(\phi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}\right) \mapsto \int_{\mathbb{T}^d} \mathrm{Tr}\left(\partial_y \partial_\mu \phi(m)(y)\right) dm(y)$$

Rademacher's theorem

- We can find $\mathbb{P}$ a probability measure with full support such any Lipschitz function on $\mathcal{P}(\mathbb{T}^d)$ with respect to the total variation distance is differentiable a.e. in the directions $\hat{m}^k/\overline{\hat{m}^k}$ for $k \in \mathbb{N}^d \setminus \{0\}$.

Reformulation of HJ

- Value function is Lipschitz, so rewrite HJ as

$$\partial_t \mathcal{U}(t,m) - \sum_k 4\pi^2 |k|^2 \widehat{m}^k \partial_{\widehat{m}^k}\mathcal{U}(t,m) - 2\pi^2 \int \Big|\sum_k k \partial_{\widehat{m}^k}\mathcal{U}(t,m) e^{i2\pi k\cdot y}\Big|^2 dm(y) + \mathcal{F}(m) = 0$$

- Formally, uniqueness is to expand $[\mathcal{U}_1 - \mathcal{U}_2](t,\mu_t)$ along

$$\partial_t \mu_t(x) + \mathrm{div}_x\Big[\Big(\partial_\mu \phi\Big)(t,\mu_t)(x)\mu_t(x)\Big] - \Delta_x \mu_t(x) = 0$$

Use semi-concavity

$$\mathcal{U}(t,\mathcal{L}(X+Y)) + \mathcal{U}(t,\mathcal{L}(X-Y)) - 2\mathcal{U}(t,\mathcal{L}(X)) \leq C\mathbb{E}\big[|Y|^2\big]$$

- If the initial value of μ_0 is random with absolute law w.r.t. to $\mathbb{P}$, then μ_t also has absolute law w.r.t. to $\mathbb{P}$: does not see the singular points of $\mathcal{V}_1 - \mathcal{V}_2$