

# Stochastic maximum principle for optimal control of non exchangeable mean field systems.

Samy Mekkaoui

CMAP, École Polytechnique and LPSM

Advances in Financial Mathematics, Paris

*Joint works with*

**Idris Kharroubi** (LPSM, Sorbonne Université)  
**Huyên Pham** (CMAP, École Polytechnique)  
**Xavier Warin** (EDF)

27 January 2026

# Introduction

## Motivation for non exchangeable mean field systems

### Motivations for non exchangeable mean field systems

- MFC theory : Interactions through symmetric particles and homogeneous interactions, through empirical measure  $\frac{1}{N} \sum_{j=1}^N \delta_{X_t^{j,N}}$ .
- NE-MFC : Particle  $i \in \llbracket 1, N \rrbracket$  interacts through  $\frac{\sum_{j=1}^N \xi^{i,j} \delta_{X_t^{j,N}}}{\sum_{j=1}^N \xi^{i,j}}$  where  $(\xi^{i,j})_{1 \leq j \leq N}$  refers to interactions weights between  $i$  and  $j$  assuming no isolated particle, i.e.  $\sum_{j=1}^N \xi^{i,j} > 0$ .  
→ Graphon case :  $\xi^{i,j} = G(\frac{i}{N}, \frac{j}{N})$ .
- Taking heuristically the limit as  $N \nearrow \infty$ , agent labeled by  $u \in I := [0, 1]$  interacts through weighted probability measure

$$I \ni u \mapsto \frac{\int_I G(u, v) \mathbb{P}_{X_t^v}(dx) dv}{\int_I G(u, v) dv} \in \mathcal{P}(\mathbb{R}^d), \quad 0 \leq t \leq T, \quad u \in I,$$

and dynamics of the controlled state system

$$\begin{cases} dX_t^u &= b(u, X_t^u, \alpha_t^u, \frac{\int_I G(u, v) \mathbb{P}_{X_t^v}(dx) dv}{\int_I G(u, v) dv}) dt + \sigma(u, X_t^u, \alpha_t^u, \frac{\int_I G(u, v) \mathbb{P}_{X_t^v}(dx) dv}{\int_I G(u, v) dv}) dW_t^u \\ X_0^u &= \xi^u. \end{cases}$$

# Introduction

## Motivation for non exchangeable mean field systems

### Motivations for non exchangeable mean field systems

- MFC theory : Interactions through symmetric particles and homogeneous interactions, through empirical measure  $\frac{1}{N} \sum_{j=1}^N \delta_{X_t^{j,N}}$ .
- NE-MFC : Particle  $i \in \llbracket 1, N \rrbracket$  interacts through  $\frac{\sum_{j=1}^N \xi^{i,j} \delta_{X_t^{j,N}}}{\sum_{j=1}^N \xi^{i,j}}$  where  $(\xi^{i,j})_{1 \leq j \leq N}$  refers to interactions weights between  $i$  and  $j$  assuming no isolated particle, i.e.  $\sum_{j=1}^N \xi^{i,j} > 0$ .  
→ Graphon case :  $\xi^{i,j} = G(\frac{i}{N}, \frac{j}{N})$ .
- Taking heuristically the limit as  $N \nearrow \infty$ , agent labeled by  $u \in I := [0, 1]$  interacts through weighted probability measure

$$I \ni u \mapsto \frac{\int_I G(u, v) \mathbb{P}_{X_t^v}(dx) dv}{\int_I G(u, v) dv} \in \mathcal{P}(\mathbb{R}^d), \quad 0 \leq t \leq T, \quad u \in I,$$

and dynamics of the controlled state system

$$\begin{cases} dX_t^u &= b(u, X_t^u, \alpha_t^u, \frac{\int_I G(u, v) \mathbb{P}_{X_t^v}(dx) dv}{\int_I G(u, v) dv}) dt + \sigma(u, X_t^u, \alpha_t^u, \frac{\int_I G(u, v) \mathbb{P}_{X_t^v}(dx) dv}{\int_I G(u, v) dv}) dW_t^u \\ X_0^u &= \xi^u. \end{cases}$$

# Introduction

Mean-field approach to large population stochastic control : A strong formulation

## A strong formulation

**Dynamics of the controlled state processes:**

$$\begin{cases} dX_t^u &= b(u, X_t^u, \alpha_t^u, (\mathbb{P}_{X_t^v})_{v \in I}) dt + \sigma(u, X_t^u, \alpha_t^u, (\mathbb{P}_{X_t^v})_{v \in I}) dW_t^u, \\ X_0^u &= \xi^u. \end{cases} \quad 0 \leq t \leq T, u \in I, \quad (1)$$

**Cost Functional** : Aim to minimize over a collection of processes  $\alpha = (\alpha^u)_{u \in I}$  in a suitable class  $\mathcal{A}$

$$J^S(\alpha) = \int_I \mathbb{E} \left[ \int_0^T f(u, X_t^u, \alpha_t^u, (\mathbb{P}_{X_t^v})_{v \in I}) dt + g(u, X_T^u, (\mathbb{P}_{X_T^v})_{v \in I}) \right] du \quad (2)$$

→ Compute  $V_0^S = J^S(\alpha^*)$  where  $\alpha^*$  is a minimizer of  $J^S$ .

- Maps  $(b, \sigma, f, g)$  are defined over the space

$$L^2(I; \mathcal{P}_2(\mathbb{R}^d)) = \{u \rightarrow \mu^u \text{ is measurable and } \int_I \mathcal{W}_2(\mu^u, \delta_0)^2 du < +\infty\}.$$

- Lack of measurability of the map  $(u, \omega) \mapsto X^u(\omega)$  on the space product  $(I \times \Omega, \mathcal{B}(I) \otimes \mathcal{A})$ . Control problem is defined at the level of the law of the processes  $\mathbb{P}_{(X^u, \alpha^u, W^u)}$ .

# Introduction

Mean-field approach to large population stochastic control : A label-state formulation

Since we are working at the level of the laws of the processes and for numerical purposes, we can relax the control problem formulation (1)-(2)

## A label state formulation

**Dynamics of the controlled state processes:**

$$\begin{cases} dX_t &= b(\textcolor{blue}{U}, X_t, \alpha_t, \mathbb{P}_{(U,X_t)})dt + \sigma(\textcolor{blue}{U}, X_t, \alpha_t, \mathbb{P}_{(U,X_t)})dW_t, & 0 \leq t \leq T \\ X_0 &= \xi. \end{cases} \quad (3)$$

**Cost Functional** : Aim to minimize over  $\alpha$  over a suitable class  $\mathcal{A}$

$$J^W(\alpha) = \mathbb{E} \left[ \int_0^T f(\textcolor{blue}{U}, X_t, \alpha_t, \mathbb{P}_{(U,X_t)})dt + g(\textcolor{blue}{U}, X_T, \mathbb{P}_{(U,X_T)}) \right] \quad (4)$$

→ Compute  $V_0^W = J^W(\alpha^*)$  where  $\alpha^*$  is a minimizer of  $J^W$ .

Maps  $(b, \sigma, f, g)$  are defined over the space

$$\mathcal{P}_2^\lambda(I \times \mathbb{R}^d) = \{\mu \in \mathcal{P}_2(I \times \mathbb{R}^d) : \text{pr}_{1\sharp}\mu = \lambda\}.$$

# Introduction

Connection between the two formulations

## Connection between strong and label-state formulation

- We prove  $V_0^S = V_0^W$ . It relies essentially on

$$\mathbb{P}(X_t^u, \alpha_t^u) = \mathbb{P}(X_t, \alpha_t) | U=u, \quad du \text{ a.e}, \quad (5)$$

when given the same policy map  $\hat{\alpha}$ .

- Label-state formulation is more suitable for numerical methods.
- Strong formulation is more suitable for path-wise interpretation.

## Objectives

- Adapt the Pontryagin Maximum Principle to mean field control for non exchangeable mean field systems (NE-MFC) to find necessary and sufficient conditions for an admissible optimal control  $\alpha$ .
- Propose an illustration in the Linear Quadratic (LQ) case with numerical illustrations.

# Analysis tools on $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$

Differentiability and convexity

## Gateaux derivative on $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$

Let  $f : \mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \rightarrow \mathbb{R}$ . For  $U, X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$  such that  $\mathbb{P}_{(U, X)}, \mathbb{P}_{(U, Y)} \in \mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$ , we have

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (f(\mathbb{P}_{(U, X+\epsilon Y)}) - f(\mathbb{P}_{(U, X)})) = \tilde{\mathbb{E}} \left[ \partial_{\tilde{x}} \frac{\delta}{\delta m} f(\mathbb{P}_{(U, X)}) (\tilde{U}, \tilde{X}) \cdot \tilde{Y} \right] \quad (6)$$

where  $(\tilde{U}, \tilde{X}, \tilde{Y})$  is an independent copy of  $(U, X, Y)$  on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ .

→ Such function  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \times I \times \mathbb{R}^d \ni (\mu, \tilde{u}, \tilde{x}) \mapsto \frac{\delta}{\delta m} f(\mu)(\tilde{u}, \tilde{x}) \in \mathbb{R}$  is called linear functional derivative of  $f$ .

# Analysis tools on $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$

Differentiability and convexity

## Gateaux derivative on $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$

Let  $f : \mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \rightarrow \mathbb{R}$ . For  $U, X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$  such that  $\mathbb{P}_{(U, X)}, \mathbb{P}_{(U, Y)} \in \mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$ , we have

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (f(\mathbb{P}_{(U, X+\epsilon Y)}) - f(\mathbb{P}_{(U, X)})) = \tilde{\mathbb{E}} \left[ \partial_{\tilde{x}} \frac{\delta}{\delta m} f(\mathbb{P}_{(U, X)})(\tilde{U}, \tilde{X}) \cdot \tilde{Y} \right] \quad (6)$$

where  $(\tilde{U}, \tilde{X}, \tilde{Y})$  is an independent copy of  $(U, X, Y)$  on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ .

→ Such function  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \times I \times \mathbb{R}^d \ni (\mu, \tilde{u}, \tilde{x}) \mapsto \frac{\delta}{\delta m} f(\mu)(\tilde{u}, \tilde{x}) \in \mathbb{R}$  is called linear functional derivative of  $f$ .

## Convexity on $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$

Let  $f : I \times \mathbb{R}^d \times \mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \rightarrow \mathbb{R}$ .  $f$  is said to be convex if for every  $u \in I$ ,  $x, x' \in \mathbb{R}^d$ ,  $\mu, \mu' \in L^2(\mathcal{P}_2(\mathbb{R}^d))$ , we have :

$$\begin{aligned} f(u, x', \mu') - f(u, x, \mu) &\geq \partial_x f(u, x, \mu). (x' - x) \\ &+ \mathbb{E} \left[ \partial_{\tilde{x}} \frac{\delta}{\delta m} f(u, x, \mu)(U, X). (X' - X) \right]. \end{aligned} \quad (7)$$

where  $(U, X') \sim \mu'$  and  $(U, X) \sim \mu$ .

# The Pontryagin formulation

## Definition of the Hamiltonian map $H$

### Definition of the Hamiltonian $H$

The Hamiltonian  $\mathbb{R}$ -valued function  $H$  of the stochastic optimization problem is defined as :

$$H(\textcolor{blue}{u}, x, \textcolor{red}{\mu}, y, z, a) = b(\textcolor{blue}{u}, x, \textcolor{red}{\mu}, a) \cdot y + \sigma(\textcolor{blue}{u}, x, \textcolor{red}{\mu}, a) : z + f(\textcolor{blue}{u}, x, \textcolor{red}{\mu}, a) \quad (8)$$

where  $(\textcolor{blue}{u}, x, \textcolor{red}{\mu}, y, z, a) \in I \times \mathbb{R}^d \times \mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d \times n} \times A$ .

- Compute an optimality criterion involving the Hamiltonian  $H$  assuming differentiability and convexity as defined previously over the space  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$ .
- In the following,  $A$  will denote a convex subset of  $\mathbb{R}^m$  for  $m \in \mathbb{N}^*$ .

# Probabilistic set-up for non exchangeable mean field SDEs

## Adjoint Equations to $X$

### Adjoint Equations to $X$

We call adjoint processes of  $X$  any pair  $(Y, Z)$  in  $\mathbb{S}^2([0, T]; \mathbb{R}^d) \times \mathbb{H}^2([0, T]; \mathbb{R}^{d \times n})$  such that  $(Y, Z)$  is solution to the adjoint equation

$$\begin{cases} dY_t = -\partial_x H(\mathcal{U}, X_t, \mathbb{P}_{(U, X_t)}, Y_t, Z_t, \alpha_t) dt + Z_t dW_t \\ \quad - \tilde{\mathbb{E}} \left[ \partial_x \frac{\delta}{\delta m} H(\tilde{\mathcal{U}}, \tilde{X}_t, \mathbb{P}_{(U, X_t)}, \tilde{Y}_t, \tilde{Z}_t, \tilde{\alpha}_t)(\mathcal{U}, X_t) \right] dt, \quad t \in [0, T], \\ Y_T = \partial_x g(\mathcal{U}, X_T, \mathbb{P}_{(U, X_T)}) + \tilde{\mathbb{E}} \left[ \partial_x \frac{\delta}{\delta m} g(\tilde{\mathcal{U}}, \tilde{X}_T, \mathbb{P}_{(U, X_T)})(\mathcal{U}, X_T) \right], \end{cases} \quad (9)$$

where  $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{\alpha})$  is an independent copy of  $(X, Y, Z, \alpha)$  defined on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ .

# Probabilistic set-up for non exchangeable mean field SDEs

## Adjoint Equations to $X$

### Adjoint Equations to $X$

We call adjoint processes of  $X$  any pair  $(Y, Z)$  in  $\mathbb{S}^2([0, T]; \mathbb{R}^d) \times \mathbb{H}^2([0, T]; \mathbb{R}^{d \times n})$  such that  $(Y, Z)$  is solution to the adjoint equation

$$\begin{cases} dY_t = -\partial_x H(\textcolor{blue}{U}, X_t, \mathbb{P}_{(U, X_t)}, Y_t, Z_t, \alpha_t) dt + Z_t dW_t \\ \quad - \tilde{\mathbb{E}} \left[ \partial_x \frac{\delta}{\delta m} H(\tilde{U}, \tilde{X}_t, \mathbb{P}_{(U, X_t)}, \tilde{Y}_t, \tilde{Z}_t, \tilde{\alpha}_t)(\textcolor{blue}{U}, X_t) \right] dt, \quad t \in [0, T], \\ Y_T = \partial_x g(\textcolor{blue}{U}, X_T, \mathbb{P}_{(U, X_T)}) + \tilde{\mathbb{E}} \left[ \partial_x \frac{\delta}{\delta m} g(\tilde{U}, \tilde{X}_T, \mathbb{P}_{(U, X_T)})(\textcolor{blue}{U}, X_T) \right], \end{cases} \quad (9)$$

where  $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{\alpha})$  is an independent copy of  $(X, Y, Z, \alpha)$  defined on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ .

→ We retrieve the adjoint equations of the standard Pontryagin formulation but here with the addition of the label randomization  $\textcolor{blue}{U}$ , i.e via  $\mathbb{P}_{(U, X_t)}$  and extension of Lions's derivative over  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$ .

# Derivation of a Pontryagin Optimality Condition

A necessary condition

We now state the main results which are obtained under some regularity assumptions on  $b$ ,  $\sigma$ ,  $f$  and  $g$ .

## Gâteaux derivative of $J$

For  $\beta \in \mathcal{A}$  such that  $\alpha + \epsilon\beta \in \mathcal{A}$  for  $\epsilon > 0$  small enough, we have :

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (J^W(\alpha + \epsilon\beta) - J^W(\alpha)) = \mathbb{E} \left[ \int_0^T \left( \partial_\alpha H(\textcolor{blue}{U}, X_t, \mathbb{P}_{(\textcolor{red}{U}, X_t)}, \textcolor{yellow}{Y}_t, \textcolor{orange}{Z}_t, \alpha_t) \cdot \beta_t \right) dt \right]$$

where  $X$  is given by (3),  $(Y, Z)$  are given by (9) and the Hamiltonian function  $H$  is given by (8).

## Necessary condition for optimality of $\alpha$

Moreover, if we assume that  $H$  is convex in  $a \in A$ , that  $\alpha = (\alpha_t)_{0 \leq t \leq T}$  is optimal, that  $X = (X_t)_{0 \leq t \leq T}$  is the associated optimal control state given by (3) and that  $(Y, Z) = (Y_t, Z_t)_{0 \leq t \leq T}$  are the associated adjoint processes solving (9), then we have :

$$\forall a \in A, \quad H(\textcolor{blue}{U}, X_t, \mathbb{P}_{(\textcolor{red}{U}, X_t)}, \textcolor{yellow}{Y}_t, \textcolor{orange}{Z}_t, \alpha_t) \leq H(\textcolor{blue}{U}, X_t, \mathbb{P}_{(\textcolor{red}{U}, X_t)}, \textcolor{yellow}{Y}_t, \textcolor{orange}{Z}_t, a) \quad dt \otimes d\mathbb{P} \text{ a.e } (10)$$

# Sufficient condition for optimality of $\alpha$

A sufficient condition

## Sufficient condition for optimality of $\alpha$

Let  $\alpha \in \mathcal{A}$ ,  $X = (X_t)_{0 \leq t \leq T}$  the corresponding controlled state process and  $(Y, Z) = (Y_t, Z_t)_{0 \leq t \leq T}$  the corresponding adjoint processes.

- (1)  $\mathbb{R}^d \times \mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \ni (x, \mu) \rightarrow g(\textcolor{blue}{U}, x, \mu)$  is convex  $d\mathbb{P}$  a.e
- (2)  $\mathbb{R}^d \times \mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \times A \ni (x, \mu, a) \rightarrow H(\textcolor{blue}{U}, x, \mu, \textcolor{orange}{Y}_t, \textcolor{brown}{Z}_t, a)$  is convex  $dt \otimes d\mathbb{P}$  a.e

If we assume also following the necessary condition for optimality :

$$H(\textcolor{blue}{U}, X_t, \mathbb{P}_{(U, X_t)}, \textcolor{orange}{Y}_t, \textcolor{brown}{Z}_t, \alpha_t) = \inf_{\beta \in A} H(\textcolor{blue}{U}, X_t, \mathbb{P}_{(U, X_t)}, \textcolor{orange}{Y}_t, \textcolor{brown}{Z}_t, \beta), \quad dt \otimes d\mathbb{P} \text{ a.e}$$

Then,  $\alpha$  is an optimal control in the sense that  $J(\alpha) = \inf_{\alpha' \in \mathcal{A}} J(\alpha')$

# Linear quadratic control problem

The non exchangeable LQ model

## Linear quadratic optimal control problem

We consider the following class of models (assuming for sake of simplicity  $\sigma$  constant and  $A = \mathbb{R}^m$ ).

$$\begin{cases} dX_t &= \left[ \beta(\mathbf{U}) + A(\mathbf{U})X_t + \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}_{(\mathbf{U}, X_t)}} [G_A(\mathbf{U}, \tilde{U})\tilde{X}_t] + B(\mathbf{U})\alpha_t \right] dt + \gamma(\mathbf{U})dW_t, t \in [0, T] \\ &= \left[ \beta(\mathbf{U}) + A(\mathbf{U})X_t + \int_{I \times \mathbb{R}^d} [G_A(\mathbf{U}, v)x] \mathbb{P}_{(\mathbf{U}, X_t)}(dv, dx) + B(\mathbf{U})\alpha_t \right] + \gamma(\mathbf{U})dW_t, t \in [0, T] \\ X_0 &= \xi, \end{cases}$$

where  $\beta \in L^\infty(I; \mathbb{R}^d)$ ,  $\gamma \in L^\infty(I; \mathbb{R}^d)$ ,  $A \in L^\infty(I; \mathbb{R}^{d \times d})$ ,  $B \in L^\infty(I; \mathbb{R}^{d \times m})$  and  $G_A \in L^2(I \times I; \mathbb{R}^{d \times d})$ .

# Linear quadratic control problem

The non exchangeable LQ model

## Linear quadratic optimal control problem

We consider the following class of models (assuming for sake of simplicity  $\sigma$  constant and  $A = \mathbb{R}^m$ ).

$$\begin{cases} dX_t &= \left[ \beta(\mathbf{U}) + A(\mathbf{U})X_t + \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}_{(\mathbf{U}, X_t)}} [G_A(\mathbf{U}, \tilde{U})\tilde{X}_t] + B(\mathbf{U})\alpha_t \right] dt + \gamma(\mathbf{U})dW_t, t \in [0, T] \\ &= \left[ \beta(\mathbf{U}) + A(\mathbf{U})X_t + \int_{I \times \mathbb{R}^d} [G_A(\mathbf{U}, v)x] \mathbb{P}_{(\mathbf{U}, X_t)}(dv, dx) + B(\mathbf{U})\alpha_t \right] + \gamma(\mathbf{U})dW_t, t \in [0, T] \\ X_0 &= \xi, \end{cases}$$

where  $\beta \in L^\infty(I; \mathbb{R}^d)$ ,  $\gamma \in L^\infty(I; \mathbb{R}^d)$ ,  $A \in L^\infty(I; \mathbb{R}^{d \times d})$ ,  $B \in L^\infty(I; \mathbb{R}^{d \times m})$  and  $G_A \in L^2(I \times I; \mathbb{R}^{d \times d})$ .

## Quadratic cost functional

$$\begin{aligned} J(\boldsymbol{\alpha}) &= \mathbb{E} \left[ \int_0^T \left[ Q(\mathbf{U})(X_t - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}_{(\mathbf{U}, X_t)}} [\tilde{G}_Q(\mathbf{U}, \tilde{U})\tilde{X}_t]) \cdot (X_t - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}_{(\mathbf{U}, X_t)}} [\tilde{G}_Q(\mathbf{U}, \tilde{U})\tilde{X}_t]) \right. \right. \\ &\quad + \alpha_t \cdot R(\mathbf{U})\alpha_t + 2\alpha_t \cdot \Gamma(\mathbf{U})X_t + 2\alpha_t \cdot \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}_{(\mathbf{U}, X_t)}} [G_I(\mathbf{U}, \tilde{U})\tilde{X}_t] \Big] dt \\ &\quad \left. \left. + H(\mathbf{U})(X_T - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_T) \sim \mathbb{P}_{(\mathbf{U}, X_T)}} [\tilde{G}_H(\mathbf{U}, \tilde{U})\tilde{X}_T]) \cdot (X_T - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_T) \sim \mathbb{P}_{(\mathbf{U}, X_T)}} [\tilde{G}_H(\mathbf{U}, \tilde{U})\tilde{X}_T]) \right] \right] \quad (11) \end{aligned}$$

# A linear quadratic model

## Characterization of optimal control

### Proposition : Optimal control in the LQ case

In the Linear quadratic model, the unique optimal control  $\hat{\alpha} = (\hat{\alpha}_t)_{0 \leq t \leq T}$  is given by

$$\hat{\alpha}_t = S_t(U)X_t + \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}(U, X_t)} [\bar{S}_t(U, \tilde{U})\tilde{X}_t] + \Gamma_t(U), \quad 0 \leq t \leq T, \quad (12)$$

where  $X = (X_t)_{0 \leq t \leq T}$  is the unique solution to the SDE obtained after replacing  $\hat{\alpha}_t$  by (12) and where we denoted

$$\begin{cases} S_t(U) = -(R(U))^{-1}((B(U))^\top K_t(U) + \Gamma(U)), \\ \bar{S}_t(U, \tilde{U}) = -(R(U))^{-1}((B(U))^\top \bar{K}_t(U, \tilde{U}) + G_I(U, \tilde{U})) \\ \Gamma_t(U) = -(R(U))^{-1}(B(U))^\top \Lambda_t(U) \end{cases}$$

where  $K \in C^1([0, T]; L^\infty(I; \mathbb{S}_+^d))$ ,  $\bar{K} \in C^1([0, T], L^2(I \times I; \mathbb{R}^{d \times d}))$  and  $\Lambda \in C^1([0, T]; L^2(I; \mathbb{R}^d))$  are to be determined through infinite dimensional Riccati equations.

# A linear quadratic model

Systemic risk model: Extension of Carmona, Fouque, Sun (2015)

## Systemic risk model with heterogeneous banks

Dynamics of the controlled state processes:

$$\begin{cases} dX_t &= \left[ \kappa(X_t - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}(\mathbf{U}, \mathbf{X}_t)} [G_\kappa(\mathbf{U}, \tilde{U}) \tilde{X}_t] + \alpha_t \right] dt + \sigma dW_t, \quad 0 \leq t \leq T, \\ X_0 &= \xi. \end{cases}$$

Cost Functional :

$$\begin{aligned} J(\alpha) = \mathbb{E} \Bigg[ & \int_0^T \left\{ \eta(\mathbf{U}) \left( X_t - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}(\mathbf{U}, \mathbf{X}_t)} [G_\eta(\mathbf{U}, \tilde{U}) \tilde{X}_t] \right)^2 \right. \\ & + q(\mathbf{U}) \alpha_t \left( X_t - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_t) \sim \mathbb{P}(\mathbf{U}, \mathbf{X}_t)} [G_q(\mathbf{U}, \tilde{U}) \tilde{X}_t] \right) + \alpha_t^2 \Big\} dt \\ & \left. + r(\mathbf{U}) \left( X_T - \tilde{\mathbb{E}}_{(\tilde{U}, \tilde{X}_T) \sim \mathbb{P}(\mathbf{U}, \mathbf{X}_T)} [G_r(\mathbf{U}, \tilde{U}) \tilde{X}_T] \right)^2 \right]. \end{aligned} \tag{13}$$

# Numerical methods for learning the optimal control

## Numerical methods for learning the optimal feedback control map $\hat{a}$ .

### Deep Graphon :

- Direct parametrization of the control via a Neural Network in [feedback form](#) in view of (12) solved by standard gradient descent algorithm.
- To learn functions defined over  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \rightarrow$  [conditional moment neural network](#) where we approximate  $\mu \in \mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$  by its conditional moments.

# Numerical methods for learning the optimal control

## Numerical methods for learning the optimal feedback control map $\hat{a}$ .

### Deep Graphon :

- Direct parametrization of the control via a Neural Network in **feedback form** in view of (12) solved by standard gradient descent algorithm.
- To learn functions defined over  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \rightarrow$  **conditional moment neural network** where we approximate  $\mu \in \mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$  by its conditional moments.

### Deep BSDE Graphon :

- Optimal control learnt in view of (10) since

$$\hat{a}_t = a(U, t, X_t, \mathbb{P}_{(U, X_t)}, Y_t) = \arg \min_{a \in A} H(U, X_t, \mathbb{P}_{(U, X_t)}, Y_t, a).$$

→ Learn  $(X, Y)$  by exploiting the FBSDE equation.

- We use 2 neural networks  $\mathcal{U}_\theta(\mu)(u, x)$  and  $\mathcal{Z}_\theta(t, \mu)(u, x)$  to approximate initial value of  $Y$  and the  $Z$  component and we minimiser over  $\theta$  the cost functional

$$\theta \mapsto L(\theta) = \mathbb{E} \left[ |Y_T^\theta - G(X_T^\theta, \mathbb{P}_{(U, X_T^\theta)})|^2 \right], \quad (14)$$

Starting from  $\mathcal{U}_\theta(\mathbb{P}_{(U, X_0)})(U, X_0)$ , we diffuse

$$\begin{cases} X_{t_{i+1}}^\theta &= X_{t_i}^\theta + B(U, X_{t_{i+1}}^\theta, Y_{t_{i+1}}^\theta, \mathbb{P}_{(U, X_{t_{i+1}}^\theta)})dt + \sigma \Delta W_{t_{i+1}} \\ Y_{t_{i+1}}^\theta &= Y_{t_i}^\theta + H(U, X_{t_{i+1}}^\theta, Y_{t_{i+1}}^\theta, \mathcal{Z}_\theta(t_{i+1}, \mathbb{P}_{(U, X_{t_{i+1}}^\theta)})(U, X_{t_{i+1}}^\theta), \mathbb{P}_{(U, X_{t_{i+1}}^\theta)}, \mathbb{P}_{(U, Y_{t_{i+1}}^\theta)})\Delta t + \mathcal{Z}_\theta(t_{i+1}, \mathbb{P}_{(U, X_{t_{i+1}}^\theta)})(U, X_{t_{i+1}}^\theta) \end{cases} \quad (15)$$

for certains maps  $B, H$  depending on model coefficients.

# Numerical methods for learning the optimal control

## Numerical methods for learning the optimal feedback control map $\hat{a}$ .

### Deep Graphon :

- Direct parametrization of the control via a Neural Network in **feedback form** in view of (12) solved by standard gradient descent algorithm.
- To learn functions defined over  $\mathcal{P}_2^\lambda(I \times \mathbb{R}^d) \rightarrow$  **conditional moment neural network** where we approximate  $\mu \in \mathcal{P}_2^\lambda(I \times \mathbb{R}^d)$  by its conditional moments.

### Deep BSDE Graphon :

- Optimal control learnt in view of (10) since

$$\hat{a}_t = a(U, t, X_t, \mathbb{P}_{(U, X_t)}, Y_t) = \arg \min_{a \in A} H(U, X_t, \mathbb{P}_{(U, X_t)}, Y_t, a).$$

→ Learn  $(X, Y)$  by exploiting the FBSDE equation.

- We use 2 neural networks  $\mathcal{U}_\theta(\mu)(u, x)$  and  $\mathcal{Z}_\theta(t, \mu)(u, x)$  to approximate initial value of  $Y$  and the  $Z$  component and we minimiser over  $\theta$  the cost functional

$$\theta \mapsto L(\theta) = \mathbb{E} \left[ |Y_T^\theta - G(X_T^\theta, \mathbb{P}_{(U, X_T^\theta)})|^2 \right], \quad (14)$$

Starting from  $\mathcal{U}_\theta(\mathbb{P}_{(U, X_0)})(U, X_0)$ , we diffuse

$$\begin{cases} X_{t_{i+1}}^\theta &= X_{t_i}^\theta + B(U, X_{t_{i+1}}^\theta, Y_{t_{i+1}}^\theta, \mathbb{P}_{(U, X_{t_{i+1}}^\theta)})dt + \sigma \Delta W_{t_{i+1}} \\ Y_{t_{i+1}}^\theta &= Y_{t_i}^\theta + H(U, X_{t_{i+1}}^\theta, Y_{t_{i+1}}^\theta, \mathcal{Z}_\theta(t_{i+1}, \mathbb{P}_{(U, X_{t_{i+1}}^\theta)})(U, X_{t_{i+1}}^\theta), \mathbb{P}_{(U, X_{t_{i+1}}^\theta)}, \mathbb{P}_{(U, Y_{t_{i+1}}^\theta)})\Delta t + \mathcal{Z}_\theta(t_{i+1}, \mathbb{P}_{(U, X_{t_{i+1}}^\theta)})(U, X_{t_{i+1}}^\theta) \end{cases} \quad (15)$$

for certains maps  $B, H$  depending on model coefficients.

# A linear quadratic model

Numerical experiments

**Application with  $\sigma = 1$ ,  $\eta = 0.73$ ,  $q = 0.8$ ,  $r = 0.22$  and  $\kappa = 0.62$  :**

Method	Riccati	Deep Graphon	Deep BSDE Graphon
Value	0.58830	0.58826	0.58820

Table: Expected cost function using  $M = 10000$  in simulation with  $G(u, v) = e^{-uv}$ .

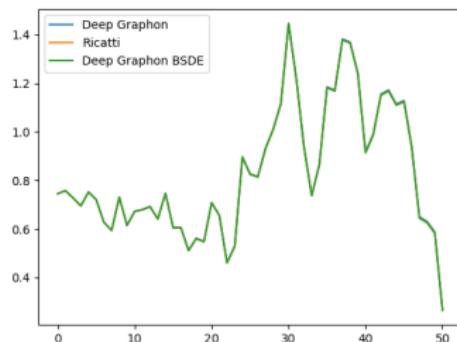


Figure: Optimal trajectory of  $X$  with  $u = 0.708$

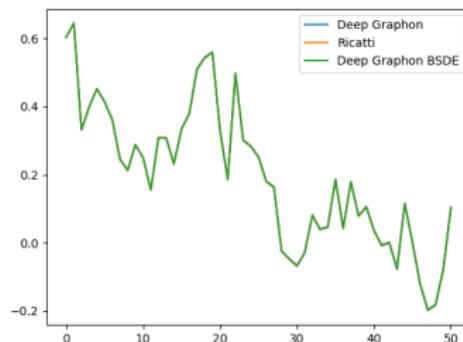


Figure: Optimal trajectory of  $X$  with  $u = 0.599$

Figure: Comparison between NN solvers and the Riccati one with  $G(u, v) = e^{-uv}$

# Conclusion

Some concluding remarks

## Summary of the talk

- We introduced a general class of non exchangeable mean field systems.
- We present an application of our framework to the case of LQ optimal control problem where we exhibit a new infinite-dimensional system of Riccati equations and we show numerically how to solve the optimal control problem through Deep learning algorithms.

# Conclusion

Some concluding remarks

## Summary of the talk

- We introduced a general class of **non exchangeable mean field systems**.
- We present an application of our framework to the case of **LQ optimal control problem** where we exhibit a new **infinite-dimensional** system of Riccati equations and we show numerically how to solve the optimal control problem through Deep learning algorithms.

## Future Works

- In the present setting, agents interact through a specified **graph/graphon** structure but it could be interesting to add a control perspective on the agent's interactions.
- Adding some randomness in the graph structure would lead to the study of dynamical systems with random interactions  $\implies$  Bridge with **Random Matrix Theory** and **Operator-Theory**.

## References

-  I. Kharroubi, S.Mekkaoui and H.Pham. *Stochastic maximum principle for optimal control problem of non exchangeable mean field systems*, arXiv preprint arXiv:2506.05595
-  F.de Feo and S.Mekkaoui. *Optimal control of heterogeneous mean-field stochastic differential equations with common noise and applications to financial models*, arXiv preprint arXiv:2511.18636
-  S. Mekkaoui, H.Pham and X.Warin *Learning mappings on labeled conditional distributions*, Work in Progress
-  S.Mekkaoui, H.Pham *Analysis of Non-Exchangeable Mean Field Markov Decision Processes with common noise : From Bellman equation to quantitative propagation of chaos*. Work in Progress

THANK YOU FOR YOUR ATTENTION