Mean-field control and mean-field type games of BSDEs

Alexander Aurell [1]

Based on joint work with Boualem Djehiche

[1] Department of Mathematics, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden (aaurell@kth.se)

Mean-field games?

Mean-field type games?

Mean-field control theory?

Backward stochastic differential equations?

Introduction: deterministic optimal control (finite horizon):

Minimize or maximize

$$J(u(\cdot)) = \int_0^T f(x(t), u(t))dt + h(x(T)) \tag{1}$$

with respect to $u : [0, T] \to U$, subject to

$$\begin{cases} \dot{x}(t) = b(x(t), u(t)), & 0 < t \leq T, \\ x(0) = x_0, \end{cases} \tag{2}$$

where U is a given set of control values.

Introduction: stochastic optimal control (finite horizon)[1] :

Minimize or maximize

$$J(u_\cdot) = E\left[\int_0^T f(X_t, u_t)dt + h(X_T)\right], \tag{3}$$

with respect to $u : [0, T] \to U$, subject to

$$\begin{cases} dX_t = b(X_t, u_t)dt + \sigma(X_t, u_t)dW_t, \quad 0 < t \leq T, \\ X_0 = x_0. \end{cases} \tag{4}$$

[1] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations.* Vol. 43. Springer Science & Business Media, 1999.

Introduction: mean-field optimal control (finite horizon)[1]:

Minimize or maximize

$$J(u_\cdot) = E\left[\int_0^T f(X_t, E[X_t], u_t)dt + h(X_T, E[X_T])\right], \tag{5}$$

with respect to $u : [0, T] \to U$, subject to

$$\begin{cases} dX_t = b(X_t, E[X_t], u_t)dt + \sigma(X_t, E[X_t], u_t)dW_t, & 0 < t \leq T, \\ X_0 = x_0. \end{cases} \tag{6}$$

"Control of SDEs of mean-field type"
"Control of McKean-Vlasov equations"

Example (non-linear in expectation):

$$\begin{aligned} J(u_\cdot) &= Var(X_T) \\ &= E\left[X_T^2 - E[X_T]^2\right] \end{aligned} \tag{7}$$

[1] Daniel Andersson and Boualem Djehiche. "A maximum principle for SDEs of mean-field type". In: *Applied Mathematics & Optimization* 63.3 (2011), pp. 341–356.

Introduction

Optimal control theory tries to answer two questions:

- Existence of a minimum/maximum of the performance functional J.
- Explicit computation of such a minimum/maximum.
    - The Bellman principle, which yields the Hamilton-Jacobi-Bellman equation (HJB) for the value function.
    - Pontryagin's maximum principle which yields the Hamiltonian system for "the derivative" of the value function.

Consider N agents with state dynamics

$$\begin{cases} dX_t^i = b(X_t^i, \mu_t^N, u_t^i)dt + \sigma(X_t^i, \mu_t^N, u_t^i)dW_t^i, \quad 0 < t \leq T, \\ X_0^i = x_0^i, \end{cases} \tag{8}$$

where $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$, cooperating to minimize/maximize

$$J^{i,N}(u_\cdot^1, \dots, u_\cdot^N) = \frac{1}{N}\sum_{i=1}^N E\left[\int_0^T f(X_t^i, \mu_t^N, u_t^i)dt + h(X_T^i, \mu_T^N)\right]. \tag{9}$$

Under some conditions...

- The control found by solving the mean-field optimal control problem (previous slide) approximates the solution to (8)-(9) .
- Results exists on the commutation of optimization and limit taking.[1]

[1] Daniel Lacker. "Limit Theory for Controlled McKean–Vlasov Dynamics". In: *SIAM Journal on Control and Optimization* 55.3 (2017), pp. 1641–1672.

Instead of cooperating, let the N agents compete. Given the control chosen by all other agents, $u_\cdot^{-i}$, agent i wants to minimize

$$j^{i,N}(u_\cdot^i; u_\cdot^{-i}) = E\left[\int_0^T f(X_t^i, \mu_t^N, u_t^i)dt + h(X_T^i, \mu_T^N)\right]. \tag{10}$$

A Nash equilibrium $(\hat{u}_\cdot^1, \ldots, \hat{u}_\cdot^N)$ for this differential game is given by

$$j^{i,N}(u_\cdot; \hat{u}_\cdot^{-i}) \geq j^{i,N}(\hat{u}_\cdot^i; \hat{u}_\cdot^{-i}), \quad \forall u_\cdot, \ \forall i = 1, \ldots, N. \tag{11}$$

A Nash equilibrium can be approximated by a fixed point scheme

i) Fix a deterministic function $\mu_t : [0, T] \to \mathcal{P}_2(\mathbb{R}^d)$.

ii) Solve the stochastic control problem (single agent!):

$$\hat{u}_\cdot = \underset{u_\cdot}{\operatorname{argmin}}\ E\left[\int_0^T f(X_t, \mu_t, u_t)dt + h(X_T, \mu_T)\right] \tag{12}$$

iii) Determine the function $\hat{\mu}_t : [0, T] \to \mathcal{P}_2(\mathbb{R}^d)$ such that $\hat{\mu}_t = \mathbb{P} \circ (\hat{X}_t)^{-1}$ for all $t \in [0, T]$, $\hat{X}_\cdot$ being the dynamic corresponding to $\hat{u}_\cdot$.

This matching problem (often in PDE form) is called a "Mean-Field Game".[1][2]

[1] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle". In: *Communications in Information & Systems* 6.3 (2006), pp. 221–252.

[2] Jean-Michel Lasry and Pierre-Louis Lions. "Mean field games". In: *Japanese journal of mathematics* 2.1 (2007), pp. 229–260.

Let there be N agents with dynamics

$$\begin{cases} dX_t^i = b^i(X_t^1, \mathbb{P}\circ(X_t^1)^{-1}, u_t^1, \dots, X_t^i, \mathbb{P}\circ(X_t^i)^{-1}, u_t^i, \dots, u_t^N)dt \\ \qquad\qquad +\sigma^i(\dots, X_t^i, \mathbb{P}\circ(X_t^i)^{-1}, u_t^i, \dots)dW_t^i, \\ X_0^i = x_0^i, \end{cases} \tag{13}$$

Agent i replies to the other agents choice of control $u_\cdot^{-i}$ by minimizing its *best reply* functional

$$J^i(u_\cdot^i; u_\cdot^{-i}) = E\Big[\int_0^T f^i(\dots, X_t^i, \mathbb{P}\circ(X_t^i)^{-1}, u_t^i, \dots)dt \\ + h^i(\dots, X_T^i, \mathbb{P}\circ(X_T^i)^{-1}, \dots)\Big]. \tag{14}$$

Players not identical (exchangeable) anymore!

A mean-field type game consists of *major players*, that can influence their distributions, and asks: what is the equilibrium behavior of these agents?

A variation in control gives a variation in the marginal distribution, and thus we must be able to handle variation of measure-valued functions.

Underlying probability space is *rich enough, so that for every* $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, *there exists a square-integrable random variable* X *whose distribution is* μ.

Consider $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$. We can write $f(\mu) =: F(X)$ and differentiate F is Frechét sense, whenever there exists a linear functional $DF[X] : L^2(\mathcal{F}; \mathbb{R}^d) \to \mathbb{R}$ such that

$$F(X+Y) - F(X) = \langle DF[X], Y \rangle + o(\|Y\|_2). \tag{15}$$

By Riesz' representation theorem, $DF[X]$ is unique and there exists a Borel function $\phi[\mu] : \mathbb{R}^d \to \mathbb{R}$ such that $\phi[\mu](X) = DF[X]$, therefore[1]

$$f(\mu') - f(\mu) = E\left[\phi[\mathbb{P} \circ (X)^{-1}](X)(X' - X)\right] + o(\|X' - X\|_2). \tag{16}$$

Denote $\partial_\mu f(\mu; x) := \phi[\mu](x)$, and we have the identity

$$DF[X] = \partial_\mu f(\mathbb{P} \circ (X)^{-1}; X) =: \partial_\mu f(\mathbb{P} \circ (X)^{-1}). \tag{17}$$

[1] Rainer Buckdahn, Juan Li, and Jin Ma. "A stochastic maximum principle for general mean-field systems". In: *Applied Mathematics & Optimization* 74.3 (2016), pp. 507–534.

Example: If $f(\mu) = \left(\int_{\mathbb{R}^d} x d\mu(x)\right)^2$ *then*

$$E[X + tY]^2 - E[X]^2 = E[2E[X]Y] + o(t) \tag{18}$$

and $\partial_\mu f(\mu) = 2\int_{\mathbb{R}^d} x d\mu(x)$.

If f takes another argument, ξ, then (with $\mu = \mathbb{P} \circ (X)^{-1}$)

$$f(\xi, \mu') - f(\xi, \mu) = E\left[\partial_\mu f(\widetilde{\xi}, \mu; X)(X' - X)\right] + o(\|X' - X\|_2), \tag{19}$$

where the expectation is **not taken over** $\widetilde{\xi}$. To shorten notation,

$$E\left[\partial_\mu f(\widetilde{\xi}, \mu; X)(X' - X)\right] =: E\left[(\partial_\mu f(\xi, \mu))^*(X' - X)\right]. \tag{20}$$

For expectations over "the other arguments", we write

$$\widetilde{E}\left[\partial_\mu f(\widetilde{\xi}, \mu; X)\right] =: E\left[{}^*(\partial_\mu f(\xi, \mu))\right] \tag{21}$$

Deterministic setting: reverse time to get control problem with state constraint at $t = T$. **Stochastic setting:** time reversal destroys adaptedness!

Given filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$, any $x_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^D)$ induces an $\mathbb{F}$-martingale

$$X_t := E[x_T \mid \mathcal{F}_t]. \tag{22}$$

If $\mathbb{F}$ is generated by a Wiener process $W_\cdot$, the martingale representation theorem then gives existence of a unique square-integrable process Z_t such that

$$X_t = x_T + \int_t^T Z_t dW_t. \tag{23}$$

$Z_\cdot$ works as a projection and makes $X_\cdot$ progressively measurable!

In this fashion, we can construct BSDEs with general drift.[1] Given a suitable driver-terminal condition pair (f, x_T), $(X_\cdot, Z_\cdot)$ solves the BSDE

$$dX_t = fdt + Z_t dW_t, \quad X_T = x_T \tag{24}$$

if (together with some regularity)

$$X_t = x_T - \int_t^T fdt - \int_t^T Z_s dW_s. \tag{25}$$

[1] Jianfeng Zhang. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory.* Vol. 86. Springer, 2017.

Introduction: control/games of mean-field BSDEs

Two problems to be presented in this talk:

1. A mean-field type control based model for pedestrian motion, where the state dynamics is a BSDE:

$$\begin{cases} \text{Find } \hat{u}_\cdot \text{ such that } J(u_\cdot) \geq J(\hat{u}_\cdot),\ \forall u_\cdot \in \mathcal{U}, \\ \text{Given a control, the state } X_\cdot \text{ satisfies a mean-field BSDE.} \end{cases}$$

2. A mean-field type game between two players whose state dynamics are BSDEs:

$$\begin{cases} \text{Find } (\hat{u}^1_\cdot, \hat{u}^2_\cdot) \text{ such that } J^i(u_\cdot; \hat{u}^{-i}_\cdot) \geq J^i(\hat{u}^i_\cdot; \hat{u}^{-i}_\cdot),\ \forall u_\cdot \in \mathcal{U}^i,\ i = 1, 2, \\ \text{Given controls, the state } X^i_\cdot \text{ satisfies a mean-field BSDE.} \end{cases}$$

Pedestrian crowd motion: quick facts

Empirical studies of human crowds have been conducted since the ’50s[1].

Basic guidelines for pedestrian behavior: will to reach specific targets, repulsion from other individuals and deterministic if the crowd is sparse but partially random if the crowd is dense[2].

Humans motion is decision-based.

Classical particles

- Robust - interaction only through collisions
- Blindness - dynamics ruled by inertia
- Local - interaction is pointwise
- Isotropy - all directions equally influential

"Smart agents"

- Fragile - avoidance of collisions and obstacles
- Vision - dynamics ruled at least partially by decision
- Nonlocal - interaction at a distance
- Anisotropy - some directions more influential than others

[1] BD Hankin and R Wright. “Passenger flow in subways”. In: *Journal of the Operational Research Society* 9.2 (1958), pp. 81–88.

[2] E Cristiani, B Piccoli, and A Tosin. “Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints”. In: *Mathematical modeling of collective behavior in socio-economic and life sciences*. Springer, 2010, pp. 337–364.

Microscopic

D Helbing and P Molnar. "Social force model for pedestrian dynamics". In: *Physical review E* 51.5 (1995), p. 4282
A Schadschneider. "Cellular automaton approach to pedestrian dynamics-theory". In: *Pedestrian and Evacuation Dynamics* (2002), pp. 75–85
S Okazaki. "A study of pedestrian movement in architectural space, part 1: Pedestrian movement by the application on of magnetic models". In: *Trans. AIJ* 283 (1979), pp. 111–119

Macroscopic

LF Henderson. "The statistics of crowd fluids". In: *Nature* 229.5284 (1971), p. 381
R Hughes. "The flow of human crowds". In: *Annual review of fluid mechanics* 35.1 (2003), pp. 169–182
S Hoogendoorn and P Bovy. "Pedestrian route-choice and activity scheduling theory and models". In: *Transportation Research Part B: Methodological* 38.2 (2004), pp. 169–190

Mesoscopic/Kinetic

C Dogbe. "On the modelling of crowd dynamics by generalized kinetic models". In: *Journal of Mathematical Analysis and Applications* 387.2 (2012), pp. 512–532
G Albi et al. "Mean field control hierarchy". In: *Applied Mathematics & Optimization* 76.1 (2017), pp. 93–135

Mean-field games:
a macroscopic approximation of a microscopic model

Mean-field type games/control:
a macroscopic approximation of a microscopic model
or
a distribution dependent microscopic model

Pedestrian crowd modeling: heuristics of the mean-field approach

- The dynamics of a pedestrians is given by
  - *change in position* $=$ *velocity* $+$ *noise*

  The pedestrian controls it's velocity.

- The pedestrian controls it's velocity rationally, it minimizes
  - *Expected cost* $= E\left[\int_0^T f\,(\text{energy use}(t), \text{interaction}(t))\, dt + \text{deviation from final target}\right]$

- The interaction is assumed to depend on an aggregate of distances to other pedestrians:
  - *Lots of pedestrians in my neighborhood - congestion cost*
  - *Seeking the company of others - social gain*

- To evaluate its interaction cost, the pedestrian anticipates the movement of other pedestrians via the distribution of the crowd.

Many possible extensions:
controlled noise, multiple interacting crowds, fast exit times, interaction with the environment, common noise, hard congestion.

Pedestrian crowd motion: mean-field models

Early works

S Hoogendoorn and P Bovy. "Pedestrian route-choice and activity scheduling theory and models". In: *Transportation Research Part B: Methodological* 38.2 (2004), pp. 169–190

C Dogbé. "Modeling crowd dynamics by the mean-field limit approach". In: *Mathematical and Computer Modelling* 52.9-10 (2010), pp. 1506–1520

Aversion and congestion

A Lachapelle and M-T Wolfram. "On a mean field game approach modeling congestion and aversion in pedestrian crowds". In: *Transportation research part B: methodological* 45.10 (2011), pp. 1572–1589

Y Achdou and M Laurière. "Mean field type control with congestion". In: *Applied Mathematics & Optimization* 73.3 (2016), pp. 393–418

Fast exits (evacuation)

M Burger et al. "On a mean field game optimal control approach modeling fast exit scenarios in human crowds". In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE. 2013, pp. 3128–3133

M Burger et al. "Mean field games with nonlinear mobilities in pedestrian dynamics". In: *Discrete and Continuous Dynamical Systems-Series B* (2014)

B Djehiche, A Tcheukam, and H Tembine. "A Mean-Field Game of Evacuation in Multilevel Building". In: *IEEE Transactions on Automatic Control* 62.10 (2017), pp. 5154–5169

Multi-population

E Feleqi. "The derivation of ergodic mean field game equations for several populations of players". In: *Dynamic Games and Applications* 3.4 (2013), pp. 523–536

M Cirant. "Multi-population mean field games systems with Neumann boundary conditions". In: *Journal de Mathématiques Pures et Appliquées* 103.5 (2015), pp. 1294–1315

Y Achdou, M Bardi, and M Cirant. "Mean field games models of segregation". In: *Mathematical Models and Methods in Applied Sciences* 27.01 (2017), pp. 75–113

Pedestrian crowd motion: rationality

Another model categorization: *level of rationality*[1].

Rationality level	Information structure	Area of application
Irrational	-	Panic situations
Basic	Destination and environment	Movement in large unfamiliar environments
Rational	Current position of other pedestrians	Movement in small and well-known environment
Highly rational	Forecast of other pedestrians movement	Movement in small and well-known environment
Optimal	Omnipotent central planner	"Soldiers"

Mean field games can model highly rational pedestrians.

Mean-field optimal control can model optimal pedestrians.

[1] E Cristiani, F Priuli, and A Tosin. "Modeling rationality to control self-organization of crowds: an environmental approach". In: *SIAM Journal on Applied Mathematics* 75.2 (2015), pp. 605–629.

Tagged pedestrian motion: control of mean-field BSDEs

Stochastic dynamics with initial condition cannot model motion that *has to terminate in a target location at time horizon T*, such as:

- Guards moving to a security threat
- Medical personnel moving to a patient
- Fire-fighters moving to a fire
- Deliveries

Control of mean-field BSDEs can be a tool for *centrally planned decision-making for pedestrian groups*, who are forced to reach a target position.

Recall, mean-field control is suitable for pedestrian crowd modeling when

- the central planner is rational and has the ability to anticipate the behaviour of other pedestrians
- aggegate effects are considered

Tagged pedestrian motion: control of mean-field BSDEs

The motion of our representative agent is described by a BSDE,

$$\begin{cases} dX_t = b(t, X_t, \mathbb{P} \circ (X_t)^{-1}, Z_t, u_t)dt + Z_t dW_t, \\ X_T = x_T. \end{cases} \tag{26}$$

The central planner faces the optimization problem

$$\begin{cases} \min_{u.} & E\left[\int_0^T f(t, X_t, \mathbb{P} \circ (X_t)^{-1}, u_t)dt + h(X_0, \mathbb{P} \circ (X_0)^{-1})\right] \\ \text{s.t.} & (X_\cdot, Z_\cdot) \text{ solves (26)}, \\ & u_\cdot \in \mathcal{U}. \end{cases} \tag{27}$$

From a modeling point of view, the tagged pedestrian uses two controls:

- $(u_t)_{t\in[0,T]}$ - picked by an optimization procedure to reduce energy use, movement in densely crowded areas
- $(Z_t)_{t\in[0,T]}$ - to predict the best path to y_T given $(u_t)_{t\in[0,T]}$, given implicitly by the martingale representation theorem.

A spike pertubation technique leads to a Pontryagin type maximum principle[1].

[1] A Aurell and B Djehiche. "Modeling tagged pedestrian motion: a mean-field type control approach". In: *arXiv preprint arXiv:1801.08777v2* (2018).

Tagged pedestrian motion: control of mean-field BSDEs

Assumptions: i) $u \mapsto b(\cdot, \cdot, \cdot, \cdot, u)$ is Lipschitz and its y-,z- and μ-derivatives are bounded ii) $b(\cdot, 0, \delta_0, 0, u)$ is square-integrable for all $u \in U$ iii) $y_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^d)$ iv) admissible controls ($\mathcal{U}[0, T]$) take values in the compact set U and are square-integrable.

Theorem - necessary conditions

Suppose that $(\hat{X}_\cdot, \hat{Z}_\cdot, \hat{u}_\cdot)$ solves the control problem. Let H be the Hamiltonian

$$H(t, x, \mu, z, u, p) := b(t, x, \mu, z, u)p - f(t, x, \mu, u), \tag{28}$$

and let $p_\cdot$ solve the adjoint equation (where $\mathbb{P}_{\hat{X}_t} := \mathbb{P} \circ (X_t)^{-1}$),

$$\begin{cases} dp_t = -\left\{\partial_x H(t, \hat{X}_t, \mathbb{P}_{\hat{X}_t}, \hat{Z}_t, \hat{u}_t, p_t) + E\left[{}^*(\partial_\mu H(t, \hat{X}_t, \mathbb{P}_{\hat{X}_t}, \hat{Z}_t, \hat{u}_t, p_t))\right]\right\} dt \\ \qquad\qquad - \partial_z H(t, \hat{X}_t, \mathbb{P}_{X_t}, \hat{Z}_t, \hat{u}_t, p_t) dW_t, \\ p_0 = \partial_x h(\hat{X}_0, \mathbb{P}_{\hat{X}_0}) + E\left[{}^*(\partial_\mu h(\hat{X}_0, \mathbb{P}_{\hat{X}_t}))\right]. \end{cases} \tag{29}$$

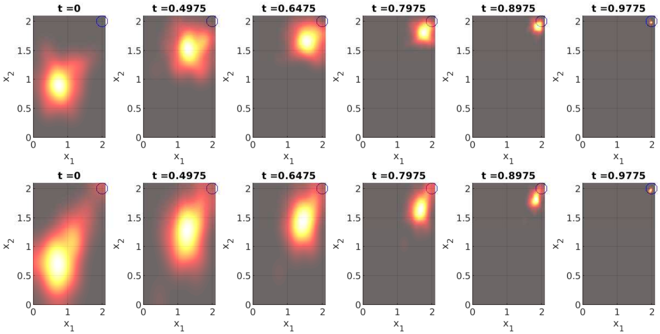
Then for a.e. t, $\mathbb{P}$-a.s.,

$$\hat{u}_t = \underset{\alpha \in U}{argmax}\ H(t, \hat{X}_t, \mathbb{P}_{\hat{X}_t}, \hat{Z}_t, \alpha, p_t). \tag{30}$$

Theorem - sufficient conditions

Suppose that H is concave in (x, μ, z, u), h is convex in (x, μ) and $\hat{u}_\cdot$ satisfies (30) $\mathbb{P}$-a.s. for a.e. t. Then $(\hat{X}_\cdot, \hat{Z}_\cdot, \hat{u}_\cdot)$ solves the control problem.

$$\begin{cases} \min_{u.\in\mathcal{U}} & \frac{1}{2}E\left[\int_0^1 \lambda_1 u_t^2 + \lambda_2(X_t - E[X_t])^2 dt + \lambda_3(X_0 - [0.2, 0.2]^T)^2\right], \\ \text{s.t.} & dX_t = (u_t + W_t)dt + Z_t dW_t, \;\; Y_1 = [2,2]^T. \end{cases} \tag{31}$$



Upper row: $(\lambda_1, \lambda_2, \lambda_3) = (50, 50, 10)$.
Lower row: $(\lambda_1, \lambda_2, \lambda_3) = (50, 0, 10)$.
Simulations based on the least-square Monte Carlo method[1].

[1] C Bender and J Steiner. "Least-squares Monte Carlo for backward SDEs". In: *Numerical methods in finance*. Springer, 2012, pp. 257–289.

Nash equilibrium: for $i = 1, \dots, \#\text{players}$,

$$\begin{aligned} &\text{Best reply}_i(\text{own eq. control; other's eq. controls}) \\ &\quad \leq \text{Best reply}_i(\text{any control; other's eq. controls}). \end{aligned} \tag{32}$$

In what follows,

- Best reply functional depends on marginal state distributions
- State dynamics are mean-field BSDEs

We start with an example...

Mean-field type games with BSDE dynamics: LQ example

Two players seek the Nash equilibrium: player 1 has state dynamics

$$\begin{cases} dX_t^1 = (a_1 u_t^1 + c_{11} W_t^1 + c_{12} W_t^2)dt + Z_t^{11} dW_t^1 + Z_t^{12} dW_t^2, \\ X_T^1 = x_T^1, \end{cases} \tag{33}$$

and wants to minimize

$$J^1(u_\cdot^1; u_\cdot^2) = E\left[\int_0^T \frac{r_1}{2}(u_t^1)^2 + \frac{\rho_1}{2}(X_t^1 - E[X_t^2])^2 dt + \frac{\nu_1}{2}(X_0^1 - x_0^1)^2\right]. \tag{34}$$

Player 2 has state dynamics

$$\begin{cases} dX_t^2 = (a_2 u_t^2 + c_{21} W_t^1 + c_{22} W_t^2)dt + Z_t^{21} dW_t^1 + Z_t^{22} dW_t^2, \\ X_T^2 = x_T^2, \end{cases} \tag{35}$$

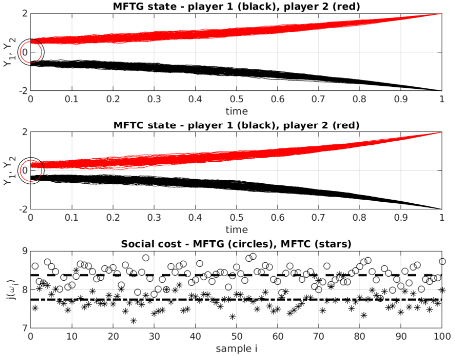
and wants to minimize

$$J^2(u_\cdot^2; u_\cdot^1) = E\left[\int_0^T \frac{r_2}{2}(u_t^2)^2 + \frac{\rho_2}{2}(X_t^2 - E[X_t^1])^2 dt + \frac{\nu_2}{2}(X_0^2 - x_0^2)^2\right]. \tag{36}$$

Alongside, a central planner wants to minimize the social cost

$$J(u_\cdot^1, u_\cdot^2) = \sum_{i=1}^{2} J^i(u_\cdot^i; u_\cdot^{-i}). \tag{37}$$

Mean-field type games with BSDE dynamics: LQ example



x_T^1	a_1	c_{11}	c_{12}	r_1	ρ_1	ν_1	x_0^1
-2	1	.3	0	1	1	1	$\mathcal{N}(0, 0.1)$

x_T^2	a_2	c_{21}	c_{22}	r_2	ρ_2	ν_2	x_0^2
2	1	0	.3	1	1	1	$\mathcal{N}(0, 0.1)$

Game: find $\hat{u}^1_\cdot, \hat{u}^2_\cdot$ such that for $i = 1, 2$,

$$J^i(u_\cdot; \hat{u}_\cdot^{-i}) \geq J^i(\hat{u}_\cdot^i; \hat{u}_\cdot^{-i}), \ \forall u_\cdot \in \mathcal{U}^i,$$

where

$$\begin{cases} dX_t^i = (a_i u_t^i + c_{i1} W_t^1 + c_{i2} W_t^2)dt \\ \qquad\qquad + Z_t^{i1} dW_t^1 + Z_t^{i2} dW_t^2, \\ X_T^i = x_T^i, \end{cases}$$

$$J^i(u_\cdot^i; u_\cdot^{-i}) =$$

$$E\left[\int_0^T \frac{r_2}{2}(u_t^2)^2 + \frac{\rho_2}{2}(X_t^2 - E[X_t^1])^2 dt + \frac{\nu_2}{2}(X_0^2 - x_0^2)^2\right]$$

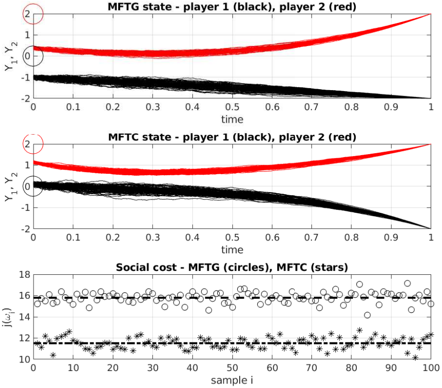
Optimal control: find $\bar{u}^1_\cdot, \bar{u}^2_\cdot$ such that

$$J(u_\cdot, v_\cdot) \geq J(\bar{u}^1_\cdot, \bar{u}^2_\cdot)$$

for all $(u_\cdot, v_\cdot) \in \mathcal{U}^1 \times \mathcal{U}^2$, where $J = J^1 + J^2$.

Social cost: Mean player cost.

Mean-field type games with BSDE dynamics: LQ example



x_T^1	a_1	c_{11}	c_{12}	r_1	ρ_1	ν_1	x_0^1
-2	1	.3	0	1	4	1	$\mathcal{N}(0, 0.1)$

x_T^2	a_2	c_{21}	c_{22}	r_2	ρ_2	ν_2	x_0^2
2	1	0	.3	1	0	1	$\mathcal{N}(2, 0.1)$

Game: find $\hat{u}_\cdot^1, \hat{u}_\cdot^2$ such that for $i = 1, 2$,

$$J^i(u_\cdot; \hat{u}_\cdot^{-i}) \geq J^i(\hat{u}_\cdot^i; \hat{u}_\cdot^{-i}), \ \forall u_\cdot \in \mathcal{U}^i,$$

where

$$\begin{cases} dX_t^i = (a_i u_t^i + c_{i1} W_t^1 + c_{i2} W_t^2)dt \\ \qquad + Z_t^{i1} dW_t^1 + Z_t^{i2} dW_t^2, \\ X_T^i = x_T^i, \end{cases}$$

$$J^i(u_\cdot^i; u_\cdot^{-i}) =$$

$$E\Big[\int_0^T \frac{r_2}{2}(u_t^2)^2 + \frac{\rho_2}{2}(X_t^2 - E[X_t^1])^2 dt + \frac{\nu_2}{2}(X_0^2 - x_0^2)^2 \Big]$$

Optimal control: find $\bar{u}_\cdot^1, \bar{u}_\cdot^2$ such that

$$J(u_\cdot, v_\cdot) \geq J(\bar{u}_\cdot^1, \bar{u}_\cdot^2)$$

for all $(u_\cdot, v_\cdot) \in \mathcal{U}^1 \times \mathcal{U}^2$, where $J = J^1 + J^2$.

Social cost: Mean player cost.

On $(\Omega, \mathcal{F}, \mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, satisfying the usual conditions, lives

- d_1- and d_2-dimensional Wiener processes $W^1_\cdot$ and $W^2_\cdot$
- two terminal values $x^1_T, x^2_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^d)$
- $\mathcal{F}_0$-measurable ξ (additional randomness at $t = 0$)

These five objects are independent and $(W^1_\cdot, W^2_\cdot, \xi)$ generate $\mathbb{F}$.

Let (U^i, d_{U^i}) be separable metric space, admissible controls for player i are

$$\mathcal{U}^i = \left\{ u : [0, T] \to U^i \mid \mathbb{F} - adapted,\ E \int_0^T d_{U^i}(u_s)^2 ds < \infty \right\} \qquad (38)$$

Given a pair of admissible controls $(u^1_\cdot, u^2_\cdot)$, the state dynamics are

$$dX^i_t = b^i(\Theta^i_t, \Theta^{-i}_t, Z_t)dt + Z^{i,1}_t dW^1_t + Z^{i,2}_t dW^2_t,\ X^i_T = x^i_T,\ i = 1, 2, \qquad (39)$$

where $\Theta^i_t := (X^i_t, \mathbb{P} \circ (X^i_t)^{-1}, u^i_t)$ and $Z_t = [Z^{11}_t Z^{12}_t Z^{21}_t Z^{22}_t]$.

Assumption 1: $b^i(\cdot, 0, \ldots, 0)$ is square integrable and given v, $b^i(\cdot, v)$ is $\mathcal{F}_t$-progressively measurable.
Assumption 2: Given a pair of admissible controls, b^i is Lipschitz-continuous in all other arguments (Wasserstein 2-metric for measures, trace-norm for matrices).
These assumptions implies existence and uniqueness.[1]

Theorem

Under assumption 1 and 2, there exists a unique solution $(X^i_\cdot, [Z^{i1}_\cdot, Z^{i2}_\cdot])$, $i = 1, 2$, to the mean-field BSDE system modelling player state dynamics. Furthermore, $Z^{ij}_\cdot$ is square integrable and $E[\sup_{t\in[0,T]} X_t^2] < \infty$.

The *best reply* ('cost') functional of player i is

$$J^i(u^i_\cdot; u^{-i}_\cdot) = E\left[\int_0^T f^i(\Theta^i_t, \Theta^{-i}_t)dt + h^i(\theta^i_0, \theta^{-i}_0)\right], \tag{40}$$

where $\theta^i_t = (X^i_t, \mathbb{P} \circ (X^i_t)^{-1})$. Goal: characterize Nash equilibria to this game.

[1] Rainer Buckdahn, Juan Li, and Shige Peng. "Mean-field backward stochastic differential equations and related partial differential equations". In: *Stochastic Processes and their Applications* 119.10 (2009), pp. 3133–3154.

1. Assume that there exists an equilibrium control pair $\hat{u}^1_\cdot, \hat{u}^2_\cdot$. Make a spike variation of $\hat{u}^1_\cdot$; for some $u_\cdot \in \mathcal{U}^1$ and $E_\epsilon \subset [0, T]$ of size $|E_\epsilon| = \epsilon$,

$$\bar{u}^{\epsilon,1}_t := \begin{cases} \hat{u}^1_t, & t \in [0, T] \backslash E_\epsilon, \\ u_t, & t \in E_\epsilon. \end{cases} \tag{41}$$

Whenever player 1 uses $\bar{u}^{\epsilon,1}_\cdot$, denote state dynamics by $\bar{X}^{i,\epsilon}_\cdot$, $i = 1, 2$.

2. Compare the perturbed control's best reply to the equilibrium,

$$J^1(\bar{u}^{\epsilon,1}_\cdot; \hat{u}^2_\cdot) - J^1(\hat{u}^1_\cdot, \hat{u}^2_\cdot) = E\left[\int_0^T \bar{f}^{\epsilon,1}_t - \hat{f}^1_t dt + \bar{h}^{\epsilon,1}_0 - \hat{h}^1_0\right]. \tag{42}$$

3. Approximate the cost difference,

$$\begin{aligned} \bar{h}^{\epsilon,i}_0 - \hat{h}^i_0 = & \sum_{j=1}^{2} \left\{ \partial_{x^j} \hat{h}^i_0 (\bar{X}^{\epsilon,j}_0 - \hat{X}^j_0) + \mathbb{E}\left[(\partial_{\mu^j} \hat{h}^i_0)^* (\bar{X}^{\epsilon,j}_0 - \hat{X}^j_0) \right] \right\} \\ & + \sum_{j=1}^{2} \left\{ o\left(|\bar{X}^{\epsilon,j}_0 - \hat{X}^j_0| \right) + o\left(E[|\bar{X}^{\epsilon,j}_0 - \hat{X}^j_0|^2]^{1/2} \right) \right\}. \end{aligned} \tag{43}$$

Assumption 3: b^i, f^i, h^i are, for all t, a.s. differentiable at the equilibrium, where their derivatives are a.s. uniformly bounded for all t and $\partial_{y^j}\hat{h}^i_0 + E[^*(\partial_{\mu^j}\hat{h}^i_0)] \in L^2_{\mathcal{F}_0}(\Omega;\mathbb{R}^d)$.
Assumption 4: b^i is a.s. Lipschitz in the controls, for all t.

4. Find the first variation processes:
Let assumptions 1-4 be in place and let $(Y^i_\cdot, [V^{i1}_\cdot, V^{i2}_\cdot])$, $i=1,2$ solve the linear BSDE system

$$\begin{cases} dY^i_t = \Big(\sum_{j=1}^2 \{\partial_{x^j}\hat{b}^i_t Y^j_t + E[(\partial_{\mu^j}\hat{b}^i_t)^* Y^j_t]\} + \sum_{j,k=1}^2 \partial_{z^{j,k}}\hat{b}^i_t V^{jk}_t \\ \qquad +\delta_1 b^i(t) 1_{E_\epsilon}(t)\Big) dt + \sum_{j=1}^2 V^{ij}_t dW^j_t, \\ Y^i_T = 0 \end{cases} \tag{44}$$

where $\delta_i\phi(t) := \phi(\hat{\theta}^i_t, \bar{u}^{\epsilon,i}_t, \hat{\Theta}^{-i}_t, \hat{Z}_t) - \hat{\phi}_t$. Then

$$\begin{aligned} &\sup_{0\le t\le T} E\left[|Y^i_t|^2 + \sum_{j=1}^2 \int_0^t \|V^{ij}_s\|^2_F ds\right] \le C\epsilon^2 \\ &\sup_{0\le t\le T} E\left[|\bar{X}^{\epsilon,i}_t - \hat{X}^i_t - Y^i_t|^2 + \sum_{j=1}^2 \int_0^t \|\bar{Z}^{\epsilon,ij}_s - \hat{Z}^{ij}_s - V^{ij}_s\|^2_F ds\right] \le C\epsilon^2. \end{aligned} \tag{45}$$

Using step 4,

$$E[\bar{h}_0^{\epsilon,1} - \hat{h}_0^1] = E\left[\sum_{j=1}^{2} \partial_{x^j} \hat{h}_0^1 Y_0^j + E[(\partial_{\mu^j} \hat{h}_0^1)^* Y_0^j]\right] + o(\epsilon). \tag{46}$$

5. Find the duality relation by introducing the adjoint process:
Let assumptions 1-3 hold and let $p_\cdot^{1j}$ be given by

$$\begin{cases} dp_t^{1j} = -\left\{\partial_{x^j} \hat{H}_t^1 + E[^*(\partial_{\mu^j} \hat{H}_t^1)]\right\} dt - \sum_{k=1}^{2} \partial_{z^{jk}} \hat{H}_t^1 dW_t^k. \\ p_0^{1j} = \partial_{x^j} \hat{h}_0^1 + E[^*(\partial_{\mu^j} \hat{h}_0^1)] \end{cases} \tag{47}$$

where $\hat{H}^1 := \hat{b}_t^1 p_t^{11} + \hat{b}_t^2 p_t^{12} - \hat{f}_t^1$ is player 1's Hamiltonian, evaluated at the equilibrium. Then the following duality relation holds

$$E\left[\sum_{j=1}^{2} p_0^{1j} Y_0^j\right] = -E\left[\int_0^T \sum_{j=1}^{2} p_t^{1j} \delta_1 b^j(t) 1_{E_\epsilon}(t) + Y_t^j \left(\partial_{x^j} \hat{f}_t^1 + E[^*(\partial_{\mu^j} \hat{f}_t^1)]\right) dt\right]$$

Use step 5 to conclude that

$$E\left[\bar{h}_0^{\epsilon,1} - \hat{h}_0^1\right] = E\left[\sum_{j=1}^{2} p_0^{1j} Y_0^j\right] + o(\epsilon). \tag{48}$$

6. Approximate the running cost difference, and get

$$J^1(\bar{u}_\cdot^{\epsilon,1}; \hat{u}_\cdot^2) - J^1(\hat{u}_\cdot^1, \hat{u}_\cdot^2) = -E\left[\int_0^T \delta_1 H^1(t) 1_{E_\epsilon}(t) dt\right] + o(\epsilon). \tag{49}$$

Step 1-6 has lead us from functional minimization to pointwise minimization![1]

Step 1-6 can be done for a spike pertubation of player 2's control. The last relationship between best reply difference and Hamiltonian difference yields *necessary and sufficient conditions for Nash equilibria*.

[1] Alexander Aurell. "Mean-Field Type Games between Two Players Driven by Backward Stochastic Differential Equations". In: (2018).

Mean-field type games with BSDE dynamics: SMP

Necessary conditions

Suppose that $(\hat{X}^i_\cdot, [\hat{Z}^{i,1}_\cdot, \hat{Z}^{i,2}_\cdot])$, $i = 1, 2$, is an equilibrium for the MFTG and that $p^{ij}_\cdot$, $i, j = 1, 2$, solves the adjoint equations. Then, for $i = 1, 2$,

$$\hat{u}^i_t = \max_{\alpha \in U^i} H^i(\hat{\theta}^i_t, \alpha, \hat{\Theta}^{-i}_t, \hat{Z}_t, p^{i,1}_t, p^{i,2}_t), \quad a.s., \ a.e.t \tag{50}$$

Sufficient conditions

Suppose $\hat{u}^i$ satisfies (50). Suppose furthermore that

$$(x^1, \mu^1, u^1, x^2, \mu^2, u^2) \mapsto H^i(x^i, \mu^i, u^i, x^{-i}, \mu^{-i}, u^{-i}, z, p^{i1}, p^{i2}) \tag{51}$$

is concave a.s. and

$$(x^1, \mu^1, x^2, mu^2) \mapsto h^i(x^i, \mu^i, x^{-i}, \mu^{-i}) \tag{52}$$

is convex a.s. Then $\hat{u}^1_\cdot, \hat{u}^2_\cdot$ constitute a Nash equilibrium control.

Revisiting the LQ example

Our LQ example satisfies the sufficient conditions. Pointwise minimization of the Hamiltonian yields

$$\hat{u}_t^i = \frac{a_i}{r_i} p_t^{ii}. \tag{53}$$

Steps 1-6 can be carried out for the central planner problem, though the first variation and adjoint processes and the Hamiltonian will have different forms. The central planner's optimal control for player i is

$$\hat{u}_t^i = \frac{a_i}{r_i} p_t^{i}. \tag{54}$$

Both (53) and (54) can be found explicitly (up to a set of Ricatti ODEs).

Improvement on a societal level can be quantified by *the price of anarchy*[1]

$$PoA := \sup_{(\hat{u}^1_\cdot, \hat{u}^2_\cdot)\ \text{Nash}} J(\hat{u}^1_\cdot, \hat{u}^1_\cdot) \Big/ \min_{u^i_\cdot \in \mathcal{U}^i, i=1,2} J(u^1_\cdot, u^2_\cdot). \tag{55}$$

x_T^1	a_1	c_{11}	c_{12}	r_1	ρ_1	ν_1	x_0^1
-2	1	0.3	0	1	1	1	$\mathcal{N}(0, 0.1)$
x_T^2	a_2	c_{21}	c_{22}	r_2	ρ_2	ν_2	x_0^2
2	1	0	0.3	1	1	1	$\mathcal{N}(0, 0.1)$

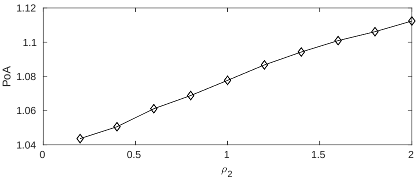


Figure: Variation of ρ_2, weight for mean-field cost, in $[0.2, 2]$.

[1] Christos Papadimitriou. "Algorithms, games, and the internet". In: *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. ACM. 2001, pp. 749–753.

Revisiting the LQ example

x_T^1	a_1	c_{11}	c_{12}	r_1	ρ_1	ν_1	x_0^1
-2	1	0.3	0	1	1	1	$\mathcal{N}(0, 0.1)$
x_T^2	a_2	c_{21}	c_{22}	r_2	ρ_2	ν_2	x_0^2
2	1	0	0.3	1	1	1	$\mathcal{N}(0, 0.1)$

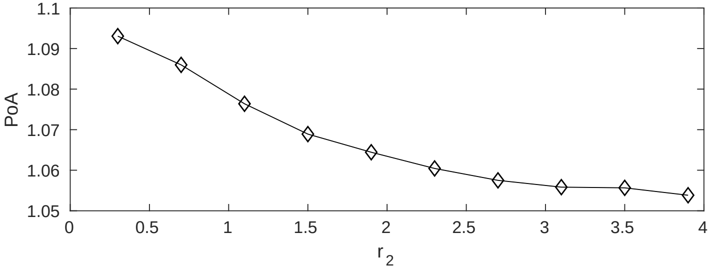


Figure: Variation of r_2, weight on control, in $[0.2, 4]$.

Revisiting the LQ example

x_T^1	a_1	c_{11}	c_{12}	r_1	ρ_1	ν_1	x_0^1
-2	1	0.3	0	1	1	1	$\mathcal{N}(0, 0.1)$
x_T^2	a_2	c_{21}	c_{22}	r_2	ρ_2	ν_2	x_0^2
2	1	0	0.3	1	1	1	$\mathcal{N}(0, 0.1)$

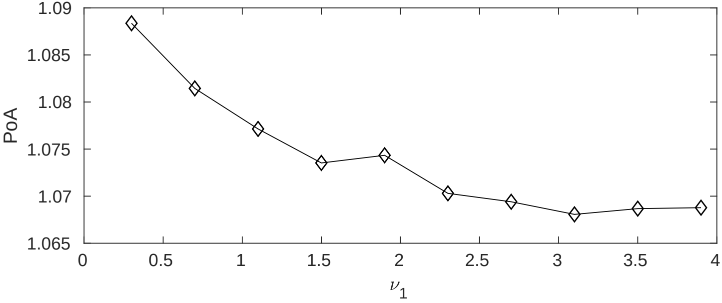


Figure: Variation of ν_1, weight on initial cost, in $[0.2, 4]$.

Revisiting the LQ example

x_T^1	a_1	c_{11}	c_{12}	r_1	ρ_1	ν_1	x_0^1
-2	1	0.3	0	1	1	1	$\mathcal{N}(0, 0.1)$

x_T^2	a_2	c_{21}	c_{22}	r_2	ρ_2	ν_2	x_0^2
2	1	0	0.3	1	1	1	$\mathcal{N}(0, 0.1)$

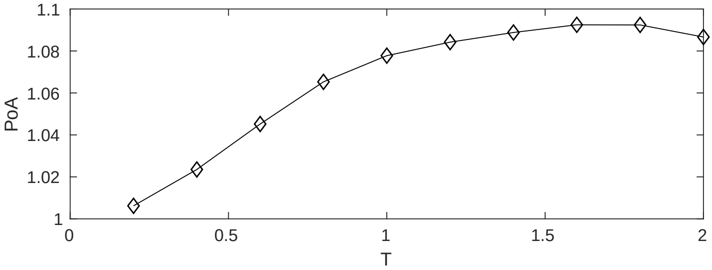


Figure: Variation of time horizon T in $[0.2, 2]$.

Conclusion

- Many variations on control problems involving control-dependent marginal distribution out there.
- Model suggested for certain pedestrian movement.
- Mean-field type game of players evolving according to BSDEs.

Thank you!