MEAN FIELD GAMES WITH MAJOR AND MINOR PLAYERS

René Carmona

Department of Operations Research & Financial Engineering
PACM
Princeton University

CEMRACS - Luminy, July 17, 2017

MFG with Major and Minor Players Set-Up

R.C. - G. Zhu, R.C. - P. Wang

State equations

$$\begin{cases} dX_t^0 & = b_0(t, X_t^0, \mu_t, \alpha_t^0)dt + \sigma_0(t, X_t^0, \mu_t, \alpha_t^0)dW_t^0 \\ dX_t & = b(t, X_t, \mu_t, X_t^0, \alpha_t, \alpha_t^0)dt + \sigma(t, X_t, \mu_t, X_t^0, \alpha_t, \alpha_t^0 dW_t, \end{cases}$$

Costs

$$\begin{cases} J^0(\boldsymbol{\alpha}^0, \boldsymbol{\alpha}) & = \mathbb{E}\left[\int_0^T f_0(t, X_t^0, \mu_t, \alpha_t^0)dt + g^0(X_T^0, \mu_T)\right] \\ J(\boldsymbol{\alpha}^0, \boldsymbol{\alpha}) & = \mathbb{E}\left[\int_0^T f(t, X_t, \mu_t^N, X_t^0, \alpha_t, \alpha_t^0)dt + g(X_T, \mu_T)\right], \end{cases}$$

Open Loop Version of the MFG Problem

The controls used by the major player and the representative minor player are of the form:

$$\alpha_t^0 = \phi^0(t, W^0_{[0,T]}), \quad \text{and} \quad \alpha_t = \phi(t, W^0_{[0,T]}, W_{[0,T]}), \tag{1}$$

for deterministic progressively measurable functions

$$\phi^0 : [0, T] \times \mathcal{C}([0, T]; \mathbb{R}^{d_0}) \mapsto A_0$$

and

$$\phi : [0, T] \times \mathcal{C}([0, T]; \mathbb{R}^{d}) \times \mathcal{C}([0, T]; \mathbb{R}^{d}) \mapsto A$$

The Major Player Best Response

Assume representative minor player uses the open loop control given by $\phi : (t, w^0, w) \mapsto \phi(t, w^0, w)$,

Major player minimizes

$$J^{\phi,0}(\boldsymbol{\alpha}^0) = \mathbb{E}\Big[\int_0^T f_0(t, X_t^0, \mu_t, \alpha_t^0)dt + g^0(X_T^0, \mu_T)\Big]$$

under the dynamical constraints:

$$\begin{cases} dX_t^0 &= b_0(t, X_t^0, \mu_t, \alpha_t^0)dt + \sigma_0(t, X_t^0, \mu_t, \alpha_t^0)dW_t^0 \\ dX_t &= b(t, X_t, \mu_t, X_t^0, \phi(t, W^0_{[0,T]}, W_{[0,T]}), \alpha_t^0)dt \\ & \qquad + \sigma(t, X_t, \mu_t, X_t^0, \phi(t, W^0_{[0,T]}, W_{[0,T]}), \alpha_t^0)dW_t, \end{cases}$$

$\mu_t = \mathcal{L}(X_t | W^0_{[0,t]})$ conditional distribution of X_t given $W^0_{[0,t]}$.

Major player problem as the search for:

$$\phi^{0,*}(\phi) = \arg \inf_{\alpha_t^0 = \phi^0(t, W^0_{[0,T]})} J^{\phi,0}(\boldsymbol{\alpha}^0) \qquad (2)$$

Optimal control of the *conditional* McKean-Vlasov type!

The Rep. Minor Player Best Response

System against which **best response** is sought comprises

- a major player
- a field of minor players different from the representative minor player
- Major player uses strategy $\alpha_t^0 = \phi^0(t, W^0_{[0,T]})$
- Representative of the field of minor players uses strategy $\alpha_t = \phi(t, W^0_{[0,T]}, W_{[0,T]})$.

State dynamics

$$\begin{cases} dX_t^0 = b_0(t, X_t^0, \mu_t, \phi^0(t, W^0_{[0,T]}))dt + \sigma_0(t, X_t^0, \mu_t, \phi^0(t, W^0_{[0,T]}))dW_t^0 \\ dX_t = b(t, X_t, \mu_t, X_t^0, \phi(t, W^0_{[0,T]}, W_{[0,T]}), \phi^0(t, W^0_{[0,T]}))dt \\ \qquad\qquad +\sigma(t, X_t, \mu_t, X_t^0, \phi(t, W^0_{[0,T]}, W_{[0,T]}), \phi^0(t, W^0_{[0,T]}))dW_t, \end{cases}$$

where $\mu_t = \mathcal{L}(X_t | W^0_{[0,t]})$ is the conditional distribution of X_t given $W^0_{[0,t]}$.

Given ϕ^0 and ϕ, **SDE of (conditional) McKean-Vlasov type**

The Rep. Minor Player Best Response (cont.)

Representative minor player chooses a strategy $\overline{\alpha}_t = \overline{\phi}(t, W^0_{[0,T]}, W_{[0,T]})$ to minimize

$$J^{\phi^0,\phi}(\bar{\alpha}) = \mathbb{E}\Big[\int_0^T f(t, \overline{X}_t, X^0_t, \mu_t, \bar{\alpha}_t, \phi^0(t, W^0_{[0,T]}))dt + g(\overline{X}_T, \mu_t)\Big],$$

where the dynamics of the virtual state $\overline{X}_t$ are given by:

$$d\overline{X}_t = b(t, \overline{X}_t, \mu_t, X^0_t, \bar{\phi}(t, W^0_{[0,T]}, W_{[0,T]}), \phi^0(t, W^0_{[0,T]}))dt$$
$$+ \sigma(t, \overline{X}_t, \mu_t, X^0_t, \bar{\phi}(t, W^0_{[0,T]}, W_{[0,T]}), \phi^0(t, W^0_{[0,T]}))d\overline{W}_t,$$

for a Wiener process $\overline{\mathbf{W}} = (\overline{W}_t)_{0\le t\le T}$ independent of the other Wiener processes.

- Optimization problem **NOT** of McKean-Vlasov type.
- **Classical** optimal control problem with **random coefficients**

$$\overline{\phi}^*(\phi^0, \phi) = \arg \inf_{\overline{\alpha}_t = \overline{\phi}(t, W^0_{[0,T]}, W_{[0,T]})} J^{\phi^0,\phi}(\bar{\alpha})$$

Nash Equilibrium

Search for **Best Response Map Fixed Point**

$$(\hat{\phi}^0, \hat{\phi}) = (\phi^{0,*}(\hat{\phi}), \bar{\phi}^*(\hat{\phi}^0, \hat{\phi})).$$

Fixed point in a **space of controls**, not **measures** !!!

Closed Loop Versions of the MFG Problem

- **Closed Loop Version**

  Controls of the major player and the representative minor player are of the form:

$$\alpha_t^0 = \phi^0(t, X^0_{[0,T]}, \mu_t), \quad \text{and} \quad \alpha_t = \phi(t, X_{[0,T]}, \mu_t, X^0_{[0,T]}),$$

  for deterministic progressively measurable functions
  $\phi^0 : [0,T] \times \mathcal{C}([0,T]; \mathbb{R}^{d_0}) \times \mathcal{P}_2(\mathbb{R}^d) \mapsto A_0$ and
  $\phi : [0,T] \times \mathcal{C}([0,T]; \mathbb{R}^{d}) \times \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{C}([0,T]; \mathbb{R}^{d}) \mapsto A$.

- **Markovian Version**

  Controls of the major player and the representative minor player are of the form:

$$\alpha_t^0 = \phi^0(t, X^0_t, \mu_t), \quad \text{and} \quad \alpha_t = \phi(t, X_t, \mu_t, X^0_t),$$

  for deterministic feedback functions $\phi^0 : [0,T] \times \mathbb{R}^{d_0} \times \mathcal{P}_2(\mathbb{R}^d) \mapsto A_0$ and
  $\phi : [0,T] \times \mathbb{R}^{d} \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^{d_0} \mapsto A$.

Nash Equilibrium

Search for **Best Response Map Fixed Point**

$$(\hat{\phi}^0, \hat{\phi}) = \left(\phi^{0,*}(\hat{\phi}), \bar{\phi}^*(\hat{\phi}^0, \hat{\phi})\right).$$

CONTRACT THEORY: A STACKELBERG VERSION

R.C. - D. Possamaï - N. Touzi

State equation

$$dX_t = \sigma(t, X_t, \nu_t, \alpha_t)[\lambda(t, X_t, \nu_t, \alpha_t)dt + dW_t],$$

- X_t Agent output
- α_t agent effort (control)
- ν_t distribution of output and effort (control) of agent

Rewards

$$\begin{cases} J^0(\xi) &= \mathbb{E}\big[U_P(X_{[0,T]}, \nu_T, \xi)\big] \\ J(\xi, \boldsymbol{\alpha}) &= \mathbb{E}\big[-\int_0^T f(t, X_t, \nu_t, \alpha_t)dt + U_A(\xi)\big], \end{cases}$$

- Given the choice of a contract ξ by the Principal
  - Each agent in the field of *exchangeable agents*
    - chooses an effort level α_t
    - meets his/her *reservation price*
    - get the field of agents in a (MF) Nash equilibrium
- Principal chooses the contract to maximize his/her expected utility

Linear Quadratic Models

State dynamics

$$\begin{cases} dX_t^0 &= (L_0 X_t^0 + B_0 \alpha_t^0 + F_0 \bar{X}_t)dt + D_0 dW_t^0 \\ dX_t &= (L X_t + B\alpha_t + F\bar{X}_t + G X_t^0)dt + D dW_t \end{cases}$$

where $\bar{X}_t = \mathbb{E}[X_t | \mathcal{F}_t^0]$, $(\mathcal{F}_t^0)_{t\geq 0}$ filtration generated by $\mathbf{W}^0$

Costs

$$J^0(\boldsymbol{\alpha}^0, \boldsymbol{\alpha}) = \mathbb{E}\left[\int_0^T [(X_t^0 - H_0\bar{X}_t - \eta_0)^\dagger Q_0 (X_t^0 - H_0\bar{X}_t - \eta_0) + \alpha_t^{0\dagger} R_0 \alpha_t^0] dt\right]$$

$$J(\boldsymbol{\alpha}^0, \boldsymbol{\alpha}) = \mathbb{E}\left[\int_0^T [(X_t - HX_t^0 - H_1\bar{X}_t - \eta)^\dagger Q (X_t - HX_t^0 - H_1\bar{X}_t - \eta) + \alpha_t^\dagger R \alpha_t] dt\right]$$

in which Q, Q_0, R, R_0 are symmetric matrices, and R, R_0 are assumed to be positive definite.

Equilibria

- **Open Loop Version**
  - Optimization problems + fixed point $\Longrightarrow$ large FBSDE
  - **affine** FBSDE solved by a large matrix Riccati equation
- **Closed Loop Version**
  - Fixed point step more difficult
  - Search limited to controls of the form

$$\begin{aligned} \alpha_t^0 &= \phi^0(t, X_t^0, \bar{X}_t) = \phi_0^0(t) + \phi_1^0(t) X_t^0 + \phi_2^0(t) \bar{X}_t \\ \alpha_t &= \phi(t, X_t, X_t^0, \bar{X}_t) = \phi_0(t) + \phi_1(t) X_t + \phi_2(t) X_t^0 + \phi_3(t) \bar{X}_t \end{aligned}$$

  - Optimization problems + fixed point $\Longrightarrow$ large FBSDE
  - **affine** FBSDE solved by a large matrix Riccati equation

Solutions are not the same !!!!

Application to Bee Swarming

- $V_t^{0,N}$ velocity of the (major player) streaker bee at time t
- $V_t^{i,N}$ the velocity of the i-th worker bee, $i = 1, \cdots, N$ at time t
- **Linear** dynamics

$$\begin{cases} dV_t^{0,N} = \alpha_t^0 dt + \Sigma_0 dW_t^0 \\ dV_t^{i,N} = \alpha_t^i dt + \Sigma dW_t^i \end{cases}$$

- **Minimization of Quadratic** costs

$$J^0 = \mathbb{E}\Big[\int_0^T \big(\lambda_0 \| V_t^{0,N} - \nu_t\|^2 + \lambda_1 \| V_t^{0,N} - \bar{V}_t^N\|^2 + (1 - \lambda_0 - \lambda_1)\|\alpha_t^0\|^2\big)\, dt\Big]$$

  - $\bar{V}_t^N := \frac{1}{N}\sum_{i=1}^N V_t^{i,N}$ the average velocity of the followers,
  - deterministic function $[0, T] \ni t \to \nu_t \in \mathbb{R}^d$ (leader's free will)
  - λ_0 and λ_1 are positive real numbers satisfying $\lambda_0 + \lambda_1 \leq 1$

$$J^i = \mathbb{E}\Big[\int_0^T \big(l_0 \| V_t^{i,N} - V_t^{0,N}\|^2 + l_1 \| V_t^{i,N} - \bar{V}_t^N\|^2 + (1 - l_0 - l_1)\|\alpha_t^i\|^2\big)\, dt\Big]$$

$l_0 \geq 0$ and $l_1 \geq 0$, $l_0 + l_1 \leq 1$.

Sample Trajectories in Equilibrium

$$\nu(t) := [-2\pi \sin(2\pi t), 2\pi \cos(2\pi t)]$$

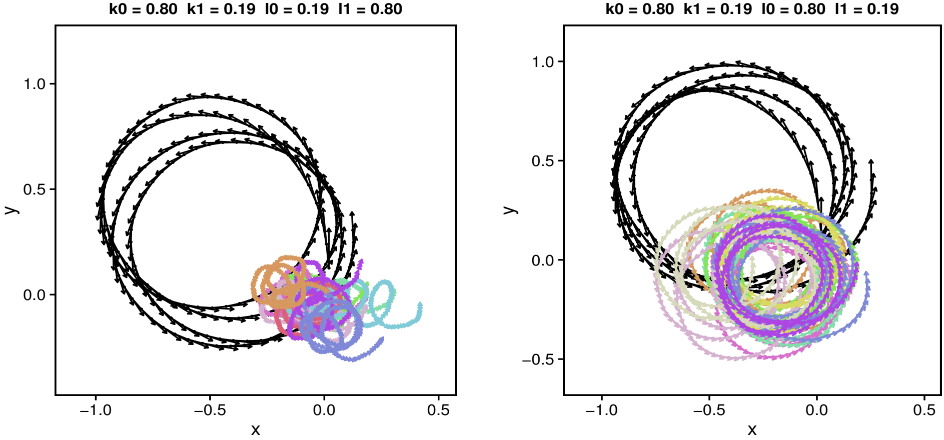


FIGURE: Optimal velocity and trajectory of follower and leaders

Sample Trajectories in Equilibrium

$$\nu(t) := [-2\pi \sin(2\pi t), 2\pi \cos(2\pi t)]$$

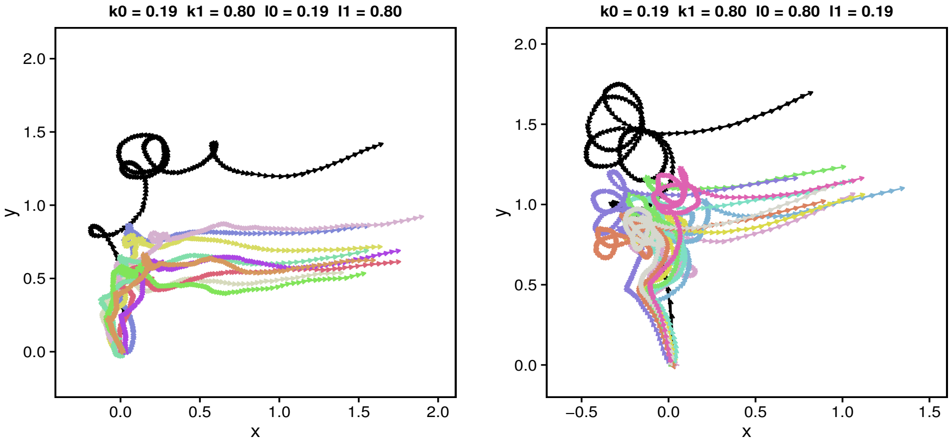


Figure: Optimal velocity and trajectory of follower and leaders

Conditional Propagation of Chaos

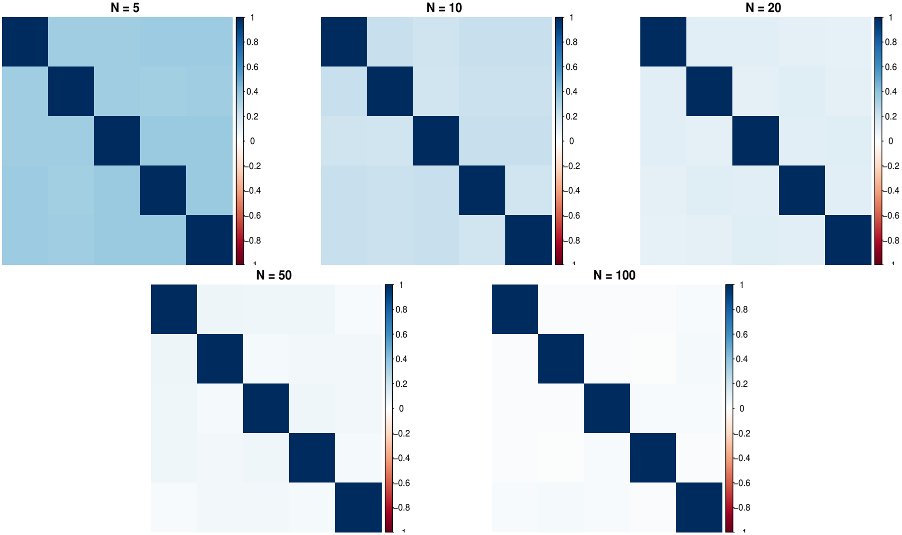


FIGURE: Conditional correlation of 5 followers' velocities

Finite State Spaces: A Cyber Security Model

Kolokolstov - Bensoussan, R.C. - P. Wang

- N computers in a network (**minor players**)
- One hacker / attacker (**major player**)
- Action of major player affect minor player states (even when $N >> 1$)
- Major player feels only μ_t^N the empirical distribution of the minor players' states

Finite State Space: each computer is in one of 4 states

- protected & infected
- protected & susceptible to be infected
- unprotected & infected
- unprotected & susceptible to be infected

Continuous time Markov chain in $E = \{DI, DS, UI, US\}$

Each **player's action** is intended to affect the **rates of change** from one state to another to minimize **expected costs**

$$J(\boldsymbol{\alpha}^0, \boldsymbol{\alpha}) = \mathbb{E}\Big[\int_0^T (k_D \mathbf{1}_D + k_I \mathbf{1}_I)(X_t) dt\Big]$$

$$J^0(\boldsymbol{\alpha}^0, \boldsymbol{\alpha}) = \mathbb{E}\Big[\int_0^T \big(-f_0(\mu_t) + k_H \phi^0(\mu_t)\big) dt\Big]$$

Finite State Mean Field Games

State Dynamics

$(X_t)_{t\geq 0}$ continuous time Markov chain in E with Q-matrix $(q_t(x,x'))_{t\geq 0, x,x'\in E}$.

Mean Field structure of the Q-matrix

$$q_t(x,x') = \lambda_t(x,x',\mu,\alpha)$$

Control Space

$A \subset \mathbb{R}^k$, sometime A finite, e.g. $A=\{0,1\}$, or a function space

Control Strategies in **feedback** form

$$\alpha_t = \phi(t,X_t), \qquad \text{for some } \phi : [0,T]\times E \mapsto A$$

Mean Field Interaction through Empirical Measures

$$\mu \in \mathcal{P}(E) = (\mu(\{x\}))_{x\in E}$$

Kolmogorov-Fokker-Planck equation: if $\mu_t = \mathcal{L}(X_t)$

$$\begin{aligned}
\partial_t \mu_t(\{x\}) &= [L_t^{\mu_t,\phi(t,\,\cdot\,),\dagger}\mu_t](\{x\}) \\
&= \sum_{x'\in E} \mu_t(\{x'\})\hat{q}_t^{\mu_t,\phi(t,\,\cdot\,)}(x',x), \\
&= \sum_{x'\in E} \mu_t(\{x'\})\lambda_t(x,x',\mu_t,\phi(t,x)) \qquad x\in E,
\end{aligned}$$

Finite State Mean Field Games: Optimization

Hamiltonian

$$H(t, x, \mu, h, \alpha) = \sum_{x' \in E} \lambda_t(x, x', \mu, \alpha) h(x') \; + f(t, x, \mu, \alpha).$$

Hamiltonian minimizer

$$\hat{\alpha}(t, x, \mu, h) = \arg \inf_{\alpha \in A} H(t, x, \mu, h, \alpha),$$

Minimized Hamiltonian

$$H^*(t, x, \mu, h) = \inf_{\alpha \in A} H(t, x, \mu, h, \alpha) = H(t, x, \mu, h, \hat{\alpha}(t, x, \mu, h)).$$

HJB Equation

$$\partial_t u^{\boldsymbol{\mu}}(t, x) + H^*(t, x, \mu_t, u^{\boldsymbol{\mu}}(t, \cdot)) = 0, \qquad 0 \le t \le T, \; x \in E,$$

with terminal condition $u^{\boldsymbol{\mu}}(T, x) = g(x, \mu_T)$.

Transition Rates Q-matrices

For $\alpha = 0$

$\lambda_t(\cdot, \cdot, \mu, \alpha^0, 0) =$

$$\begin{array}{c} \\ \mathrm{DI} \\ \mathrm{DS} \\ \mathrm{UI} \\ \mathrm{US} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{DI} & \mathrm{DS} & \mathrm{UI} & \mathrm{US} \end{array} \\ \left[\begin{array}{cccc} \cdots & q^{\mathrm{D}}_{\mathrm{rec}} & 0 & 0 \\ \alpha^0 q^{\mathrm{D}}_{\mathrm{inf}} + \beta_{\mathrm{DD}}\mu(\{\mathrm{DI}\}) + \beta_{\mathrm{UD}}\mu(\{\mathrm{UI}\}) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & q^{\mathrm{U}}_{\mathrm{rec}} \\ 0 & 0 & \alpha^0 q^{\mathrm{U}}_{\mathrm{inf}} + \beta_{\mathrm{UU}}\mu(\{\mathrm{UI}\}) + \beta_{\mathrm{DU}}\mu(\{\mathrm{DI}\}) & \cdots \end{array}\right] \end{array}$$

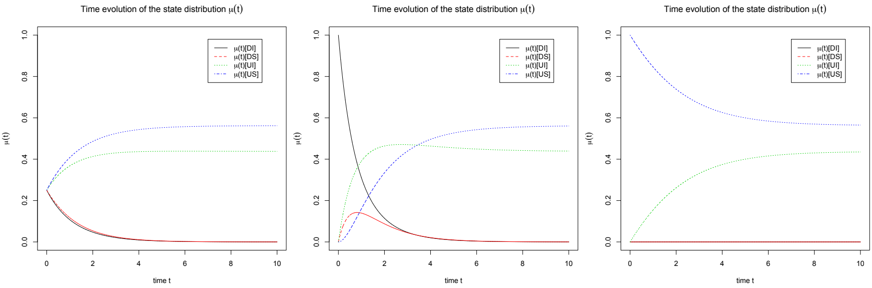
and for $\alpha = 1$:

$\lambda_t(\cdot, \cdot, \mu, \alpha^0, 1) =$

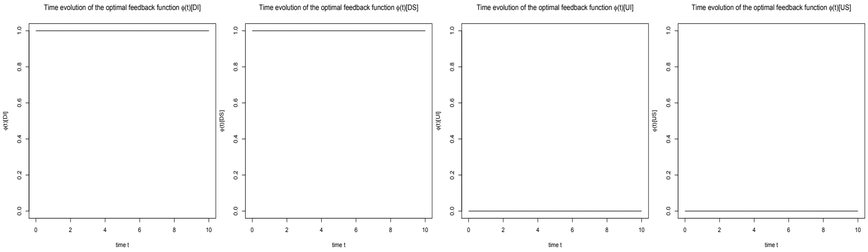
$$\begin{array}{c} \\ \mathrm{DI} \\ \mathrm{DS} \\ \mathrm{UI} \\ \mathrm{US} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{DI} & \mathrm{DS} & \mathrm{UI} & \mathrm{US} \end{array} \\ \left[\begin{array}{cccc} \cdots & q^{\mathrm{D}_{\mathrm{rec}}} & \lambda & 0 \\ \alpha^0 q^{\mathrm{D}}_{\mathrm{inf}} + \beta_{\mathrm{DD}}\mu(\{\mathrm{DI}\}) + \beta_{\mathrm{UD}}\mu(\{\mathrm{UI}\}) & \cdots & 0 & \lambda \\ \lambda & 0 & \cdots & q^{\mathrm{U}}_{\mathrm{rec}} \\ 0 & \lambda & \alpha^0 q^{\mathrm{U}}_{\mathrm{inf}} + \beta_{\mathrm{UU}}\mu(\{\mathrm{UI}\}) + \beta_{\mathrm{DU}}\mu(\{\mathrm{DI}\}) & \cdots \end{array}\right] \end{array}$$

where all the instances of $\cdots$ should be replaced by the negative of the sum of the entries of the row in which $\cdots$ appears on the diagonal.

Equilibrium Distribution over Time with Constant Attacker

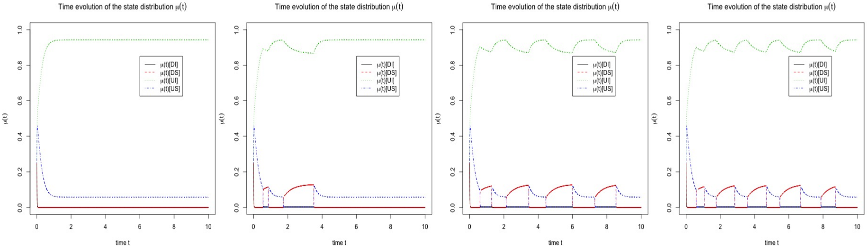


Equilibrium Optimal Feedback $\phi(t, \cdot)$



From left to right $\phi(t, \mathrm{DI})$, $\phi(t, \mathrm{DS})$, $\phi(t, \mathrm{UI})$, and $\phi(t, \mathrm{US})$.

Convergence May Be Elusive



From left to right, time evolution of the distribution μ_t for the parameters given in the text, after 1, 5, 20, and 100 iterations of the successive solutions of the HJB equation and the Kolmogorov Fokker Planck equation.

The Master Equation Equation

$$\partial_t U + H^*(t,x,\mu,U(t,\cdot,\mu)) + \sum_{x'\in E} h^*(t,\mu,U(t,\cdot,\mu))(x')\frac{\partial U(t,x,\mu)}{\partial \mu(\{x'\})} = 0,$$

where the $\mathbb{R}^E$-valued function h^* is defined on $[0,T]\times\mathcal{P}(E)\times\mathbb{R}^E$ by:

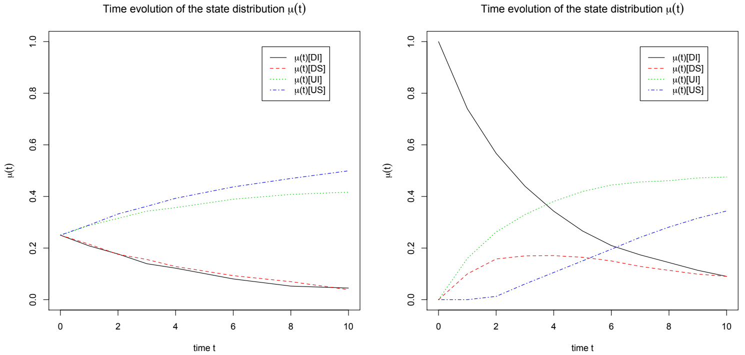
$$\begin{aligned} h^*(t,\mu,u) &= \int_E \lambda_t\big(x,\,\cdot\,,\mu,\hat\alpha(t,x,\mu,u)\big)\, d\mu(x) \\ &= \sum_{x\in E} \lambda_t\big(x,\,\cdot\,,\mu,\hat\alpha(t,x,\mu,u)\big)\mu(\{x\}). \end{aligned}$$

System of Ordinary Differential Equations (**ODEs**)

If and when the Master equation is solved

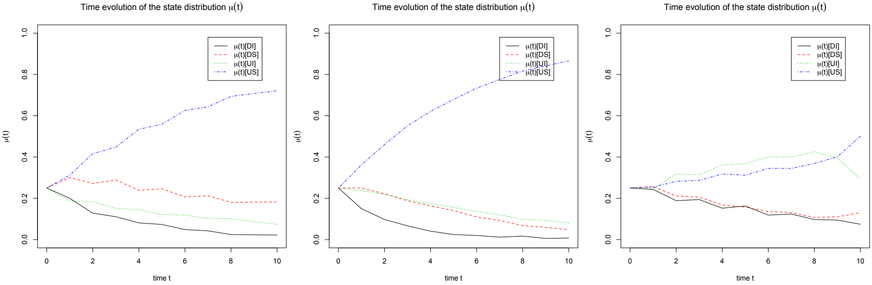
$$\partial_t \mu_t(\{x\}) = h^*(t,\mu_t,U(t,\cdot,\mu_t))(x)$$

μ_t EVOLUTION FROM THE MASTER EQUATION



As before, we used the initial conditions μ_0: $\mu_0 = (0.25, 0.25, 0.25, 0.25)$ in the left and $\mu_0 = (1, 0, 0, 0)$ on the right.

IN THE PRESENCE OF A MAJOR (HACKER) PLAYER



Time evolution in equilibrium, of the distribution μ_t of the states of the computers in the network for the initial condition μ_0: $\mu_0 = (0.25, 0.25, 0.25, 0.25)$ when the major player is not rewarded for its attacks, i.e. when $f_0(\mu) \equiv 0$ (leftmost pane), in the absence of major player and $v = 0$ (middle plot), and with $f_0(\mu) = k_0(\mu(\{UI\}) + \mu(\{DI\}))$ with $k_0 = 0.05$ (rightmost plot).

PoA Bounds for Continuous Time MFGs

Price of Anarchy Bounds

compare Social Welfare for NE to what a Central Planer would achieve

Koutsoupias-Papadimitriou

Usual **Game Model** for Cyber Security

- Zero-Sum Game between **Attacker** and **Network Manager**
- Compute Expected Cost to Network for Protection

MFG Model for Cyber Security

- Let the individual computer owners take care of their security
- Hope for a Nash Equilibrium
- Compute Expected Cost to Network for Protection

How much worse the NE does is the PoA

PoA Bounds for Continuous Time MFGs with Finite State Spaces

$X_t = (X_t^1, \cdots, X_t^N)$ state at time t, with $X_t^i \in \{e_1, \cdots, e_d\}$

- Use distributed feedback controls, for state to be a continuous time Markov Chain
- Dynamics given by Q-matrices $(q_t(x, x'))_{t \geq 0, x, x' \in E}$
- Empirical measures

$$\mu_x^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x^i} = \sum_{\ell=1}^{d} p_\ell \delta_{e_\ell}$$

  where $p_\ell = \#\{i;\ 1 \leq i \leq N,\ x^i = e_\ell\}/N$ is the proportion of elements x^i of the sample which are equal to e_ℓ.

- **Cost Functionals**

  Player i minimizes:

$$J^i(\boldsymbol{\alpha}^1, \cdots, \boldsymbol{\alpha}^N) = \mathbb{E}\Big[\int_0^T f(t, X_t^i, \mu_{X_t^{-i}}^{N-1}, \alpha_t^i)\, dt + g(X_T^i, \mu_{X_T^{-i}}^{N-1})\Big],$$

Social Cost

If the N players use distributed Markovian control strategies of the form $\alpha_t^i = \phi(t, X_t^i)$ we define the **cost (per player) to the system** as the quantity $J_\phi^{(N)}$

$$J_\phi^{(N)} = \frac{1}{N}\sum_{j=1}^N J^i(\boldsymbol{\alpha}^1, \cdots, \boldsymbol{\alpha}^N)$$

In the limit $N \to \infty$ the social cost should be

$$\begin{aligned}
\lim_{N\to\infty} J_\phi^{(N)} &= \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^N J^i(\boldsymbol{\alpha}^1, \cdots, \boldsymbol{\alpha}^N) \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^N \mathbb{E}\left[\int_0^T f(t, X_t^i, \mu_{X_t}^N, \phi(t, X_t^i))\, dt + g(X_T^i, \mu_{X_T}^N)\right], \\
&= \lim_{N\to\infty} \mathbb{E}\left[\int_0^T < f(t, \cdot, \mu_{X_t}^N, \phi(t, \cdot)), \mu_{X_t}^N > dt + < g(\cdot, \mu_{X_T}^N), \mu_{X_T}^N > \right],
\end{aligned} \tag{3}$$

if we use the notation $< \varphi, \nu >$ for the integral $\int \varphi(z)\nu(dz)$.
Now if $\mu_{X_t}^N$ converge toward a deterministic μ_t, the social cost becomes:

$$SC_\phi(\boldsymbol{\mu}) = \int_0^T < f(t, \cdot, \mu_t, \phi(t, \cdot)), \mu_t > dt + < g(\cdot, \mu_T), \mu_T >,$$

Asymptotic regime $N = \infty$

Two alternatives

- ϕ is the optimal feedback function for a **MFG equilibrium** for which $\boldsymbol{\mu}$ is the equilibrium flow of statistical distributions of the state, in which case we use the notation SC^{MFG} for $SC_\phi(\boldsymbol{\mu})$;

$$\mathbb{E}\Big[\int_0^T f(t, X_t, \mu_t, \phi(t, X_t))dt + g(X_T, \mu_T)\Big]$$
$$= \int_0^T < f(t, \cdot, \mu_t, \phi(t, \cdot)), \mathcal{L}(X_t) > dt + < g(\cdot, \mu_T), \mathcal{L}(X_T) >$$
$$= SC_\phi(\boldsymbol{\mu}) = SC^{MFG} \qquad \text{in equilibrium}$$

- ϕ is the feedback (**chosen by a central planner**) minimizing the social cost $SC_\phi(\boldsymbol{\mu})$ **without** having to be an MFG Nash equilibrium, in which case we use the notation SC^{MKV} for $SC_\phi(\boldsymbol{\mu})$;

$$\hat{\phi} = \arg\inf_\phi \int_0^T < f(t, \cdot, \mu_t, \phi(t, \cdot)), \mu_t > dt + < g(\cdot, \mu_T), \mu_t >$$

where μ_t satisfies Kolmogorov-Fokker-Planck forward dynamics

PoA: Social Cost Computation

Minimize

$$\int_0^T < f(t,\cdot,\mu_t,\phi(t,\cdot)),\mu_t > dt + < g(\cdot,\mu_T),\mu_t >$$

under the dynamical constraint

$$\partial_t \mu_t(\{x\}) = [L_t^{\mu_t,\phi(t,\,\cdot\,),\dagger}\mu_t](\{x\}) = \sum_{x'\in E} \mu_t(\{x'\})\lambda(t,x,x',\mu_t,\phi(t,x)), \qquad x\in E,$$

ODE in the d-dimensional **probability simplex** $\mathcal{S}_d \subset \mathbb{R}^d$!!!

Hamiltonian $H(t,\mu,\varphi,\phi)$ by

$$H(t,\mu,\varphi,\phi) =< \varphi,[L_t^{\mu,\phi,\dagger}\mu] > + < f(t,\,\cdot\,,\mu,\phi(\cdot)),\mu >$$
$$=< L_t^{\mu,\phi}\varphi + f(t,\,\cdot\,,\mu,\phi(\cdot)),\mu >$$

minimized Hamiltonian:

$$H^*(t,\mu,\varphi) = \inf_{\phi\in\tilde{A}} H(t,\mu,\varphi,\phi).$$

Assume infimum is attained for a unique $\hat{\phi}$:

$$H^*(t,\mu,\varphi) = H(t,\mu,\varphi,\hat{\phi}(t,\mu,\varphi)) =< L_t^{\mu,\hat{\phi}(t,\mu,\varphi)}\varphi + f(t,\,\cdot\,,\mu,\hat{\phi}(t,\mu,\varphi)(\cdot)),\mu > .$$

HJB equation

$$\partial_t v(t,\mu) + H^*(t,\mu,\frac{\delta v(t,\mu)}{\delta\mu}) = 0, \qquad v(T,\mu) =< g(\cdot,\mu),\mu > .$$

Remarks on Derivatives w.r.t. Measures

Standard identification

$$\mathcal{P}(E) \ni \mu \leftrightarrow \boldsymbol{p} = (p_1, \cdots, p_d) \in \mathcal{S}_d$$

via $\mu \leftrightarrow \boldsymbol{p} = (p_1, \cdots, p_d)$ with $p_i = \mu(\{e_i\})$ for $i = 1, \cdots, d$ i.e. $\mu = \sum_{i=1}^d p_i \delta_{e_i}$

- $\delta v / \delta \mu$ when v is defined on an open neighborhood of the probability simplex $\mathcal{S}_d$.
- $\partial v(t, \mu) / \partial \mu(\{x'\})$ is the derivative of v with respect to the weight $\mu(\{x'\})$.

Important Remark $L_t^{\mu,\phi} \varphi$ does not change if we add a constant to the function φ, so does $\hat{\phi}(t, \mu, \varphi)$.

Consequence (for numerical purposes):

$$\partial_t v(t, \mu) + H^* \Big(t, \mu, \big(\frac{\partial v(t, \mu)}{\partial \mu(\{x'\})} - \frac{\partial v(t, x, \mu)}{\partial \mu(\{x\})}\big)_{x \in E}\Big) = 0.$$

We can identify

$$\frac{\partial v(t, \mu)}{\partial \mu(\{x'\})} - \frac{\partial v(t, \mu)}{\partial \mu(\{x\})},$$

for $x' \neq x$, with the partial derivative of $v(t, \cdot)$ with respect to $\mu(\{x'\})$ whenever $v(t, \cdot)$ is regarded as a smooth function of the $(d-1)$ tuple $(\mu(\{x'\}))_{x' \in E \setminus \{x\}}$, which we can see as an element of the $(d-1)$-dimensional domain

$$\mathcal{S}_{d-1,\leq} = \{(p_1, \cdots, p_{d-1}) \in [0,1]^{d-1} : \sum_{i=1}^{d-1} p_i \leq 1\}$$

The Example of Corporate Bonds

- **Major Player** = bond issuer
  - Bond is **Callable**
  - Major Player (issuer) chooses a **stopping time** to
    - pay-off the investors
    - stop coupon payments to the investors
    - refinance his debt with better terms
- **Minor Players** = field of investors
  - Bond is **Convertible**
  - Each Minor Player (investor) chooses a **stopping time** at which to
    - convert the bond certificate in a fixed number (***conversion ratio***) of stock shares
    - if and when **owning the stock** is more profitable
    - creating **Dilution** of the stock