Mean-Field-Type Game Theory

Tembine

August 25, 2014

NEW YORK UNIVERSITY
Introduction
- What is a Mean-Field-Type Game?
- Static Mean-Field-Type Games
- Mean-Field Learning

From Micro to Macro

Mean-Field-Type Games
- Stochastic Maximum Principle
- Master System
- Cooperative Mean-Field-Type Games

Nonasymptotic Mean-Field-Type Games
- Auction with Heterogeneous Bidders
Game Theory

Strategic-form games:

- set of players
- set of actions for each player
- objective (pay-off) functions for each player

In this classical setup, a payoff function is determined by the realized action profile.

Payoff: \( r_j(a_1, \ldots, a_n), \ j \in \{1, 2, \ldots, \} \)

Equilibrium: \( r_j(a_1^*, \ldots, a_n^*) = \max_{a_j} r_j(a_1^*, \ldots, a_{j-1}^*, a_j, a_{j+1}^*, \ldots, a_n^*), \)

\( j \in \{1, 2, \ldots, \} \)
State-dependence

State-dependent strategic-form games:

- set of states,
- set of players,
- set of actions for each player (may depend on the state),
- objective (pay-off) functions for each player.

In this classical setup, a payoff function is determined by the realized state-action profile.

Payoff: \( r_j(s, a_1, \ldots, a_n), \ j \in \{1, 2, \ldots, \} \)

Mean-Field-Type Game

Any game in which the payoffs and/or the state dynamics coefficient functions involve not only the state and actions profiles but also the distributions of state-action pairs (and/or its marginals: distribution of actions or distributions of states).

— Tembine 2010

This may not be of von Neumann type.
The number of players/agents/participants is not necessarily large.

Example: Mean-Variance optimization, Failure probability and risk minimization
References: Dynamic Mean-Field Games

Main assumptions of mean-field games

- **Infinite** number of players
- **Indistinguishability**.
- An individual player has no influence on the aggregate

Questions:
- Why do we need these assumptions?
- Where are they used?
In many of real world decision-making systems, the number of decision-makers may be relatively large but still finite. There is no continuum of players.

- How to extend the mean-field game theory ideas to finite regime?
- Preliminary answers in H. Tembine, Nonasymptotic Mean-Field Games, IEEE Transactions on Cybernetics, 2014
Re-adjustment needed: non-negligible effect

- Can we drop the assumption that each individual player has a negligible influence on the aggregate (mean-field)?
- Answer: yes [this talk!]
Mean-Field-Type Game but not MFG

Mean-Field-Type Game is not a standard mean-field game.

- Here, a single player may have a big influence of the distribution (mean-field).
Mean-Field-Type Game but not MFG

Mean-Field-Type Game is different than multi-population MFG:
- In multi-population games, each representative agent has its own strategies interacting with inter-and-intra population players.
Mean-Field-Type Games: Features

- In mean-field-type games, players can be atomic.
- It does not require indistinguishability property.
- It is suitable for few number of players as well as for large number of players.
Static mean-field-type games

Aggregative Game

The payoff function of each player depends on its own-action and an aggregative term of the other actions.


The number of players/agents/participants is not necessarily large.
Public Good Provisioning

How to provide public goods? ref: Bergstrom, T., L. Blume, and H. Varian (1986): On the

**Voluntary Contribution**

\[ r_j(a_1, \ldots, a_n) = \bar{r}_j(a_j, \text{Good}) = w_j + [h(G) - a_j] \mathbb{1}_{G > 0} \]


**Financing Public Good by means of Lottery**

\[ r_j(a_1, \ldots, a_n) = \bar{r}_j(a_j, \text{Good}) = w_j + [h(G) + R \frac{a_j}{\sum_{i=1}^{n} a_i} - a_j] \mathbb{1}_{G > 0} \]

An Efficient Probabilistic Reward Scheme

Morgan’s scheme is INEFFECTIVE

\[ \bar{r}_j(a_j, G) = w_j + \left[ h(G) + R \sum_{i=1}^{n} \frac{a_j}{a_i} - a_j \right] \mathbb{1}_{G>0} \]

— ref. Morgan 2000

A modification: EFFECTIVE

\[ r_j = w_j + \left[ h_j(G) + R \sum_{i=1}^{n} \frac{d_j a_j^\alpha}{d_i a_i^\alpha} - a_j \right] \mathbb{1}_{G>0} \]
An Efficient Probabilistic Reward Scheme

- For any reward $R \geq \frac{4m^* \sigma}{(1-\sigma)^2}$, $\sigma = \frac{h_i'(m)-1}{h_j'(m)-1} > 0$, $m^* \in \arg\max[h(m) - m]$, there exists a design parameter $(d_i)$ such that the "new" lottery based scheme provides the global optimum level of contribution in the public good.

- If $h_j$ are identical, one cannot achieve the global optimum. In particular, the winning probability used in Morgan (2000) cannot achieve a global optimum.
Rent-seeking game

\[ r_j(s, a) = \mathbb{1}_{\{G \neq 0\}} \left[ \frac{c_n(s)}{G} \right] \]

where

- \( c_n(s) \) is available capacity
- \( G \) is total input
- \( a_j \) is cost for action
- \( p_n \) is price

Rent-seeking game with discontinuous payoff function

---

A class of Resource Sharing Games

Consider $n$ users in a cloud network. User $j$ has a demand $a_j \in \mathbb{R}_+$. The payoff is

$$r_j(s, a) = \left[ c_n(s) \frac{a_j^\alpha}{\sum_{j' = 1}^{n} a_{j'}^\alpha} - p_n a_j \right] 1_{\{\sum_{j' = 1}^{n} a_{j'}^\alpha \neq 0\}},$$

where $p_n > 0$ the unit price for resource utilization.
Resource sharing games $\alpha \in (0, 1]$:

Basic results I

- There is a unique Nash Equilibrium (all the users are active)
  $$a^*_{NE} = \alpha \frac{(n-1)c_n}{n^2p_n}$$

- Optimal price: $$p^*_n = \alpha \frac{(n-1)}{n}$$

- $\forall 2 \leq n < +\infty$, the Nash equilibrium is not stable in the sense of Schaffer 1988.

- The Mann-based best reply dynamics converges to Nash equilibrium.

- The game is submodular game in $(a, \bar{m})$. 
Resource sharing games $\alpha \in (0, 1]$:

**Basic results II**

- there is a unique mean-field equilibrium $\bar{m}^* = \alpha \frac{c}{p}$. The mean-field equilibrium is an ESS in the sense Maynard Smith & Price (1973).
- the mean-field equilibrium is an $\frac{1}{n}$—Nash equilibrium.
- if $\alpha = 1$, the best reply dynamics exhibits a limit cycle. The limit cycle can be eliminated via Mann’s learning scheme.
Mean-Field Learning

Model-based mean-field response (myopic):

$$a_{j,t+1} \in \arg \max_{a_j} \bar{r}_j(a_j, m_t), \quad m_t = \text{aggregate}$$

Model-based mean-field learning

$$m_{t+1} \in \arg \max_b \bar{r}(b, m_t).$$

Note: No need to know all the actions of the other players.
How to estimate the payoff? one-step memory

ref: joint with Zhu, Başar: CDC 2010

Observe that the average payoff per action \(l\),

\[ F_{j,T}(l) = \frac{1}{T} \sum_{t=1}^{T} d_{j,t} \mathbb{1}_{\{a_j,t=l\}} \]

satisfies the recursive equation

\[
F_{T+1}(l) = \frac{1 - \frac{1}{T+1}}{T} F_T(l) + \frac{1}{T+1} d_{j,T+1} \mathbb{1}_{\{a_j,T+1=l\}}
\]

\[
\hat{r}_{j,t+1}(l) = \hat{r}_{j,t}(l) + \mu t \begin{pmatrix}
\text{New Payoff} \\
\text{Old Value}
\end{pmatrix}
\]

\[
\begin{pmatrix}
\text{Learning Rate} \\
\text{New Payoff}
\end{pmatrix}
\begin{pmatrix}
\hat{r}_{j,t} - \hat{r}_{j,t}(l)
\end{pmatrix}
\]
Distributed Strategic Learning: CODIPAS

CODIPAS: COmbined fully DIstributed PAyoff and Strategy learning\(^2\).

- The consequences influence the behaviors of the players.
- The behaviors influence the outcomes.

The model is flexible enough to incorporate: random updates, variable set of interacting players, random states, noisy and delayed measurements.

Pursuit CODIPAS

\[
\begin{align*}
P[a_{j,t+1} = l] &= x_{j,t+1}(l) \sim \delta \arg \max_k \hat{r}_{j,t}(k) \\
\hat{r}_{j,t+1}(l) &= \hat{r}_{j,t}(l) + \frac{\mu_t}{\theta_{l,t}} \mathbb{1}_{\{a_{j,t}=l\}}(r_{j,t} - \hat{r}_{j,t}(l)) \\
\theta_{l,t+1} &= \theta_{l,t} + \mathbb{1}_{\{a_{j,t}=l\}}.
\end{align*}
\]

Imitative CODIPAS

\[
\begin{align*}
P[a_{j,t+1} = l] &= x_{j,t+1}(l) \sim \frac{x_{j,t}(l)(1+\lambda_t)\hat{r}_{j,t}(l)}{\sum_{k \in A} x_{j,t}(k)(1+\lambda_t)\hat{r}_{j,t}(k)} \\
\hat{r}_{j,t+1}(l) &= \hat{r}_{j,t}(l) + \frac{\mu_t}{\theta_{l,t}} \mathbb{1}_{\{a_{j,t}=l\}}(r_{j,t} - \hat{r}_{j,t}(l)) \\
\theta_{l,t+1} &= \theta_{l,t} + \mathbb{1}_{\{a_{j,t}=l\}}.
\end{align*}
\]
3 players, 2 actions:

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>up</td>
<td>0,0,0</td>
<td>0,1,0*</td>
</tr>
<tr>
<td>down</td>
<td>1,0,0*</td>
<td>0,0,1*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>up</td>
<td>0,0,1*</td>
<td>1,0,0*</td>
</tr>
<tr>
<td>down</td>
<td>0,1,0*</td>
<td>0,0,0</td>
</tr>
</tbody>
</table>

Figure: A three player road traffic game.
Convergence to one of the global optima
Satisfactory solutions

- Players, Action Spaces, Satisfaction Criterion.
- Satisfaction à la Simon$^3$: $r_j(a_1, \ldots, a_n) \geq \gamma^*_j, \quad j \in \{1, 2, \ldots\}$
- Pure satisfactory solution: $a^*$ such that $r_j(a^*_1, \ldots, a^*_n) \geq \gamma^*_j, \quad j \in \{1, 2, \ldots\}$

Goals:
- Find a satisfactory solution
- Develop speedup learning techniques

---

Why satisfactory solution?

- Games in satisfaction form can be mapped into constrained games or Generalized games with normal form.
- How to play a normal form game with coupled constraint?
Mean-field-type games in satisfaction form

- \( \gamma_j^* = \gamma^* \), \( A \subset \mathbb{R} \) nonempty, closed, convex
- \( \phi : A \times \mathcal{P}(A) \mapsto \mathbb{R}, \ r_j(a) = \phi \left( a_j, \frac{1}{n-1} \sum_{i \neq j} \delta a_i \right) = \phi(a_j, \tilde{m}) \)

Existence and well-posedness:
- Feasibility condition: \( \exists \ \epsilon \geq 0, \ \gamma^* + \epsilon \in r(A^n) \).
Introduction
From Micro to Macro
Mean-Field-Type Games
Nonasymptotic Mean-Field-Type Games

Mean-field learning for satisfactory solutions

How to find a satisfactory solution?\(^4\)

- Aggregative game (Selten 1980): \(\phi(a_j, E[\tilde{m}])\)
- Model-free mean-field learning scheme (Banach-Picard): \(\tilde{m}_{t+1} = f(\tilde{m}_t), \quad f(x) = \text{proj}_A[x \frac{\gamma^*}{\phi(x,x)}]\)
- Interesting test with \(f(x) = \frac{1}{x}, \quad x_0 \in A = [\frac{1}{4}, 4]\)
- Mann iterates:
  \(\tilde{m}_{t+1} = \lambda f(\tilde{m}_t) + (1 - \lambda)\tilde{m}_t, \quad 0 < \lambda < \bar{\lambda}, \quad \tilde{m}_0 \in A\)
- Ishikawa iterates:
  \(\tilde{m}_{t+1} = \lambda_t f(y_t) + (1 - \lambda_t)\tilde{m}_t, \quad \tilde{y}_t = \mu_t f(\tilde{m}_t) + (1 - \mu_t)\tilde{m}_t.\)

---

\(^4\) joint with Tempone, Vilanova: Mean Field Learning for Satisfactory Solutions, CDC’13.
Convergence time

\[ T_\eta = \inf \{ t > 0 \mid d(\bar{m}_t, \text{fix}(f)) \leq \eta \} \]

- **Banach/Picard iterates:** \( T_\eta = \ln \left[ \frac{d(\bar{m}_0, \text{fix}(f))}{\eta(1-\alpha_1)} \right] \cdot \frac{1}{\ln(\frac{1}{\alpha_1})} \) iterations

- **Ishikawa iterates:** \( d(\bar{m}_t, f(\bar{m}_t)) \leq 2 \frac{d(\bar{m}_0, \text{fix}(f))}{\sqrt{\pi} \sum_{k=1}^{t} \lambda_k (1-\lambda_k)} \)

\[ T_\eta = \frac{14d(\bar{m}_0, \text{fix}(f))^2}{\pi \eta^2} \]
How to speedup mean-field learning algorithms?

<table>
<thead>
<tr>
<th>Derivative – based</th>
<th>Derivative – free</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Newton</em></td>
<td><em>Reverse Ishikawa</em></td>
</tr>
<tr>
<td><em>Halley (cubic)</em></td>
<td><em>Secant</em></td>
</tr>
<tr>
<td><em>Combined Newton (twice)</em></td>
<td><em>Aitken</em></td>
</tr>
<tr>
<td><em>Householder</em></td>
<td><em>Steffensen</em></td>
</tr>
<tr>
<td></td>
<td><em>Combined Steffensen</em></td>
</tr>
</tbody>
</table>
Derivative-based speedup methods

- Newton: $\tilde{m}_{t+1} = \tilde{m}_t - \frac{g(\tilde{m}_t)}{g'(\tilde{m}_t)}$
  quadratic convergence

- Halley: $\tilde{m}_{t+1} = \tilde{m}_t - \frac{2g(\tilde{m}_t)g'(\tilde{m}_t)}{2[g'(\tilde{m}_t)]^2 - g(\tilde{m}_t)g''(\tilde{m}_t)}$
  cubic convergence

- 2-points Newton: $\bar{y}_t = \tilde{m}_t - \frac{g(\tilde{m}_t)}{g'(\tilde{m}_t)}$, $\tilde{m}_{t+1} = \bar{y}_t - \frac{g(\bar{y}_t)}{g'(\bar{y}_t)}$
  fourth order convergence

- Householder: $\tilde{m}_{t+1} = \tilde{m}_t + o \frac{(1/g)^{(o-1)}(\tilde{m}_t)}{(1/g)^{(o)}(\tilde{m}_t)}$
  $(o + 1)$-th order convergence
Derivative-free speedup method

- Secant method:
  \[
  \tilde{m}_{t+1} = \tilde{m}_t - \frac{g(\tilde{m}_t)(\tilde{m}_t - \tilde{m}_{t-1})}{g(\tilde{m}_t) - g(\tilde{m}_{t-1})}
  \]  
  order: 1.6

- Aitken\(^5\):
  \[
  \tilde{y}_t = \tilde{m}_t - \frac{(\tilde{m}_{t+1} - \tilde{m}_t)^2}{\tilde{m}_{t+2} - 2\tilde{m}_{t+1} + \tilde{m}_t},
  \]

- Steffensen: \(m_0, m_1 = g(m_0), m_2 = g(m_1)\),
  \[
  \tilde{y}_t = \tilde{m}_t - \frac{g(\tilde{m}_t)^2}{g(\tilde{m}_t + g(\tilde{m}_t)) - g(\tilde{m}_t)},
  \]

- Reverse Ishikawa:
  \[
  \tilde{m}_{t+1} = \text{proj}_A[\lambda_t f(\tilde{m}_t) + (1 - \lambda_t)\tilde{m}_t], \quad 1 < \lambda_t < 2, \quad \tilde{m}_0 \in A
  \]

Speedup mean-field learning

Different mean-field learning schemes, a comparative example.

- State-of-the-art: accuracy = $1.25 \times 10^{-5}$ after 50 iterations
- Reverse Ishikawa mean field learning: accuracy = $10^{-8}$ after 25 iterations
- Steffensen mean field learning: accuracy = $10^{-15}$ after 6 iterations

Accuracy in 5 iterations:
- State-of-the-art error = 2.11
- Our algorithm v1 error = 0.13
- Our algorithm v2 error = $10^{-15}$
Cell formation in 4G

- We consider a geographical area served by a heterogeneous network.
- The network consists of $K$ spatially and spectrally coexisting tiers.
- $M = \sum_{k=1}^{K} M_k$ BSs are randomly deployed to satisfy the traffic demand of the customers and covers the total area $(x_j, y_j, z_j)$ the spatial coordinates of BS $j$.
- $n$ users are also randomly placed following a certain user distribution function $m(x, y)$. 
4G and small-cell co-existence

Problem

Find the cell association \((C_j)_j\) such that

\[
F(C_1, \ldots, C_M) = \sum_{j=1}^{M} N_j \int_{C_j} \bar{P}_j(x, y)m(x, y) \, dx \, dy \leq F(\tilde{C}_1, \ldots, \tilde{C}_M)
\]

where \((\tilde{C}_1, \ldots, \tilde{C}_M)\) is a partition of the domain where the users are distributed according to the probability density \(m\).

\(C_j = \{(x, y) \mid \text{user located at (x, y) is served by BS } j\}\)

\(N_j = N \int_{C_j} m(x, y) \, dx \, dy\)
Explicit characterization of the solution

**Cell characterization**

\[ C_j = \left\{ (x, y) \mid N_j \bar{P}_j(x, y)m(x, y) + \int_{C_j} \bar{P}_j(x, y)m(x, y) \, dx\, dy \leq N_i \bar{P}_i(x, y)m(x, y) + \int_{C_i} \bar{P}_i(x, y)m(x, y) \, dx\, dy, \ i \neq j \right\} \]

Here: \( P_j(x, y) = C \cdot d(BS_j, (x, y, 0))^{\nu/10} \).

We use mean-field learning algorithm to solve the above fixed-point system between \( N_j \) and \( C_j \).
Numerical examples

Figure: Uniform distribution over $5 \times 5 \text{ km}^2$ for 500 users.
Figure: Users are concentrated in downtown hotspot areas and the decreases as we move away from the center (0,0). Gaussian centered with variance 1.15.
Figure: Two hotspots where users are concentrated: Two neighboring residential areas. 16 stations instead of 25.
Figure: Multi-tier network: one macrocell and 4 small cells with 150 users distributed as gaussian centered with variance 0.5.
Figure: Multi-tier network: one macrocell and 8 small cells with 150 users distributed as gaussian centered with variance 0.5.
State of the art:

- **Bianchi (2000)**\(^6\): Markov chain, i.i.d assumption and stationary regime,

- **Benaim & Le Boudec (2010)**: general mean field model that touches upon the veracity of Bianchi’s throughput formula.

- **Duffy (2010)**: overview on the conjectures and non-validity discussions

- **Cho et al. (2010)**: provides an explicit example where the decoupling assumptions is non-valid in stationary regime.

**Question:** What is the validity domain?

---

\(^6\) G. Bianchi, Performance Analysis of the IEEE 802.11 Distributed Coordination Function, IEEE JSAC, vol. 18, N. 3, March 2000
CSMA protocols

CSMA/CA protocol in single cell IEEE 802.11 network with $n$ wireless nodes. Assumptions:

- Each user has a saturated queue and can hear each other.
- Time is discrete (slotted) and time space is the set of natural numbers.
- Assume that the backoff slots are synchronized and only collisions are responsible of packet loss. Each user has its backoff which has a maximum of $K$ states denoted by $0, 1, \ldots, K$. 
State evolution

The backoff process in IEEE 802.11 is governed by a decision process

- every user in backoff state $s$ attempts transmission with probability $\phi(t, s, n)$ for every time-slot;
- if it succeeds, $s$ changes to 0;
- otherwise, $s$ changes to $(s + 1) \mod (K + 1)$.

- The backoff state process of each node is influenced by the decision of all the active users (to transmit or not).
Interactive Markov decision processes

- 1-step-Markov state-feedback strategy $\phi(t, x, n)$.
- The backoff state process is Markovian.

$M_t^n(s) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{s_j^n, t = s\}}$

**Q0:** What is the probability that none of the other nodes transmits?
**Q0\*:** What is the probability that at least one of the other nodes transmits?
**Q1:** What is the probability for a generic user to experiment a collision at a given state $s$?
Transition probabilities

for $x = 0$:

\[
q^n_{0,1} = \phi(t, 0, n)(1 - \mu^n_0(t, \phi(t, \ldots, n), m(t))) \\
q^n_{0,0} = 1 - \phi(t, 0, n) + \phi(t, 0, n)\mu^n_0(t, \phi(t, \ldots, n), m(t))
\]

$s \geq 2$, $q^n_{0,s} = 0$

for $s \in \{1, \ldots, K - 1\}$:

\[
q^n_{s,s+1} = \phi(t, s, n)(1 - \mu^n_s(t, \phi(t, \ldots, n), m(t))) \\
q^n_{s,s} = 1 - \phi(t, s, n) \\
q^n_{s,0} = \phi(t, s, n)\mu^n_s(t, \phi(t, \ldots, n), m(t))
\]

$s' \notin \{0, s, s + 1\}$, $q^n_{s,s'} = 0$

for $s = K$:

\[
q^n_{K,0} = \phi(t, K, n) \\
q^n_{K,K} = (1 - \phi(t, K, n)), s \notin \{0, K\}, q^n_{K,s} = 0
\]

$\mu^n_s(\ldots, \ldots, \ldots) := (1 - \phi(t, s, n))^{nm_s(t)-1} \prod_{s' \neq s} (1 - \phi(t, s', n))^{nm_{s'}(t)}$
\[ \phi = \frac{a_{t,s}}{\xi(n)}, \quad \frac{n}{\xi(n)} \longrightarrow \delta > 0. \quad M^n_t(s) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}\{s^n_{j,t}=s\} \]

**ODE**

For any \( \epsilon > 0, \, T < +\infty \),

\[ \lim_{n} \mathbb{P} \left( \sup_{t \in [0, T]} \| M^n_t - m_t[m_0] \| > \epsilon \right) = 0, \]

where \( m_t \) is the unique solution of the ODE \( \dot{m}_t = f(a_t, m_t) \)
starting from \( m_0 \in \Delta(S) \).

\[ f : = \lim_{n} n \langle m, (Q^n - I) \rangle \quad (\text{Drift}). \]
Mean field convergence: Multi-Class CSMA

ODE

\[ M_{n,\theta}^{s}(t/\delta_n) \] converges [in probability] to \( m_{\theta}^{s}(t) \) which is the solution of the system of ODEs:

\[
\begin{align*}
\frac{dm_{0}^{\theta}}{dt}(t) &= \bar{v}^{\theta}(t)(1 - \gamma(t)) - v_{0}^{\theta}(t)m_{0}^{\theta}(t) + v_{K}^{\theta}(t)m_{K}^{\theta}(t)\gamma(t) \\
\frac{dm_{s}^{\theta}}{dt}(t) &= v_{s-1}^{\theta}(t)m_{s-1}^{\theta}(t)\gamma(t) - v_{s}^{\theta}(t)m_{s}^{\theta}(t), \quad s \in \{1, ..., K^{\theta}\}, \\
\theta &\in \Theta
\end{align*}
\]

where \( v_{s}^{\theta} \) is a feedback strategy, \( \bar{v}^{\theta}(t) = \sum_{s=0}^{K^{\theta}} v_{s}^{\theta}(t)m_{s}^{\theta}(t) \) and \( \gamma(t) = 1 - e^{-\sum_{\theta} \bar{v}^{\theta}(t)} \) for \( \theta \in \Theta \).

Here \( \bar{v}^{\theta}(t) \) is the mean field limit of the average attempt rate and \( \gamma(t) \) is that of the collision probability.
Fixed-point equation

\[ \gamma = 1 - e^{-\bar{v}} \]

\[ \bar{v} = \frac{\sum_s \gamma_s^s}{\sum_s \frac{\gamma_s}{v_s}} \]
$\Theta = \{1, 2\}$. We plot the trajectories of $m^1, m^{17}, m^0$ for the strategy

$$(u_{1,0}, u_{1,1}, \ldots, u_{1,20}) = \left( \frac{1}{2400}, \frac{1}{480}, \frac{1}{40}, \frac{\tau}{40}, \ldots, \frac{\tau^{18}}{40} \right),$$

and

$$(u_{2,0}, u_{2,1}, \ldots, u_{2,20}) = \left( \frac{1}{3840}, \frac{1}{64}, \frac{1}{64}, \ldots, \frac{1}{64} \right)$$

with $\tau = \frac{4}{5}$. 
The challenging example?
Q4: Is it possible to stabilize the system?

Answer: The system can be stabilized for example by only changing the strategy $u : t$

$$(u_{1,0}, u_{1,1}, \ldots, u_{1,20}) = \left( \frac{1}{2400}, \frac{1}{480}, \frac{1}{40}, \frac{\tau}{40}, \ldots, \frac{\tau^{18}}{40} \right),$$

and

$$(u_{2,0}, u_{2,1}, \ldots, u_{2,20}) = \left( \frac{1}{3840}, \frac{1}{64}, \frac{1}{64}, \ldots, \frac{1}{64} \right)$$

with $\tau = \frac{1}{2}$ and the linearly stability property of the stationary point can be easily checked.
The challenging example (cont’d)?
Q4: What about the validity of the performance metrics at the limit?
In presence of limit cycle: \( \liminf_{t \to +\infty} \frac{1}{t} \sum_{t'=1}^{t} r_n(t') \),
Violation of propagation of chaos property:

\[
\lim_{t} \frac{1}{t} \int_{0}^{t} \left( \prod_{l=1}^{k} m_{s_l}(t') \right) \, dt' \neq \prod_{l=1}^{k} \left( \lim_{t} \frac{1}{t} \int_{0}^{t} m_{s_l}(t') \, dt' \right)
\]
Carrier Sense Multiple Access: summary

- The backoff processes can be correlated even in asymptotic regime.
- Presence of limit cycle for some strategies,
- The validity domain of fixed-point formula and throughput formula is restricted.
- Good news: the domain has at least two elements.
Notations

- Horizon $\mathcal{T}$, $\text{length}(\mathcal{T}) > 0$
- $s(t) \in \mathcal{S} \subseteq \mathbb{R}$ is a 1-dimensional state vector of a generic Player,
- $a(t) \in \mathcal{A}$, is the control action at time $t$ with $\mathcal{A} \subseteq \mathbb{R}^k$, the control action set $\mathcal{A}$ is non-empty,
- $B$ is a standard Brownian motion in $\mathbb{R}$.
- $m(t) = \mathcal{L}(s(t), a(t))$ is the probability distribution of the pair process $s(t), a(t)$ at time $t$. 
Mean-Field Type Game: risk-neutral case

Continuous time: $R_i^*(s_0, m_0) =$

$$
\sup_{a_i} \mathbb{E} \left[ g_i(s(T), m(T)) + \int_0^T r_i(t, s(t), m(t), a(t)) dt \right],
$$

subject to

$$
ds(t) = b(t, s(t), m(t), a(t)) dt + \sigma(t, s(t), m(t), a(t)) dB(t),$$

$s(0) = s_0 \in S \subseteq \mathbb{R}$, $s_0 \sim m_0$

$m(t) = \mathcal{L}(s(t), a(t)),$
Discrete time:

\[
R_i^*(s_0, m_0) = \left\{ \begin{array}{l}
\sup_{a_i} \mathbb{E} \left[ g_i(s(T), m(T)) + \sum_{t=0}^{T-1} r_i(t, s(t), m(t), a(t))dt \right], \\
\text{subject to}
\end{array} \right.
\]

- Kernel\(s(t + 1) \in \cdot \mid t, s(t), m(t), a(t))
- \(s(0) = s_0 \in \mathcal{S} \subseteq \mathbb{R}\),
- \(m(t) = \text{Law}(s(t), a(t))\),
Stochastic maximum principle

SMP: Derive necessary conditions for optimality
- Existence of unique control
- Well-posedness
- Conditions for uniqueness
- Solvability
- Algorithms
- Implementation
Dual Game Variables

- Pontryagin function:
  \[ \hat{H}_i(t, s, m, a, p, q) = r_i + b' p_i + \text{trace} [\sigma' q_i] \]

- Dual variable:
  \[
  u_{i,t} + \hat{H}_i(t, s, m, a^*, u_{i,s}, \frac{\sigma}{2} u_{i,ss}) + \mathbb{E} \left[ \hat{H}_{i,m} \right] = 0, \\
  u_i(T, s) = g_i(s(T), m(T, .)) + \mathbb{E} [g_{i,m}].
  \]
Dual Game Variable

Derive necessary conditions for optimality\(^7\):

\[ R_i(a_i + \epsilon d_i, a_{-i}, m_{a_i + \epsilon d_i}) - R_i(a_i, a_{-i}, m) \leq 0 \]

- As \( \epsilon \) goes to zero, if \( u_i \) solves the dual game equation, then

\[
\lim_{\epsilon \to 0^+} \frac{d}{d\epsilon} R_i = \int_{(t,s) \in T \times S} \hat{H}_{i,a_i}(t, s, m, a, u_{i,s}, u_{i,ss}) m(t, s) ds \, dt
\]

\(^7\)Bensoussan et al. 2013
First-order adjoint system

\[ dp_i = -\alpha_i dt + q_i dB(t) \]  
\[ p_i(T) = g_i, s + \mathbb{E}[g_i, m, s] \]
\[ \alpha_i = \hat{H}_{i,s} + \mathbb{E}[\hat{H}_{i,m,s}] \]
\[ \hat{H}_{i,a_i} = 0 \]

Master System for the Dual Game

**Master System**

\[ U_{i,t} + \hat{H}_i(t, s, m, a, U_{i,s}, \frac{\sigma}{2} U_{i,ss}) + \mathbb{E} \left[ \hat{H}_i, m \right] \]

\[ + \langle U_{i,m}, -\text{div}_s(bm) + \frac{1}{2} \text{trace}(\sigma \sigma' m_{ss}) \rangle = 0 \]  \hspace{0.5cm} (6)

\[ u_i(T, s) = g_i(s(T), m(T,.)) + \mathbb{E} [g_i, m]. \]  \hspace{0.5cm} (7)
Cooperative Games with Coalition Formation Cost

- Two Agents. Get $v(\{12\}, [t_0, T], s_0)$ if they work together.
- Non-cooperation: Agent 1 gets $v(\{1\}, [t_0, T], s_0)$ and Agent 2 gets $v(\{2\}, [t_0, T], s_0)$
- **Cost of coalition formation**: $c(\{12\})$
- A necessary condition for coalition formation is:

  \[
  \text{SURPLUS} = v(\{12\}) - c(\{12\}) - v(\{1\}) - v(\{2\}) \geq 0
  \]

How to share the payoff if the coalition is formed?
Equal allocation: \( \left( \frac{v(\{12\}) - c(\{12\})}{2} , \frac{v(\{12\}) - c(\{12\})}{2} \right) \) may not be sustainable.

Sharing the dividend of cooperation = Surplus:
\[ DoC = v(\{12\}) - c(\{12\}) - v(\{1\}) - v(\{2\}). \]

Agent 1 gets \( v(\{1\}) + \lambda DoC \) and Agent 2 gets \( v(\{2\}) + (1 - \lambda) DoC, \quad \lambda \in (0, 1) \).

For \( \lambda = 1/2 \), one gets the Shapley value (Lloyd Shapley introduced it in 1953. He won in 2012 Nobel Memorial Prize in Economic Sciences “for the theory of stable allocations and the practice of market design.”)
Cooperation Mean-Field Type Game

Grand coalition value

\[ v(\mathcal{N}) = \sup_{a=(a_i)_{i \in \mathcal{N}}} \mathbb{E} \left[ g(s(T), m(T)) + \int_0^T r(t, s(t), m(t), a(t)) dt \right], \]

subject to

\[
\begin{aligned}
& ds(t) = b(t, s(t), m(t), a(t)) dt + \sigma(t, s(t), m(t), a(t)) dB(t), \\
& s(0) = s_0 \in \mathcal{S} \subseteq \mathbb{R}, \ s_0 \sim m_0 \\
& m(t) = \mathcal{L}(s(t), a(t))
\end{aligned}
\]
Towards a Smarter Pollution Control

Pollution is a public bad. There is a need for a smarter control of pollution!


- Two companies/farmers. Each of them produces a good with output $q_i = F_i(a_i)$
- The production of the good emits an amount $a_i$ of pollutants.
- The stock of pollution at time $t$: $s_1(t)$.

$$ds_1 = \left[ \sum_{i=1}^{n_1} a_{1,i} - bs_1 + r_1 \frac{s_1^2}{\epsilon_2 + s_1^2} \right] dt - \sigma_{11} s_1 dB_{11}.$$

- Long-term payoff functional of $i$:

$$R_{1,i} = \mathbb{E} \left[ g(s(T), m(T)) + \int_0^T (\log a_{1,i}(t) - \lambda_1 s_1^2(t)) \ dt \right]$$

Nash equilibrium:

$$a_{1,i} \in \arg \max_{b_i} R_{1,i}(a_{1,1}, \ldots, a_{1,i-1}, b_i, a_{1,i+1}, \ldots, a_{1,n}).$$
Noncooperation leads to higher pollution level

- There is a unique Nash equilibrium.
- The stock of pollution in a noncooperative case would be larger than it would be in a fully cooperative case.
- There is a simple Taxation scheme that drives the noncooperative players to a **Global optimum**.
- However it is difficult to implement a tax that is highly time-varying. Is there a (piecewise) constant tax that leads to global optimum?
Fishing Policy

- Fish Stock: \( ds_2 = r_2[s_2(1 - \frac{s_2}{K(s_1)}) - ks_2a_2]dt + r_2s_2\sigma_{21}dB_{21}(t) - r_2\sigma_{22}s_2^2dB_{22}(t). \)
- Fokker-Planck-Kolmogorov forward equation:
  \[
  \partial_t m_2(t, s) + \partial_s [m_2(t, s)D] - \frac{1}{2} r_2^2 \partial_{ss}^2 (s_2^2 \sigma_{21}^2 + \sigma_{22}^4) m_2(t, s) = 0.
  \]
- Stationary solutions: \( s_2 = 0 \) and
  \[
  m^*_2(s) = \mu s^{2(\frac{1}{2} - 1)} \left(1 + \frac{\sigma_{22}^2}{\sigma_{21}^2} s^2\right)^{-\frac{1}{\sigma_{21}^2} - 1} e^{-\frac{2}{K\sigma_{21}\sigma_{22}}} \arctan(s\frac{\sigma_{22}}{\sigma_{21}}),
  \]
  \( h = ks_2a_2. \)
- Payoff functional of the Fish industry: \( R_2 = \int_0^T [p(h)h - c_2(a_2)]dt. \)
Towards a Smarter Fishing Policy

- **Consequence:** As the number, size and power of fishing boats has grown, an increasing number of commercial fisheries are being fished to the point of collapse.

- **Need:** to move from overfishing to a **smarter fishing strategy**.
A Cooperative Game for Smarter Cities

Smarter cities start with smarter systems, working for the benefit of both residents, industry and the environment.

**Coupled system (**):**

\[
\begin{align*}
    ds_1 &= \left[ \sum_{j=1}^{n} a_{1,j}(t) - bs_1 + r_1 \frac{s_1^2}{\varepsilon^2 + s_1^2} \right] dt - \sigma_{11} s_1 dB_{11}, \\
    ds_2 &= r_2 \left[ s_2 \left( 1 - \frac{s_2}{K(s_1)} \right) - ks_2 a_2 \right] dt + r_2 s_2 \sigma_{21} dB_{21} - r_2 \sigma_{22} s_2^2 dB_{22}.
\end{align*}
\]

Payoff functional: \( R = R_1 + R_2 \)

**Synergy**

For small cost of coalition formation, the coalition between the Fishing Industry/Farmers/Mining Industry is synergetic.
Introduction
From Micro to Macro
Mean-Field-Type Games
Nonasymptotic Mean-Field-Type Games

Stochastic Maximum Principle
Master System
Cooperative Mean-Field-Type Games

Dynamics of the model for $K(p) = 1/(p+1)$

- Red: Pollution $p$
- Blue: Fish population $x$
Assumption A0: the global payoff function is invariant by permuting the variables, i.e.,

$$r_g(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = r_g(a_{\pi(1)}, \ldots, a_{\pi(i)}, \ldots, a_{\pi(j)}, \ldots, a_{\pi(n)})$$

for every permutation $\pi: \mathcal{N} = \{1, 2, \ldots, n\} \rightarrow \mathcal{N}$, and $r_g$ is (locally) twice differentiable with respect to the variables.
\( \partial_{a_j} r_g(\bar{m} \otimes^n) = \partial_{a_1} r_g(\bar{m} \otimes^n) \) where \( \bar{m} \otimes^n := (\bar{m}, \ldots, \bar{m}) \) and

\[
\bar{m} = \frac{1}{n} \sum_{j=1}^{n} a_j = \int_{b \in A} b \left[ \frac{1}{n} \sum_{j=1}^{n} \delta_{a_j} \right] (db),
\]

(8)

\( \delta_{a_j} \) is the Dirac measure concentrated at the point \( a_j \),
\( m = \frac{1}{n} \sum_{j=1}^{n} \delta_{a_j} \).

The structure of the payoff function implies that the first order term in the Taylor expansion is cancelled out.

The cross-derivatives are independent of the labels:
\( \partial_{a_i a_j}^2 r_g(\bar{m} \otimes^n) = \partial_{a_1 a_2}^2 r_g(\bar{m} \otimes^n) \)
Example: First price auction

- Player 0, 1, \ldots, n.
- Player 0 proposes an object, a slot, a resource, depending on the state of environment etc.
- Each player has its own-valuation: \( v_j \sim F_j \)
- Each of Players 1, 2, \ldots, n chooses a bidding strategy: \( v_j \mapsto b_j(v_j) \).
- Player 0 gets \( \max_j b_j(v_j) \).
- The player with the highest price is the winner and saves \( (v_{j^*} - b_{j^*})1_{j^*} \) is winner.
THANK YOU