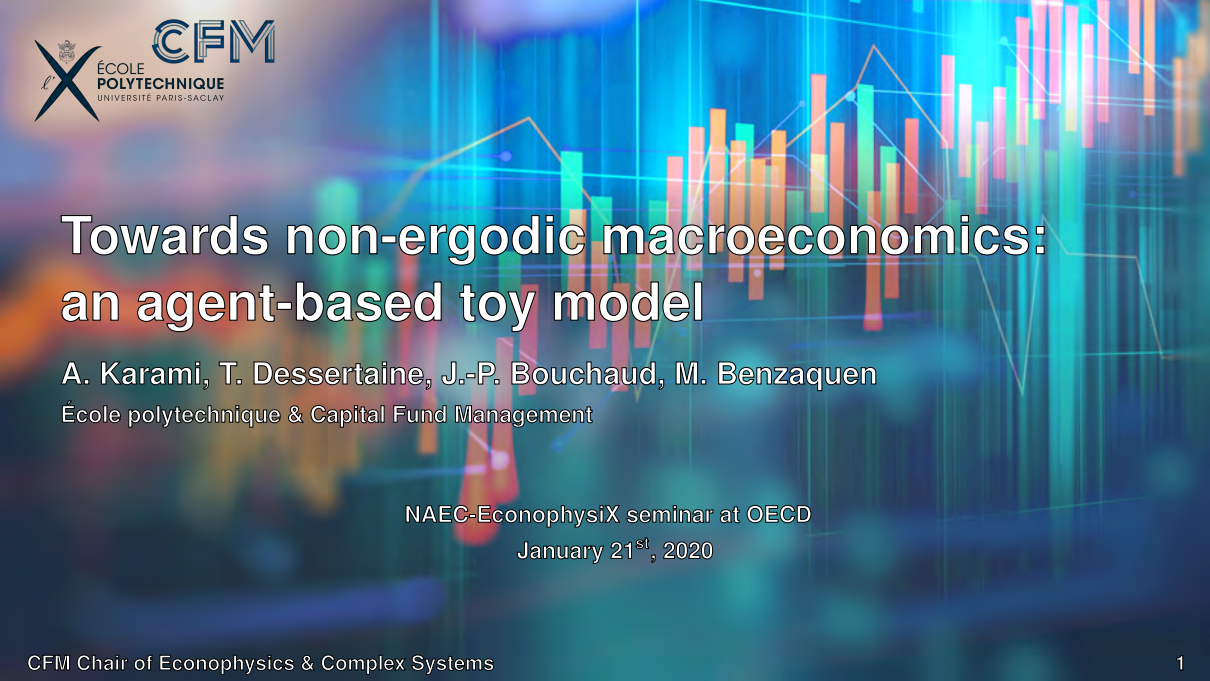


# Towards non-ergodic macroeconomics: an agent-based toy model

A. Karami, T. Dessertaine, J.-P. Bouchaud, M. Benzaquen

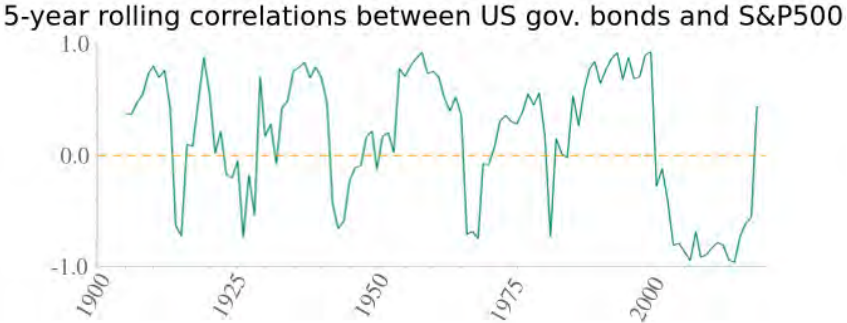
École polytechnique & Capital Fund Management

NAEC-EconophysiX seminar at OECD

January 21st, 2020

# Unique equilibrum in classical economics

- Mainstream economics: unique equilibrium around which the system oscillates



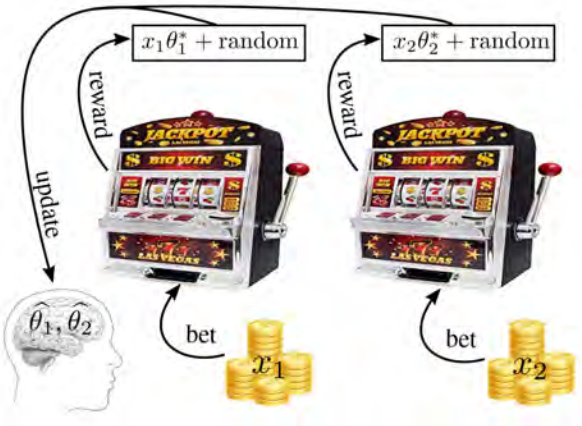

- A key assumption behind classical models (expected utility) is ergodicity
  - each agent makes decisions based on the average outcome over all possible ones
  - fundamental values take a preponderant place, no room left for self-fulfilling behavior

# Outline of this work

We present an example of a **non-ergodic agent-based** toy model

- the model is derived from a multi-agent linear bandits model
- we study agent interactions and their impact on ergodicity & synchronization
- mainly simulation results, some of which are backed by analytical arguments

# An intuitive presentation of the model




- The agent places successive bets $(x_{1,1}, x_{2,1}, \ldots, x_{1,t}, x_{2,t})$
  - doing so, he learns estimates $\widehat{\theta_{1,t}}$ and $\widehat{\theta_{2,t}}$ of $\theta_1^*$ and $\theta_2^*$ ("fundamental values")

# Linear bandits framework: agent policy

- With $d$ machines, $\theta^* = (\theta_1^*, \dots, \theta_d^*)$, $\widehat{\theta}_t = (\widehat{\theta_{1,t}}, \dots, \widehat{\theta_{d,t}})$, $x_t = (x_{1,t}, \dots, x_{1,d})$
- Agent-wise policy: at each time $t$, each agent bets as $x_t = x(\widehat{\theta}_t) = \widehat{\theta}_t / \|\widehat{\theta}_t\|$
- $\widehat{\theta_{t+1}}$ for next turn chosen as $\widehat{\theta}_{t+1} = \underset{\theta \in \mathbb{R}^d}{\arg\min} \left\{ \sum_{s=1}^{t} \left( x_s^\intercal \theta^* + \xi_s - x_s^\intercal \theta \right)^2 + \lambda \|\theta\| \right\}$

**Self-coherent set**

- $\{\theta \mid x(\theta)^\intercal \theta = x(\theta)^\intercal \theta^*\}$

avg. predicted reward = avg. observed reward $\implies$ set of possible limits of $\widehat{\theta}_t$

- Once the reward predicted with the agent's estimator is on average equal to the reward he gets, the agent effectively stops updating his estimator
  - even though $\widehat{\theta}_t \neq \theta^*$!

# Linear bandits framework: agent policy

- With $d$ machines, $\theta^* = (\theta_1^*, \dots, \theta_d^*)$, $\widehat{\theta}_t = (\widehat{\theta_{1,t}}, \dots, \widehat{\theta_{d,t}})$, $x_t = (x_{1,t}, \dots, x_{1,d})$
- Agent-wise policy: at each time $t$, each agent bets as $x_t = x(\widehat{\theta}_t) = \widehat{\theta}_t / \|\widehat{\theta}_t\|$
- $\widehat{\theta_{t+1}}$ for next turn chosen as $\widehat{\theta}_{t+1} = \underset{\theta \in \mathbb{R}^d}{\arg\min} \left\{ \sum_{s=1}^{t} \left( x_s^\intercal \theta^* + \xi_s - x_s^\intercal \theta \right)^2 + \lambda \|\theta\| \right\}$

**Self-coherent set**

- $\{\theta \mid x(\theta)^\intercal \theta = x(\theta)^\intercal \theta^*\}$

avg. predicted reward = avg. observed reward $\implies$ set of possible limits of $\widehat{\theta}_t$

- Once the reward predicted with the agent's estimator is on average equal to the reward he gets, the agent effectively stops updating his estimator
  - even though $\widehat{\theta}_t \neq \theta^*$!

# Simulation in $d = 2$ dimensions

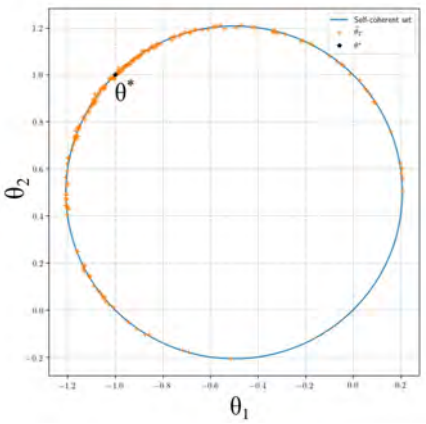


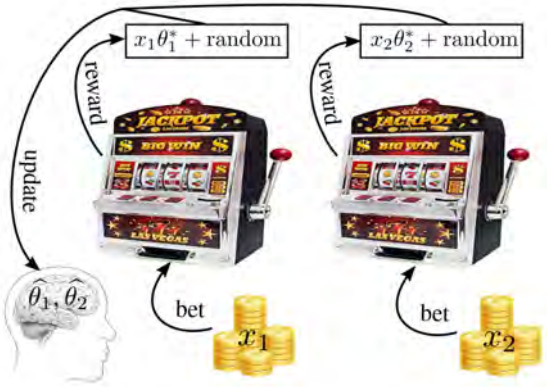

Figure: Simulation with $d = 2$, for $N = 200$ players. Parameters: $N = 200, T = 1000$, $\lambda = 0.1$, $\sigma = 0.1$. Each cross is the final position of each independant agent's estimator.

# Simulation in $d = 2$ dimensions

Simulation for $N = 200$ independent agents.

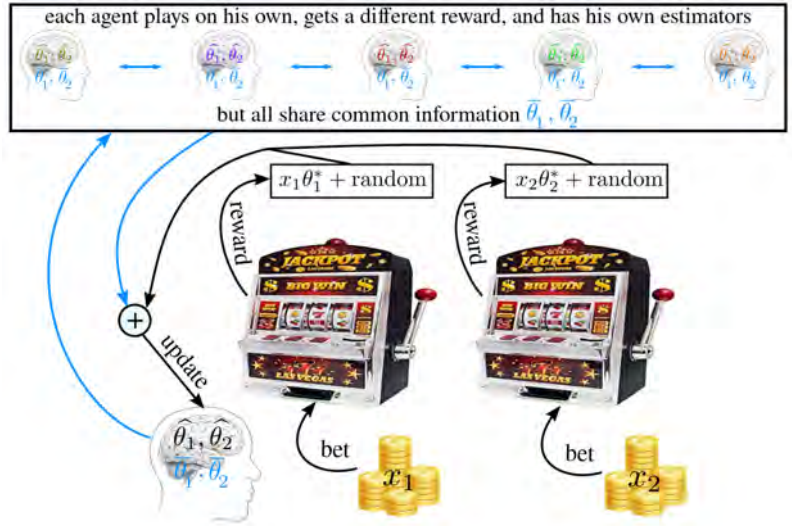
**Let us now add some interaction between the agents, in the form of cooperation.**

# Adding inter-agent interaction: cooperating agents



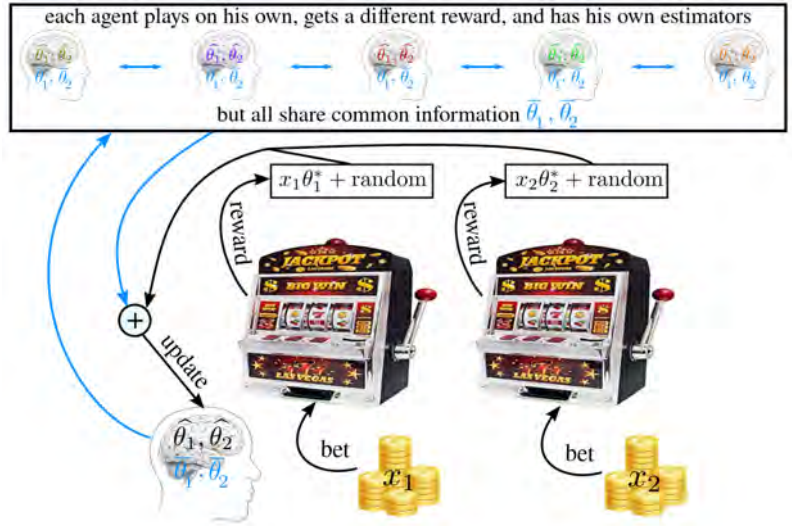

Each agent updates his $\widehat{\theta_{t+1}}$ as $(1-\alpha)$(own policy) $+ \alpha$(mean estimator $\overline{\theta_t}$)

# Adding inter-agent interaction: cooperating agents




Each agent updates his $\widehat{\theta_{t+1}}$ as $(1-\alpha)$(own policy) $+\,\alpha$(mean estimator $\overline{\theta_t}$)

# Adding inter-agent interaction: cooperating agents




**Each agent updates his $\widehat{\theta_{t+1}}$ as $(1-\alpha)$(own policy) $+ \alpha$(mean estimator $\overline{\theta_t}$)**

# Cooperating agents: simulations

- There is **synchronization**: the agents estimators tend to the same limit
- The limit is **non-ergodic**: the limiting estimator is non-deterministic

# Cooperation: non-ergodicity

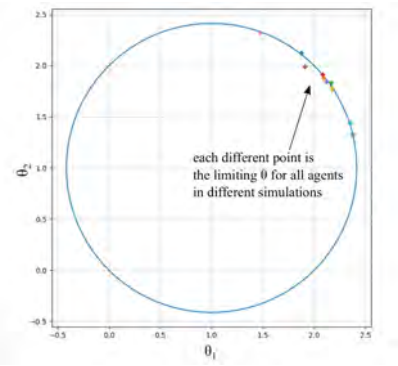
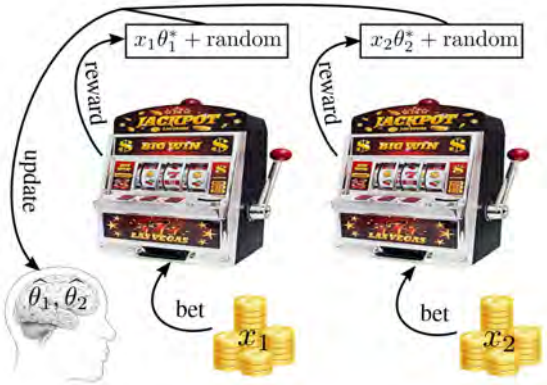



Figure: Cooperation interactions: examples of synchronized but non-ergodic behavior (each cross represent all the agents for different simulations, $\alpha = 0.8, N = 10, T = 500, \lambda = 0.1$).

- There is **synchronization**:
  the agents estimators tend to the same limit
- The limit is **non-ergodic**:
  the limiting estimator is non-deterministic

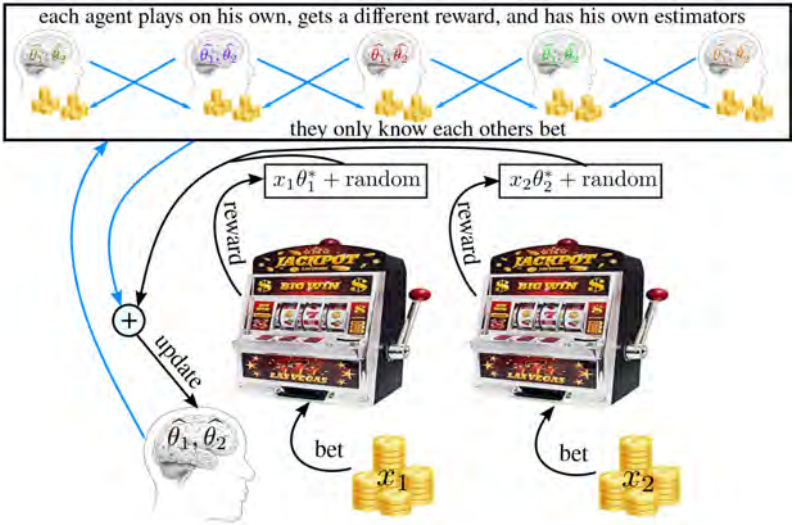
**As a different inter-agent interaction, we now consider competition.**

# Adding inter-agent interaction: competing agents



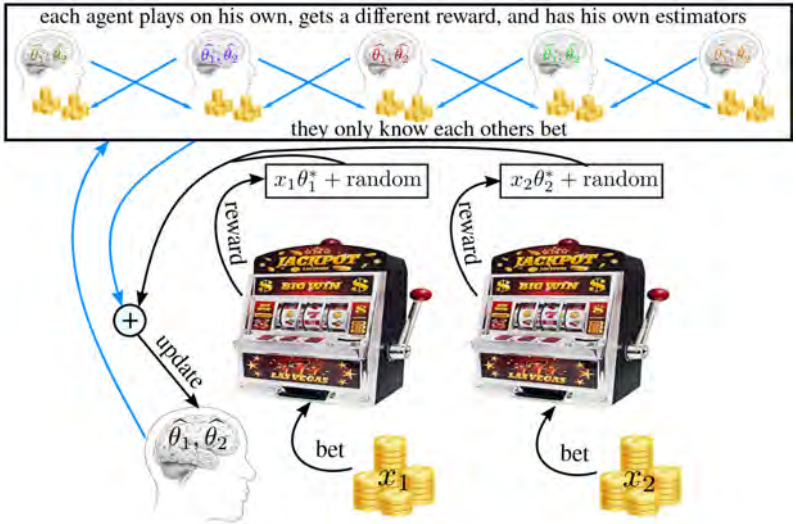

Each agent sets his bets $x_{t+1}$ as $(1-\alpha)$(own betting policy) $+ \alpha$(average bet $\overline{x_t}$)

# Adding inter-agent interaction: competing agents




Each agent sets his bets $x_{t+1}$ as $(1-\alpha)$(own betting policy) $+\,\alpha$(average bet $\overline{x}_t$)

# Adding inter-agent interaction: competing agents

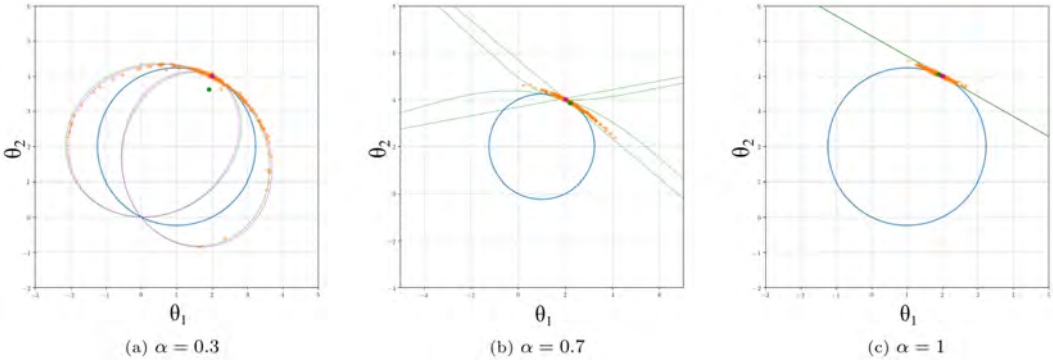



**Each agent sets his bets $x_{t+1}$ as $(1-\alpha)$(own betting policy)$+\alpha$(average bet $\overline{x_t}$)**

# Competing agents: simulations

# Competing agents: simulation results for different values of $\alpha$

- The self-coherent set has a more complex geometry and strongly depends on $\alpha$




(a) $\alpha = 0.3$ (b) $\alpha = 0.7$ (c) $\alpha = 1$

- **No ergodicity nor synchronization** (at least for reasonable $T$)

# Summary and further developments

- We have presented a toy agent-based model exhibiting non-ergodic behavior
  - although interactions result in collective phenomena, ergodicity is not enforced
- Current research: analytical arguments on (non-)ergodicity and synchronization
  - if ergodic in theory, how quick is convergence ? (“effective ergodicity“, $\sigma / T$ interplay)
- Longer term: complexifying the model and mapping it to economical situations

**Thank you.**

## Supplementary

# Linear bandits framework: setting

- $\chi \subset \mathbb{R}^n$ is set of allowable allocations, e.g., $\chi = \mathcal{B}(0, 1)$ meaning that $\sum_i |x_i| = 1$
  - think ways to bet on $n$ casino machines,
- $\theta^\star \in \mathbb{R}^n$ the unknown parameter
  - average reward factor of each machine
- $\xi$ a zero-mean, Gaussian noise of standard deviation $\sigma$
  - each machine's randomness in generated return

**Reward**

The reward generated from choosing an allocation $x \in \chi$ is randomly drawn as

$$r(x) = x^\intercal \theta^\star + \sigma\xi$$

# Linear bandits framework: setting

- $\chi \subset \mathbb{R}^n$ is set of allowable allocations, e.g., $\chi = \mathcal{B}(0, 1)$ meaning that $\sum_i |x_i| = 1$
  - think ways to bet on $n$ casino machines,
- $\theta^\star \in \mathbb{R}^n$ the unknown parameter
  - average reward factor of each machine
- $\xi$ a zero-mean, Gaussian noise of standard deviation $\sigma$
  - each machine's randomness in generated return

**Reward**

The reward generated from choosing an allocation $x \in \chi$ is randomly drawn as

$$r(x) = x^\intercal \theta^\star + \sigma\xi$$

## Estimator update algorithm

Given initial conditions and maximum time $(\widehat{\theta}_1, T)$:

1. **for** $t \in \{1, \ldots, T\}$ **do**
2. compute the optimal allocation $x_t = \widehat{\theta}_t / \|\widehat{\theta}_t\|$
3. play $x_t$ and observe reward $r(x_t)$,
4. update $\widehat{\theta}_{t+1}$ by least square criterion
5. **end for**

# Linear bandits framework: agent policy

- An allocation $x \in \chi$ is evaluated according to its expected reward $\mathbb{E}[r(x)] = x^\intercal \theta^\star$
- For any $\theta \in \mathbb{R}^d$, the optimal allocation is denoted by $x^\star(\theta) = \underset{x \in \chi}{\text{argmax}}\ x^\intercal \theta$
- If the admissible allocations is $\chi = \mathcal{B}(0, 1)$ then $x^\star(\theta) = \theta / \|\theta\|_1$.
  - we choose this example for the following simulations.

- Pairs (allocation, reward) come over time: $\{(x_1, r_1), (x_2, r_2), \ldots, (x_{t-1}, r_{t-1})\}$
  - at each timestep, the agent's estimator $\widehat{\theta}_t$ is updated:

$$\widehat{\theta}_t = \underset{\theta \in \mathbb{R}^d}{\arg\min} \sum_{s=1}^{t-1} \left(r_s - x_s^T \theta\right)^2 + \lambda \|\theta\|^2 \qquad \text{(ridge least squares)} \tag{1}$$

- In closed form,

$$\widehat{\theta}_t = V_t^{-1} \sum_{s=1}^{t-1} x_s r_s \qquad \text{where} \qquad V_t = \lambda I + \sum_{s=1}^{t-1} x_s x_s^T \tag{2}$$

Developing the closed form formula for the updated estimator $\widehat{\theta}_{t+1}$, we get:

$$\widehat{\theta}_{t+1} = \widehat{\theta}_t + \frac{V_t^{-1} x_t}{g_t + 1}\Big[ \quad x_t^\intercal \theta^\star - x_t^\intercal \widehat{\theta}_t \quad + \sigma \xi_t \Big] \qquad \text{with } g_t = x_t^\intercal V_t^{-1} x_t \tag{3}$$

**Self-coherent set**

- $\{\theta \mid x^*(\theta)^\intercal \theta = x^*(\theta)^\intercal \theta^*\}$

avg. predicted reward = avg. observed reward $\implies$ set of possible limits of $\widehat{\theta}_t$

- We have chosen admissible allocations as $\chi = B(0, 1)$
  - in this case, the self-coherent set is a circle of center $\theta^*/2$ and radius $\|\theta^*\|/2$.

# Self-Coherent Set

Developing the closed form formula for the updated estimator $\widehat{\theta}_{t+1}$, we get:

$$\widehat{\theta}_{t+1} = \widehat{\theta}_t + \frac{V_t^{-1} x_t}{g_t + 1}\Big[ \underbrace{x_t^\intercal \theta^\star - x_t^\intercal \widehat{\theta}_t}_{\substack{\textbf{reward} \text{ with optimal estimator} \\ \text{vs.} \\ \textbf{reward} \text{ with current estimator}}} + \sigma \xi_t \Big] \qquad \text{with } g_t = x_t^\intercal V_t^{-1} x_t \quad (3)$$

**Self-coherent set**

- $\{\theta \mid x^*(\theta)^\intercal \theta = x^*(\theta)^\intercal \theta^\star\}$

avg. predicted reward = avg. observed reward $\implies$ set of possible limits of $\widehat{\theta}_t$

- We have chosen admissible allocations as $\chi = B(0, 1)$
  - in this case, the self-coherent set is a circle of center $\theta^*/2$ and radius $\|\theta^*\|/2$.

Developing the closed form formula for the updated estimator $\widehat{\theta}_{t+1}$, we get:

$$\widehat{\theta}_{t+1} = \widehat{\theta}_t + \frac{V_t^{-1} x_t}{g_t + 1}\Big[ \underbrace{x_t^\intercal \theta^\star - x_t^\intercal \widehat{\theta}_t}_{0 \text{ on the self-coherent set}} + \sigma \xi_t \Big] \quad \text{with } g_t = x_t^\intercal V_t^{-1} x_t \qquad (3)$$

**Self-coherent set**

$$\blacktriangleright \ \{\theta \mid x^*(\theta)^\intercal \theta = x^*(\theta)^\intercal \theta^\star\}$$

avg. predicted reward = avg. observed reward $\implies$ set of possible limits of $\widehat{\theta}_t$

- We have chosen admissible allocations as $\chi = B(0, 1)$
  - in this case, the self-coherent set is a circle of center $\theta^*/2$ and radius $\|\theta^*\|/2$.

Developing the closed form formula for the updated estimator $\widehat{\theta}_{t+1}$, we get:

$$\widehat{\theta}_{t+1} = \widehat{\theta}_t + \frac{V_t^{-1} x_t}{g_t + 1}\Big[ \underbrace{x_t^\intercal \theta^\star - x_t^\intercal \widehat{\theta}_t}_{0 \text{ on the self-coherent set}} + \sigma \xi_t \Big] \quad \text{with } g_t = x_t^\intercal V_t^{-1} x_t \qquad (3)$$

**Self-coherent set**

- $\{\theta \mid x^*(\theta)^\intercal \theta = x^*(\theta)^\intercal \theta^\star\}$

avg. predicted reward = avg. observed reward $\implies$ set of possible limits of $\widehat{\theta}_t$

- We have chosen admissible allocations as $\chi = B(0, 1)$
  - in this case, the self-coherent set is a circle of center $\theta^*/2$ and radius $\|\theta^*\|/2$.

# Adding inter-agent interaction: cooperating agents

- **Cooperation**: agents share knowledge: each agent $j$ shares his estimator $\widehat{\theta}_t^j$.
  - an agent $j$ chooses the estimator

$$\widehat{\theta}_{t+1} = (1-\alpha)\left(\widehat{\theta}_t + \frac{V_t^{-1}x_t}{g_t+1}\left[x_t^\intercal\theta^\star - x_t^\intercal\widehat{\theta}_t + \sigma\xi_t\right]\right) + \alpha\overline{\theta_t}$$

- **Competition**: they only have access to the action made by the other agents, $x_t$.
  - the agent $j$ chooses the allocation:

$$x_t^j = (1-\alpha)x^\star(\widehat{\theta}_t^j) + \alpha\overline{x_t} \tag{4}$$

with $\overline{\theta}_t, \overline{x}_t$ the mean across the $N$ agents at time $t$ and $\alpha \in [0,1]$.

- **Competition**: agents only have access other agents' *allocation* $(x_t^j)_{1\leq j\leq N}$.
  - the allocation chosen by player $j$ is

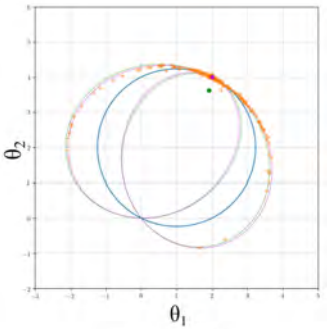
$$x_t^j = (1-\alpha)x^\star(\widehat{\theta}_t^j) + \alpha\overline{x_t}$$

  - from which we can deduce the estimator update map

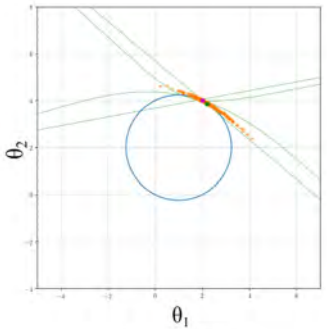
$$\widehat{\theta}_{t+1} = \widehat{\theta}_t + \frac{V_t^{-1}x_t}{g_t+1}\left[(1-\alpha)\left(x_t^\intercal\theta^\star - x_t^\intercal\widehat{\theta}_t\right) + \alpha\left(\overline{x_t}^\intercal\theta^\star - \overline{x_t}^\intercal\widehat{\theta}_t\right)\right]$$

# Competing agents: different self-coherent sets

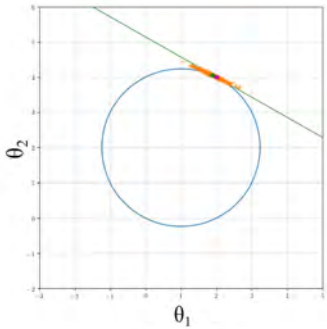
- The self-coherent set has a more complex geometry and strongly depends on $\alpha$



(a) $\alpha = 0.3$

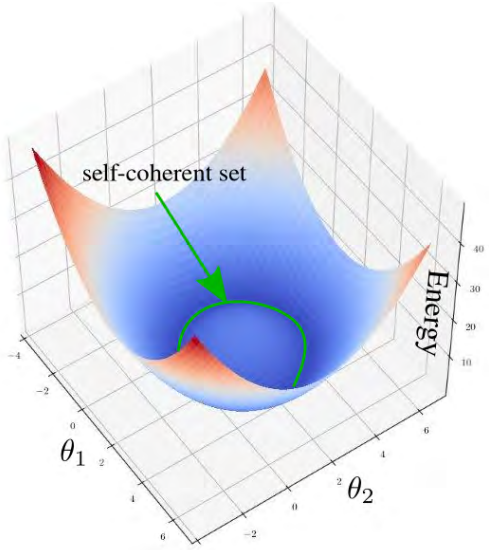


(b) $\alpha = 0.7$



(c) $\alpha = 1$

# Energy-minimization interpretation of agents behavior



# Distribution on the self-coherent set: independent agents

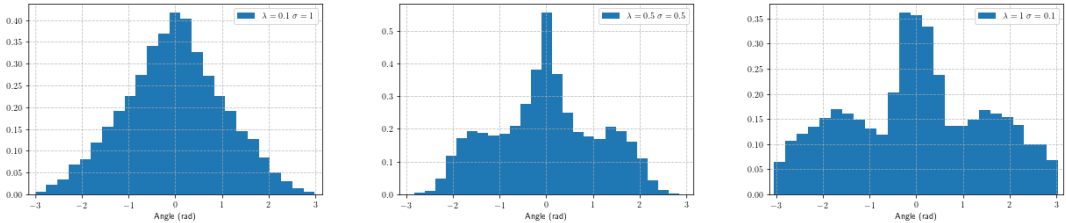



Figure: Histograms limit estimators angles. Parameters $T = 1000$, $N = 1000$.

# Cooperation: Influence of $\alpha$

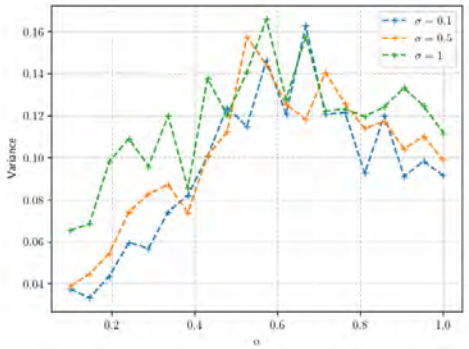



Figure: Variance of the limit distribution as a function of the interaction parameter $\alpha$. Parameters: $N = 10, T = 500$, $\lambda = 0.1$

# Cooperation: Influence of $N$

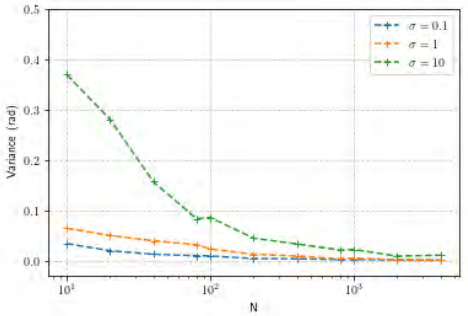



Figure: Variance of the limit distribution as a function of the number of player $N$. Parameters: $\alpha = 0.3, T = 500$, $\lambda = 0.1$