Mean Field Stochastic Control

Peter E. Caines

IEEE Control Systems Society Bode Lecture

48^{th} Conference on Decision and Control and 28^{th} Chinese Control Conference

Shanghai 2009

McGill University

Co-Authors

Minyi Huang

Roland Malhamé

Collaborators & Students

Arman Kizilkale

Arthur Lazarte

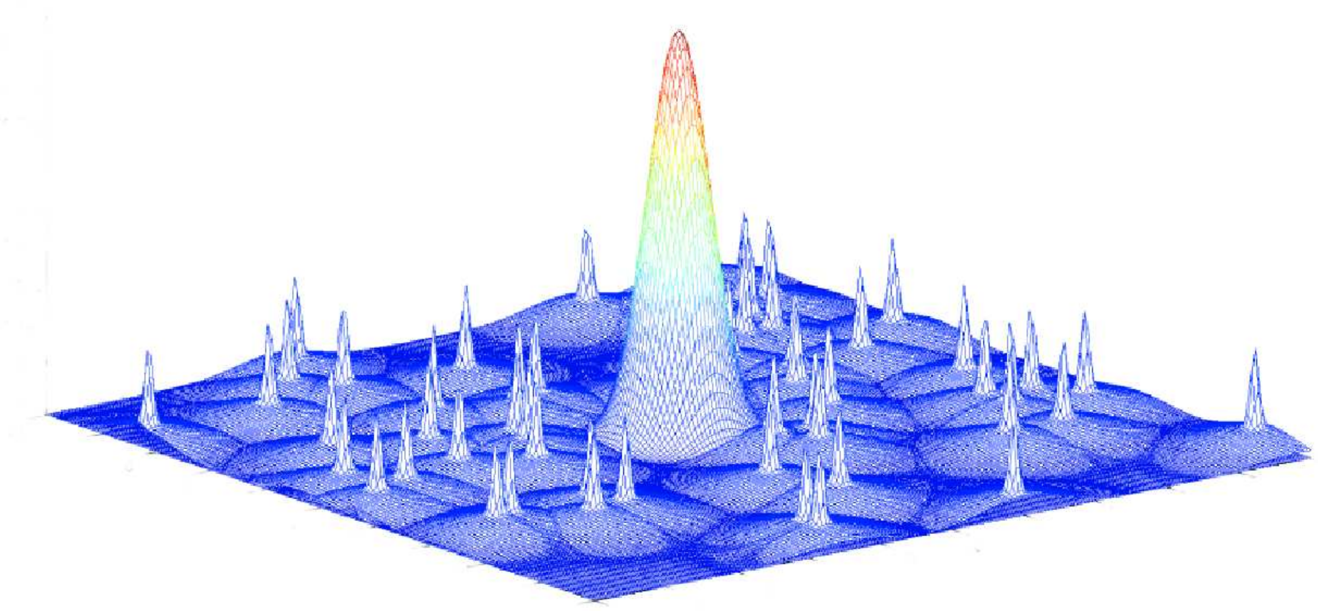
Zhongjing Ma

Mojtaba Nourian

Overview

- Overall Objective:
  - Present a theory of decentralized decision-making in stochastic dynamical systems with many competing or cooperating agents
- Outline:
  - Present a motivating control problem from Code Division Multiple Access (CDMA) uplink power control
  - Motivational notions from Statistical Mechanics
  - The basic notions of Mean Field (MF) control and game theory
  - The Nash Certainty Equivalence (NCE) methodology
  - Main NCE results for Linear-Quadratic-Gaussian (LQG) systems
  - NCE systems with interaction locality: Geometry enters the models
  - Present adaptive NCE system theory

Base Station & Individual Agents



Part I – CDMA Power Control

- Lognormal channel attenuation: $1 \leq i \leq N$

  i_{th} channel: $dx_i = -a(x_i + b)dt + \sigma dw_i, \qquad 1 \leq i \leq N$

  Transmitted power = channel attenuation $\times$ power

  = $e^{x_i(t)} p_i(t)$

  (Charalambous, Menemenlis; 1999)

  Signal to interference ratio (Agent i) at the base station

  = $e^{x_i} p_i / \left[(\beta/N) \sum_{j \neq i}^{N} e^{x_j} p_j + \eta \right]$

- How to optimize all the individual SIR's?
  - Self defeating for everyone to increase their power
  - Humans display the "Cocktail Party Effect": Tune hearing to frequency of friend's voice

Part I – CDMA Power Control

- Can maximize $\sum_{i=1}^{N} SIR_i$ with centralized control.
  (HCM, 2004)

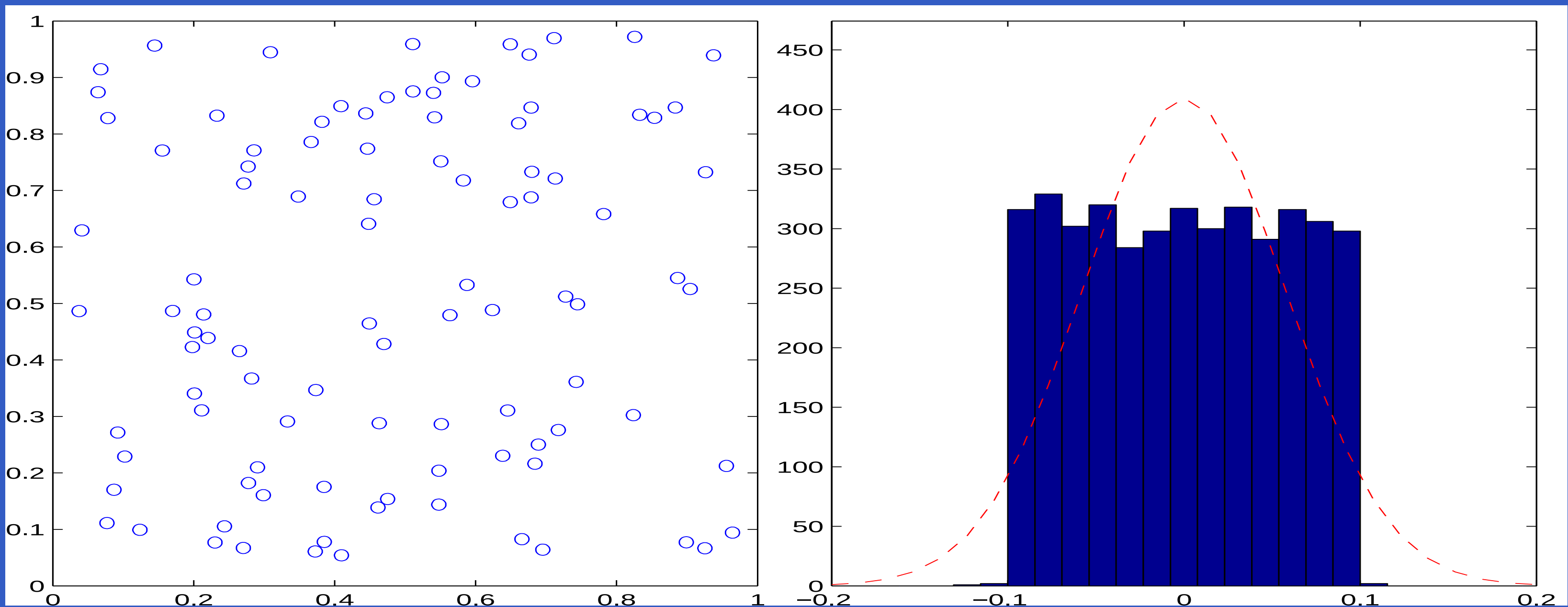
- Since centralized control is not feasible for complex systems, how can such systems be optimized using decentralized control?

- Idea: Use large population properties of the system together with basic notions of game theory.

The Statistical Mechanics Connection

Part II – Statistical Mechanics

- A foundation for thermodynamics was provided by the Statistical Mechanics of Boltzmann, Maxwell and Gibbs.
- Basic Ideal Gas Model describes the interaction of a huge number of essentially identical particles.
- SM explains very complex individual behaviours by PDEs for a continuum limit of the mass of particles.



Animation of Particles

Start from the equations for the perfectly elastic (i.e. hard) sphere mechanical collisions of each pair of particles:

Velocities before collision: $\mathbf{v}, \mathbf{V}$

Velocities after collision: $\mathbf{v}' = \mathbf{v}'(\mathbf{v}, \mathbf{V}, t), \quad \mathbf{V}' = \mathbf{V}'(\mathbf{v}, \mathbf{V}, t)$

These collisions satisfy the conservation laws, and hence:

$$\text{Conservation of} \begin{cases} \text{Momentum} & m(\mathbf{v}' + \mathbf{V}') = m(\mathbf{v} + \mathbf{V}) \\[2em] \text{Energy} & \frac{1}{2}m\left(\|\mathbf{v}'\|^2 + \|\mathbf{V}'\|^2\right) = \frac{1}{2}m\left(\|\mathbf{v}\|^2 + \|\mathbf{V}\|^2\right) \end{cases}$$

Part II – Boltzmann's Equation

The assumption of Statistical Independence of particles (Propagation of Chaos Hypothesis) gives Boltzmann's PDE for the behaviour of an infinite population of particles

$$\frac{\partial p_t(\mathbf{v},\mathbf{x})}{\partial t} + \nabla_x p_t(\mathbf{v},\mathbf{x}) \cdot \mathbf{v} = \iiint Q(\theta,\psi) d\theta d\psi \|\mathbf{v} - \mathbf{V}\| \cdot$$
$$\Big(p_t(\mathbf{v}',\mathbf{x}) p_t(\mathbf{V}',\mathbf{x}) - p_t(\mathbf{v},\mathbf{x}) p_t(\mathbf{V},\mathbf{x}) \Big) d^3\mathbf{V}$$

$$\mathbf{v}' = \mathbf{v}'(\mathbf{v},\mathbf{V},t), \quad \mathbf{V}' = \mathbf{V}'(\mathbf{v},\mathbf{V},t)$$

Entropy H:

$$H(t) \stackrel{\text{def}}{=} -\iiint p_t(\mathbf{v},x)(\log p_t(\mathbf{v},x))\, d^3x d^3\mathbf{v}$$

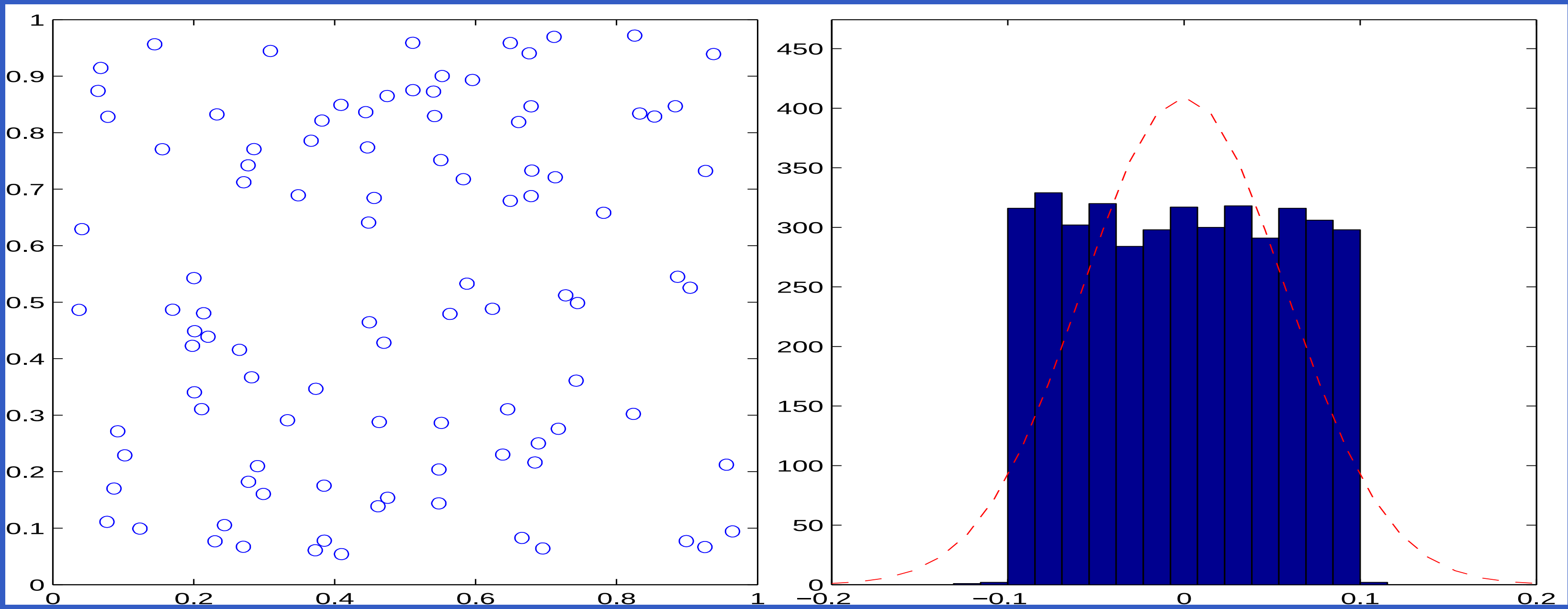
The H-Theorem:
$$\frac{dH(t)}{dt} \geq 0$$

$$H_\infty \stackrel{\text{def}}{=} \sup_{t\geq 0} H(t)$$ occurs at $$p_\infty = N(\mathbf{v}_\infty, \Sigma_\infty)$$

(The Maxwell Distribution)

Part II – Statistical Mechanics: Key Intuition

Extremely complex large population particle systems
may have simple continuum limits
with distinctive bulk properties.

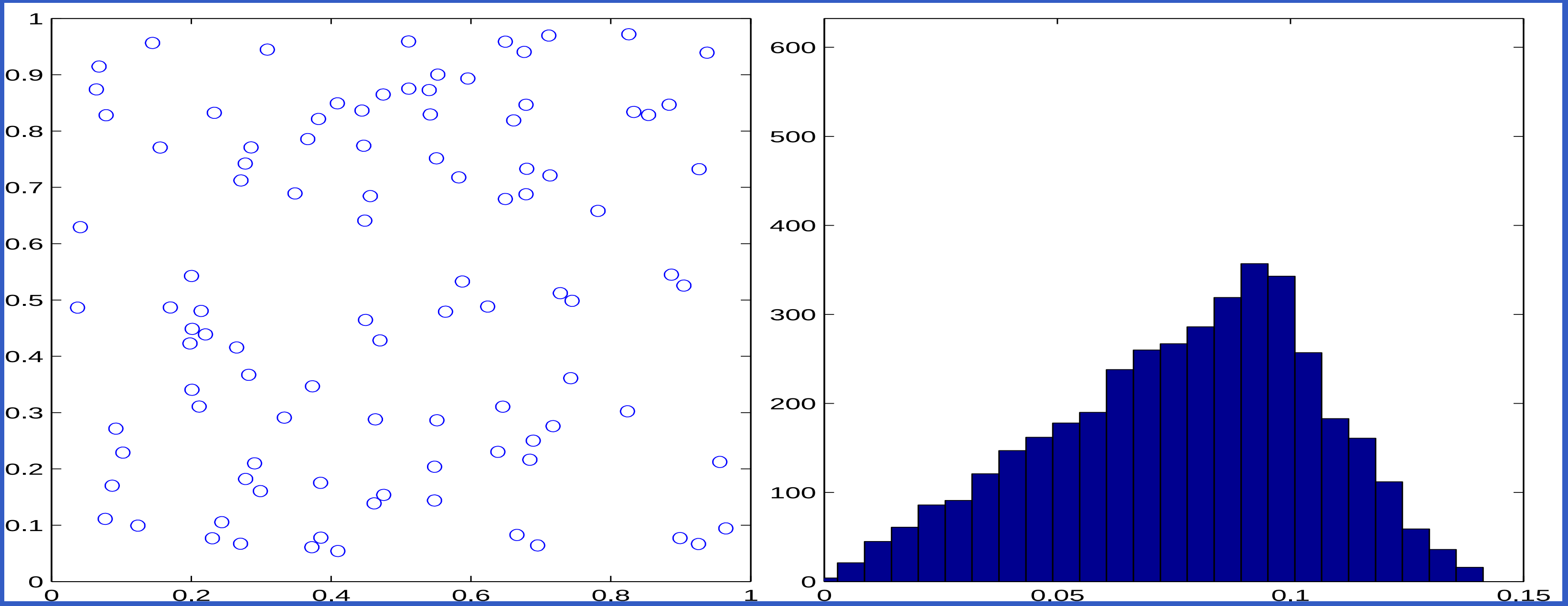


Animation of Particles

Control of Natural Entropy Increase

Feedback Control Law (Non-physical Interactions):

- At each collision, total energy of each pair of particles is shared equally while physical trajectories are retained.
- Energy is conserved.



Animation of Particles

Part II – Key Intuition

- A sufficiently large mass of individuals may be treated as a continuum.
- Local control of particle (or agent) behaviour can result in (partial) control of the continuum.

- Game Theoretic Control Systems of interest will have many competing agents
  - A large ensemble of players seeking their individual interest
  - Fundamental issue: The relation between the actions of each individual agent and the resulting mass behavior

Part III – Basic LQG Game Problem

Individual Agent's Dynamics:

$$dx_i = (a_i x_i + b_i u_i)dt + \sigma_i dw_i, \quad 1 \leq i \leq N.$$

(scalar case only for simplicity of notation)

- x_i: state of the ith agent
- u_i: control
- w_i: disturbance (standard Wiener process)
- N: population size

Part III – Basic LQG Game Problem

Individual Agent's Cost:

$$J_i(u_i, \nu) \triangleq E \int_0^\infty e^{-\rho t}[(x_i - \nu)^2 + r u_i^2] dt$$

Basic case: $\nu \triangleq \gamma.(\frac{1}{N} \sum_{k \neq i}^N x_k + \eta)$

More General case: $\nu \triangleq \Phi(\frac{1}{N} \sum_{k \neq i}^N x_k + \eta)$ Φ Lipschitz

Main feature:

- Agents are coupled via their costs
- Tracked process ν:

(i) stochastic

(ii) depends on other agents' control laws

(iii) not feasible for x_i to track all x_k trajectories for large N

Part III – Background & Current Related Work

Background:

- 40+ years of work on stochastic dynamic games and team problems: Witsenhausen, Varaiya, Ho, Basar, *et al.*

Current Related Work:

- Industry dynamics with many firms: Markov models and Oblivious Equilibria (Weintraub, Benkard, Van Roy, Adlakha, Johari & Goldsmith, 2005, 2008 -)

- Mean Field Games: Stochastic control of many agent systems with applications to finance (Lasry and Lions, 2006 -)

Part III – Large Popn. Models with Game Theory Features

- Economic models (e.g, oligopoly production output planning): Each agent receives average effect of others via the market. Cournot-Nash equilibria (Lambson)
- Advertising competition game models (Erickson)
- Wireless network resource allocation (Alpcan, Basar, Srikant, Altman; HCM)
- Admission control in communication networks; (point processes) (Ma, RM, PEC)
- Feedback control of large scale mechanics-like systems: Synchronization of coupled oscillators (Yin, Mehta, Meyn, Shanbhag)

Part III – Large Popn. Models with Game Theory Features

- Public health: Voluntary vaccination games (Bauch & Earn)
- Biology: Pattern formation in plankton (S. Levine *et al.*); stochastic swarming (Morale *et al.*); PDE swarming models (Bertozzi *et al.*)
- Biology: "Selfish Herd" behaviour: Fish reducing individual predation risk by joining the group (Reluga & Viscido)
- Sociology: Urban economics and urban behaviour (Brock and Durlauf *et al.*)

Part III – An Example: Collective Motion of Fish (Schooling)

Part III – An Example: School of Sardines

School of Sardines

Part III – Names

Why Mean Field Theory?

Standard Term in Physics for the Replacement of Complex Hamiltonian H by an Average $\bar{H}$

Notion Used in Many Fields for Approximating Mass Effects in Complex Systems

Shall use Mean Field as General Term and Nash Certainty Equivalence (NCE) in Specific Control Theory Contexts

Part III – Names

Why Nash Certainty Equivalence Control (NCE Control)?

Because the equilibria established are Nash Equilibria

and

The feedback control laws depend upon an "equivalence" assumed between unobserved system properties and agent-computed approximations e.g. θ^0 and $\hat{\theta}_N$ (which are exact in a limit).

Part III – Preliminary Optimal LQG Tracking

Take x^* (bounded continuous) for scalar model:

$$dx_i = a_i x_i dt + b u_i dt + \sigma_i dw_i$$

$$J_i(u_i, x^*) = E\int_0^\infty e^{-\rho t}[(x_i - x^*)^2 + r u_i^2]dt$$

Riccati Equation: $\rho\Pi_i = 2a_i\Pi_i - \frac{b^2}{r}\Pi_i^2 + 1, \quad \Pi_i > 0$

Set $\beta_1 = -a_i + \frac{b^2}{r}\Pi_i$, $\beta_2 = -a_i + \frac{b^2}{r}\Pi_i + \rho$, and assume $\beta_1 > 0$

Mass Offset Control: $\rho s_i = \frac{ds_i}{dt} + a_i s_i - \frac{b^2}{r}\Pi_i s_i - x^*.$

Optimal Tracking Control: $u_i = -\frac{b}{r}(\Pi_i x_i + s_i)$

- Boundedness condition on x^* implies existence of unique solution s_i.

Part III – Key Intuition

When the tracked signal is replaced by the deterministic mean state of the mass of agents:

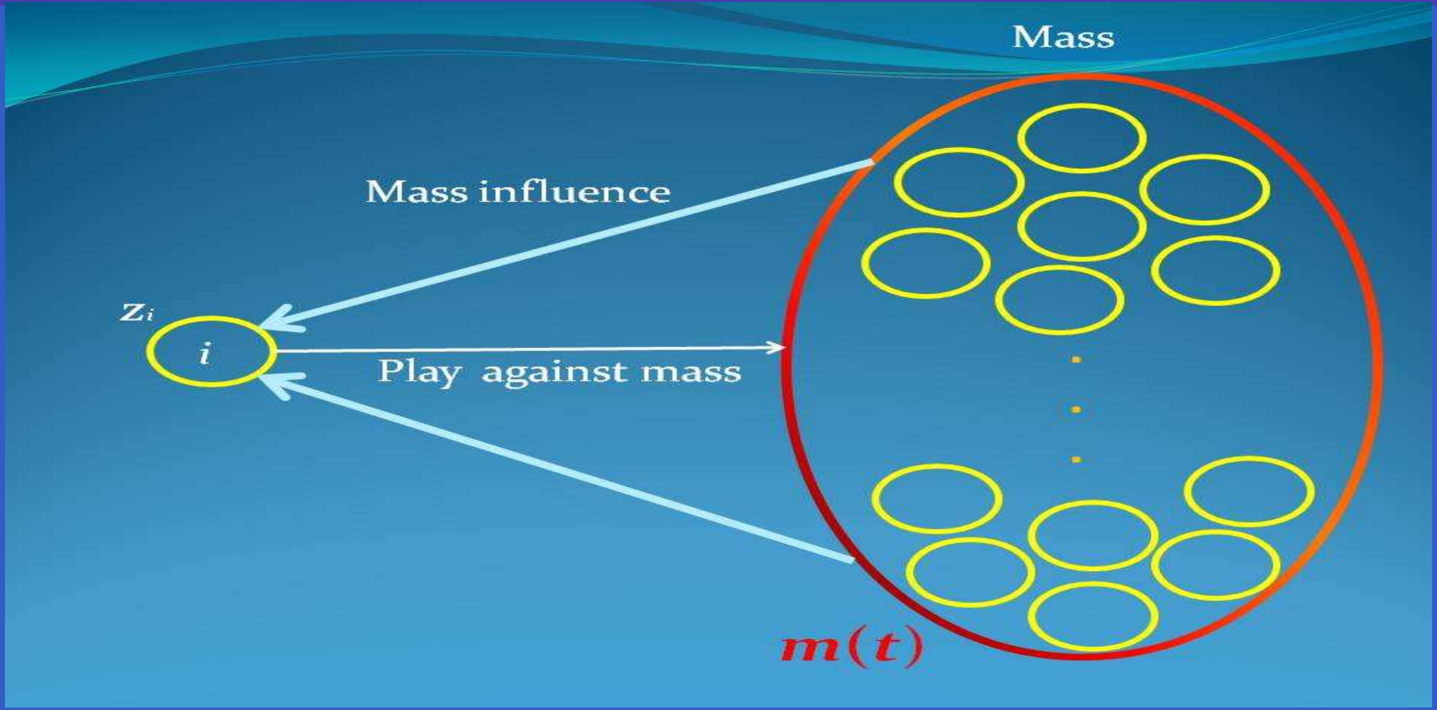
Agent's feedback = feedback of agent's local stochastic state

\+

feedback of deterministic mass offset

Think Globally, Act Locally
(Geddes, Alinsky, Rudie-Wonham)

Part III – Key Features of NCE Dynamics



- Under certain conditions, the mass effect concentrates into a deterministic - hence predictable - quantity $m(t)$.
- A given agent only reacts to its own state and the mass behaviour $m(t)$, any other individual agent becomes invisible.
- The individual behaviours collectively reproduce that mass behaviour.

Part III – Population Parameter Distribution

- Info. on other agents is available to each agent only via $F(.)$

Define empirical distribution associated with N agents

$$F_N(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(a_i \le x), \qquad x \in \mathbb{R}.$$

- H1 There exists a distribution function F on $\mathbb{R}$ such that $F_N \to F$ weakly as $N \to \infty$.

The Fundamental NCE Equation System

Continuum of Systems: $a \in \mathcal{A}$; common b for simplicity

$$\rho s_a = \frac{ds_a}{dt} + a s_a - \frac{b^2}{r}\Pi_a s_a - x^*$$

$$\frac{d\overline{x}_a}{dt} = (a - \frac{b^2}{r}\Pi_a)\overline{x}_a - \frac{b^2}{r}s_a,$$

$$\overline{x}(t) = \int_{\mathcal{A}} \overline{x}_a(t)\,dF(a),$$

$$x^*(t) = \gamma(\overline{x}(t) + \eta) \qquad t \geq 0$$

$$\textit{Riccati Equation}: \quad \rho\Pi_a = 2a\Pi_a - \frac{b^2}{r}\Pi_a^2 + 1, \quad \Pi_a > 0$$

- Individual control action $u_a = -\frac{b}{r}(\Pi_a x_a + s_a)$ is optimal w.r.t tracked x^*.
- Does there exist a solution $(\overline{x}_a, s_a, x^*; a \in \mathcal{A})$?

  Yes: Fixed Point Theorem

Part III – NCE Feedback Control

Proposed MF Solution to the Large Population LQG Game Problem

The Finite System of N Agents with Dynamics:

$$dx_i = a_i x_i dt + b u_i dt + \sigma_i dw_i, \qquad 1 \le i \le N, \qquad t \ge 0$$

Let $u_{-i} \triangleq (u_1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_N)$; then the individual cost

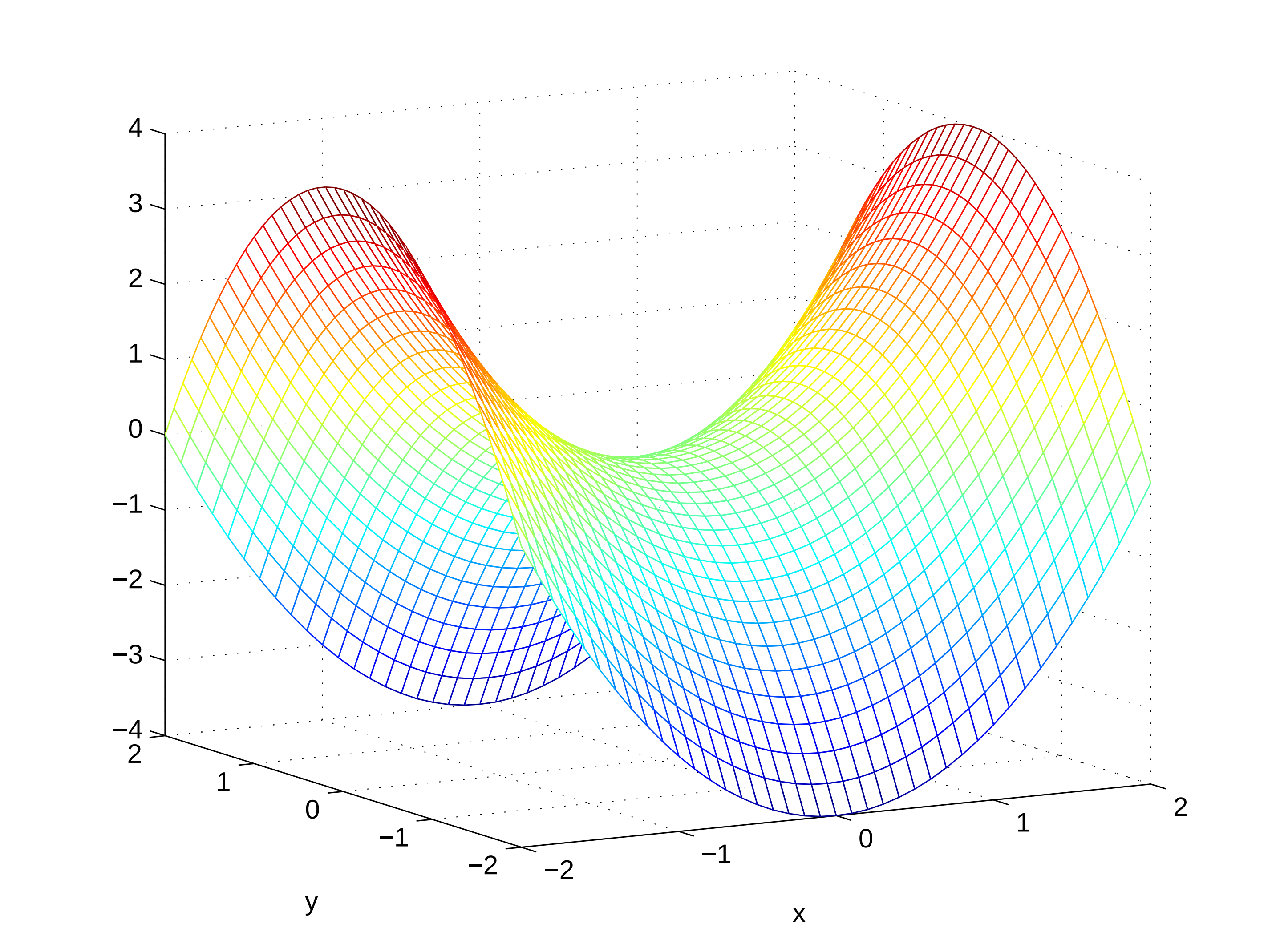
$$J_i(u_i, u_{-i}) \triangleq E \int_0^\infty e^{-\rho t}\{[x_i - \gamma(\frac{1}{N}\sum_{k \ne i}^{N} x_k + \eta)]^2 + r u_i^2\} dt$$

Algorithm: For ith agent with parameter (a_i, b) compute:

- x^* using NCE Equation System
- $$\begin{cases} \rho \Pi_i = 2a_i \Pi_i - \frac{b^2}{r}\Pi_i^2 + 1 \\ \rho s_i = \frac{ds_i}{dt} + a_i s_i - \frac{b^2}{r}\Pi_i s_i - x^* \\ u_i = -\frac{b}{r}(\Pi_i x_i + s_i) \end{cases}$$

Part III – Saddle Point Nash Equilibrium

- Agent y is a maximizer
- Agent x is a minimizer



A_i's set of control inputs $\mathcal{U}_i$ consists of all feedback controls $u_i(t)$ adapted to $\mathcal{F}^N \triangleq \sigma(x_j(\tau); \tau \le t, 1 \le j \le N)$.

Definition The set of controls $\mathcal{U}^0 = \{u_i^0; 1 \le i \le N\}$ generates a *Nash Equilibrium* w.r.t. the costs $\{J_i; 1 \le i \le N\}$ if, for each i,

$$J_i(u_i^0, u_{-i}^0) = \inf_{u_i \in \mathcal{U}_i} J_i(u_i, u_{-i}^0)$$

Part III – ε-Nash Equilibrium

A_i's set of control inputs $\mathcal{U}_i$ consists of all feedback controls $u_i(t)$ adapted to $\mathcal{F}^N \triangleq \sigma(x_j(\tau); \tau \le t, 1 \le j \le N)$.

Definition Given $\varepsilon > 0$, the set of controls $\mathcal{U}^0 = \{u_i^0; 1 \le i \le N\}$ generates an *ε-Nash Equilibrium* w.r.t. the costs $\{J_i; 1 \le i \le N\}$ if, for each i,

$$J_i(u_i^0, u_{-i}^0) - \varepsilon \le \inf_{u_i \in \mathcal{U}_i} J_i(u_i, u_{-i}^0) \le J_i(u_i^0, u_{-i}^0)$$

Part III - Assumptions

- H2 $\int_{\mathcal{A}} [b^2\gamma]/[r\beta_1(a)\beta_2(a)]dF(a) < 1$.
- H3 All agents have independent zero mean initial conditions. In addition, $\{w_i, 1 \leq i \leq N\}$ are mutually independent and independent of the initial conditions.
- H4 The interval $\mathcal{A}$ satisfies:
  - $\mathcal{A}$ is a compact set and $\inf_{a\in\mathcal{A}} \beta_1(a) \geq \epsilon$ for some $\epsilon > 0$.
  - $\sup_{i\geq 1}[\sigma_i^2 + \mathbb{E}x_i^2(0)] < \infty$.

Part III – NCE Control: First Main Result

Theorem (MH, PEC, RPM, 2003)

- The NCE Control Algorithm generates a set of controls $\mathcal{U}_{nce}^{N} = \{u_i^0; 1 \le i \le N\}, \;\; 1 \le N < \infty$, with

$$u_i^0 = -\frac{b}{r}(\Pi_i x_i + s_i)$$

  s.t.

  (i) The NCE Equations have a unique solution.

  (ii) All agent systems $S(A_i), \;\; 1 \le i \le N,$ are second order stable.

  (iii) $\{\mathcal{U}_{nce}^{N}; 1 \le N < \infty\}$ yields an ε-Nash equilibrium for all ε, i.e. $\forall \varepsilon > 0 \;\; \exists N(\varepsilon)$ s.t. $\forall N \ge N(\varepsilon)$

$$J_i(u_i^0, u_{-i}^0) - \varepsilon \le \inf_{u_i \in \mathcal{U}_i} J_i(u_i, u_{-i}^0) \le J_i(u_i^0, u_{-i}^0),$$

  where $u_i \in \mathcal{U}_i$ is adapted to $\mathcal{F}^N$.

Part III – NCE Control: Key Observations

- The information set for NCE Control is minimal and completely local since Agent A_i's control depends on:
  (i) Agent A_i's own state: $x_i(t)$
  (ii) Statistical information $F(\theta)$ on the dynamical parameters of the mass of agents.

  Hence NCE Control is truly decentralized.

- All trajectories are statistically independent for all finite population sizes N.

Part III – Key Intuition:

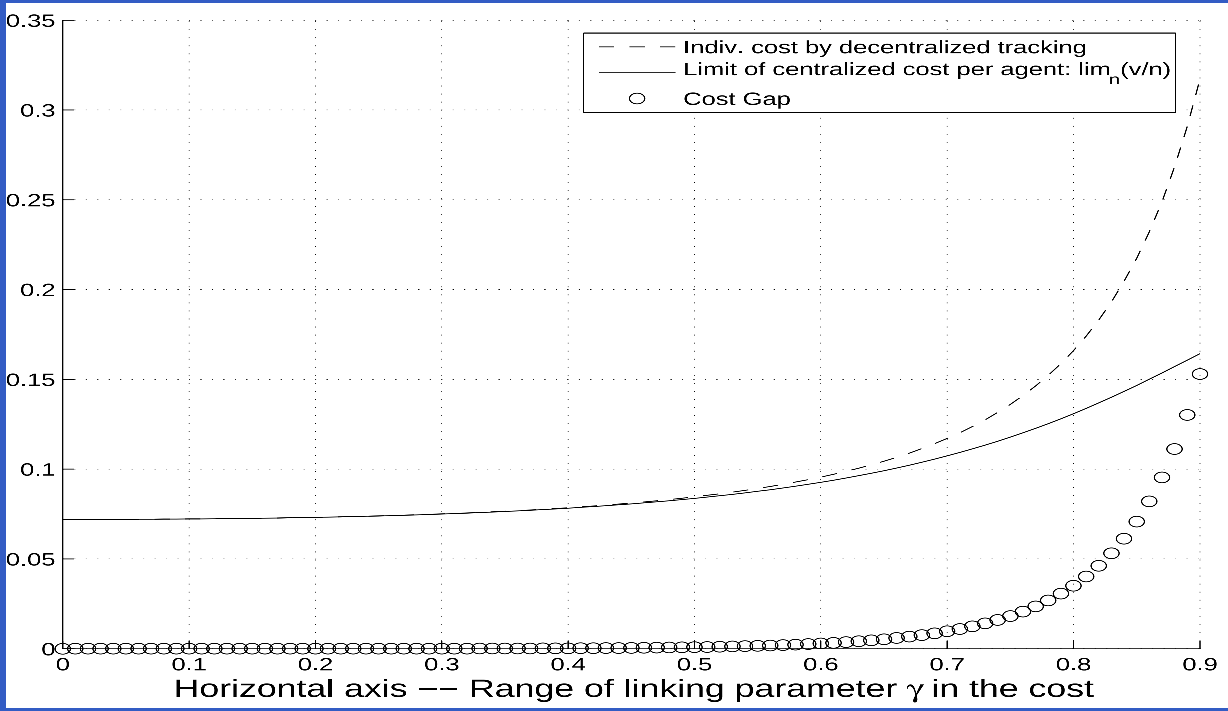
Fixed point theorem shows all agents' competitive control actions (based on local states and assumed mass behaviour) reproduce that mass behaviour

Hence:

NCE Stochastic Control results in a stochastic dynamic Nash Equilibrium

Part III – Centralized versus Decentralized NCE Costs

- Per agent infinite population NCE cost $:= v_{indiv}$ (the dashes)
- Centralized cost: $J = \sum_{i=1}^{N} J_i$
  Optimal Centralized Control Cost per agent $:= v_N/N$
  Per agent infinite population Centralized Cost $\bar{v} = \lim_{N\to\infty} v_N/N$
  (the solid curve)
- Cost gap as a function of linking parameter γ: $v_{indiv} - \bar{v}$ (the circles)



Part IV — Localized NCE Theory

- Motivation: Control of physically distributed controllers & sensors
  - Geographically distributed markets
- Individual dynamics: (note vector states and inputs!)

$$dx_{i,p_i}(t) = \mathbf{A}_i x_{i,p_i}(t)dt + \mathbf{B}_i u_{i,p_i}(t)dt + \mathbf{C} dw_{i,p_i}(t), \quad 1 \leq i \leq N, \quad t \geq 0$$

  - State of Agent A_i at p_i is denoted x_{i,p_i} and indep. hyps. assumed.
  - Dynamical parameters independent of location, hence A_{i,p_i} = A_i
- Interaction locality:

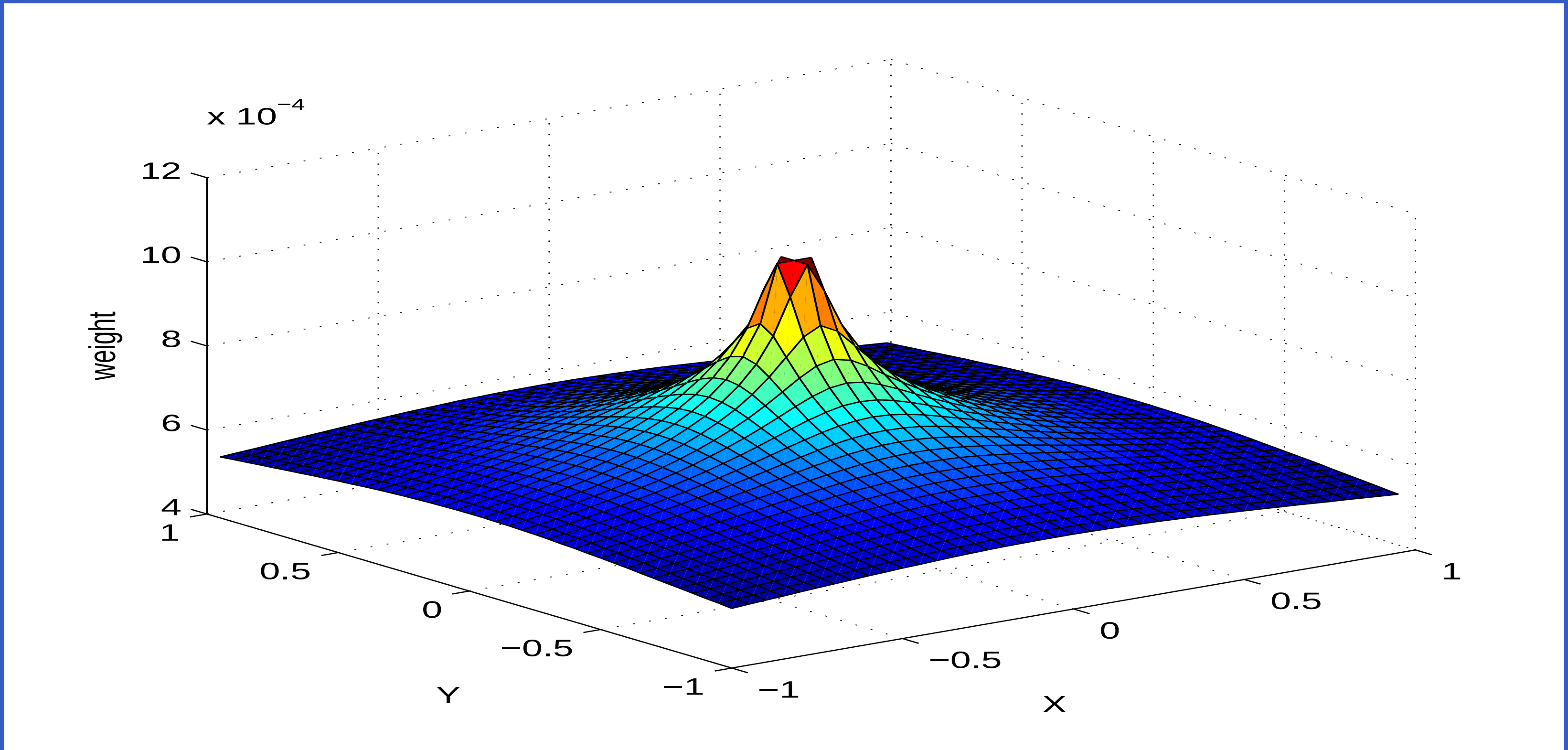
$$J_{i,p_i} = \mathbb{E} \int_0^\infty e^{-\rho t} \left\{ [x_{i,p_i} - \tilde{\Phi}_{i,p_i}]^{\mathrm{T}} \mathbf{Q} [x_{i,p_i} - \tilde{\Phi}_{i,p_i}] + u_{i,p_i}^{\mathrm{T}} \mathbf{R} u_{i,p_i} \right\} dt,$$

where $\tilde{\Phi}_{i,p_i} = \gamma(\sum_{j \neq i}^{N} \omega_{p_i p_j}^{(N)} x_{j,p_j} + \eta)$

Part IV – Example of Weight Allocation

Consider the 2-D interaction:

- Partition $[-1,1] \times [-1,1]$ into a 2-D lattice
- Weight decays with distance by the rule $\omega^{(N)}_{p_i p_j} = c|p_i - p_j|^{-\lambda}$ where c is the normalizing factor and $\lambda \in (0,2)$



Part IV – Assumptions on Weight Allocations

- C1 The weight allocation satisfies the conditions:

(a) $\omega_{p_i p_j}^{(N)} \geq 0, \quad \forall i, j,$

(b) $\sum_{j \neq i}^{N} \omega_{p_i p_j}^{(N)} = 1, \quad \forall i,$

(c) $\epsilon_N^{\omega} \triangleq \sup_{1 \leq i \leq N} \sum_{j \neq i}^{N} |\omega_{p_i p_j}^{(N)}|^2 \longrightarrow 0,$ as $N \rightarrow \infty.$

- Condition (c) implies the weights cannot become highly concentrated on a small number of neighbours.

Part IV – Localized NCE: Parametric Assumptions

- For the continuum of systems: $\theta = [\mathbf{A}_\theta, \mathbf{B}_\theta, \mathbf{C}]$, assume $\left[\mathbf{Q}^{1/2}, \mathbf{A}_\theta\right]$ observable, $[\mathbf{A}_\theta, \mathbf{B}_\theta]$ controllable

- As in the standard case, let $\mathbf{\Pi}_\theta > 0$ satisfy the Riccati equation:

$$\rho \mathbf{\Pi}_\theta = \mathbf{A}_\theta^{\mathrm{T}} \mathbf{\Pi}_\theta + \mathbf{\Pi}_\theta \mathbf{A}_\theta - \mathbf{\Pi}_\theta \mathbf{B}_\theta \mathbf{R}^{-1} \mathbf{B}_\theta^{\mathrm{T}} \mathbf{\Pi}_\theta + \mathbf{Q}$$

Denote $\mathbf{A}_{1,\theta} = \mathbf{A}_\theta - \mathbf{B}_\theta \mathbf{R}^{-1} \mathbf{B}_\theta^{\mathrm{T}} \mathbf{\Pi}_\theta$ and $\mathbf{A}_{2,\theta} = \mathbf{A}_\theta - \mathbf{B}_\theta \mathbf{R}^{-1} \mathbf{B}_\theta^{\mathrm{T}} \mathbf{\Pi}_\theta - \rho \mathbf{I}$

C2 The $\mathbf{A}_1$ eigenvalues satisfy $\lambda.(\mathbf{A}_1) < 0$ (necessarily $\lambda.(\mathbf{A}_2) < 0$), and the analogy of $\int_{\mathbf{\Theta}} [\gamma b_\theta^2]/[r\beta_{1,\theta}\beta_{2,\theta}] dF(\theta) < 1$ holds.

- Take $[\underline{p}, \overline{p}]$ as the locality index set (in $\mathbb{R}^1$).

C3 There exist location weight distributions $F_p(\cdot), p \in [\underline{p}, \overline{p}]$, satisfying certain technical conditions.

Part IV – Localized NCE: Parametric Assumptions

Convergence of Emp. Dynamical Parameter and Weight Distributions:

1. Parameter Distribution:

$$F_N(\theta|p) = \lim_{N\to\infty} \frac{1}{N} \sum_1^N \mathbb{I}(\theta_i \leq \theta) \to F(\theta) \qquad \text{weakly.}$$

e.g. $\quad F(\theta) \sim N(\bar{\theta}, \Sigma)$

2. Location Weight Distribution: $w \lim_{N\to\infty} \sum_{p_j \leq q} \omega_{pp_j}^{(N)} = F_p(q)$

$$\lim_{N\to\infty} \sum_{j=1}^N \omega_{p_i p_j}^{(N)} \mathbb{I}_{(p_j \in S, \theta_j \in H)} = \int_{q\in S} \int_{\theta \in H} dF(\theta) dF_p(q)|_{p=p_i}$$

e.g. $\quad \omega_{p_i p_j}^{(N)} = \frac{c_i(N)}{|p_j - p_i|^{1/2}} \qquad 1 \leq i \neq j \leq N$

Part IV – NCE Equations with Interaction Locality

Each agent $A_{\theta,p}$ computes:

$$\rho s_{\theta,p} = \frac{ds_{\theta,p}}{dt} + \mathbf{A}_\theta^{\mathrm{T}} s_{\theta,p} - \mathbf{\Pi}_\theta \mathbf{B}_\theta \mathbf{R}^{-1} \mathbf{B}_\theta^{\mathrm{T}} s_{\theta,p} - x_p^*$$

$$\frac{d\bar{x}_{\theta,p}}{dt} = (\mathbf{A}_\theta - \mathbf{B}_\theta \mathbf{R}^{-1} \mathbf{B}_\theta^{\mathrm{T}} \mathbf{\Pi}_\theta)\bar{x}_{\theta,p} - \mathbf{B}_\theta \mathbf{R}^{-1} \mathbf{B}_\theta^{\mathrm{T}} s_{\theta,p}, \quad \bar{x}(0) = 0$$

$$\bar{x}_p(t) = \int_{\mathbf{\Theta}} \int_P \overline{x}_{\theta',p'}(t) dF(\theta') dF_p(p')$$

$$x_p^*(t) = \gamma(\bar{x}_p(t) + \eta), \quad t \geq 0$$

- The Mean Field effect now depends on the locations of the agents.

Part IV – Localized NCE: Main Result

Theorem (HCM, 2008) Under C1-C4:

Let the Localized NCE Control Algorithm generate a set of controls

$$\mathcal{U}_{nce}^N = \{u_i^0; 1 \le i \le N\}, \ \ 1 \le N < \infty,$$

where for agent A_{i,p_i}, $\quad u_i^0 = -\mathbf{R}^{-1}\mathbf{B}_i^{\mathrm{T}}(\mathbf{\Pi}_\theta x_i + s_{i,p_i})$

s.t. s_{i,p_i} is given by the Localized NCE Equation System via $\theta := i, p := p_i$.

Then,

(i) There exists a unique bounded solution $(s_p(\cdot), \bar{x}_p(\cdot), x_p^*(\cdot), p \in \mathcal{A})$ to the NCE equation system.

(ii) All agents systems $S(A_i)$, $\ 1 \le i \le N$, are second order stable.

(iii) $\{\mathcal{U}_{nce}^N; 1 \le N < \infty\}$ yields an ε-Nash Equilibrium for all ε, i.e. $\forall \varepsilon > 0 \ \ \exists N(\varepsilon)$ s.t. $\forall N \ge N(\varepsilon)$

$$J_i(u_i^0, u_{-i}^0) - \varepsilon \le \inf_{u_i \in \mathcal{U}_i} J_i(u_i, u_{-i}^0) \le J_i(u_i^0, u_{-i}^0),$$

where $u_i^0 \in \mathcal{U}_{nce}^N$, and $u_i \in \mathcal{U}_i$ is adapted to $\mathcal{F}^N$.

- 1-D & 2-D Cases:

  Locally: $\omega_{p_i p_j}^{(N)} = |p_i - p_j|^{-0.5}$

  Connection to other population: $\omega_{p_i p_j}^{(N)} = 0$

  For Population 1: $A = 1$, $B = 1$

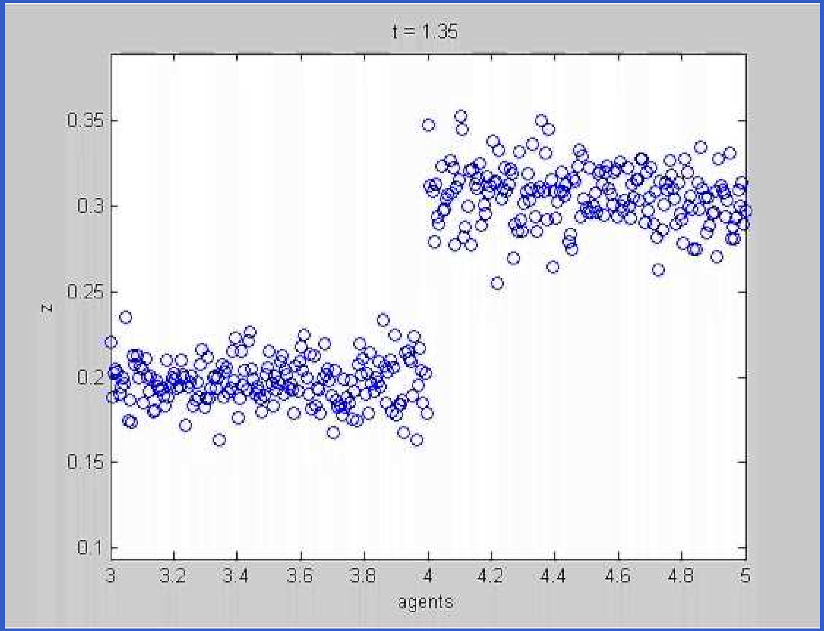
  For Population 2: $A = 0$, $B = 1$

  $\eta = 0.25$

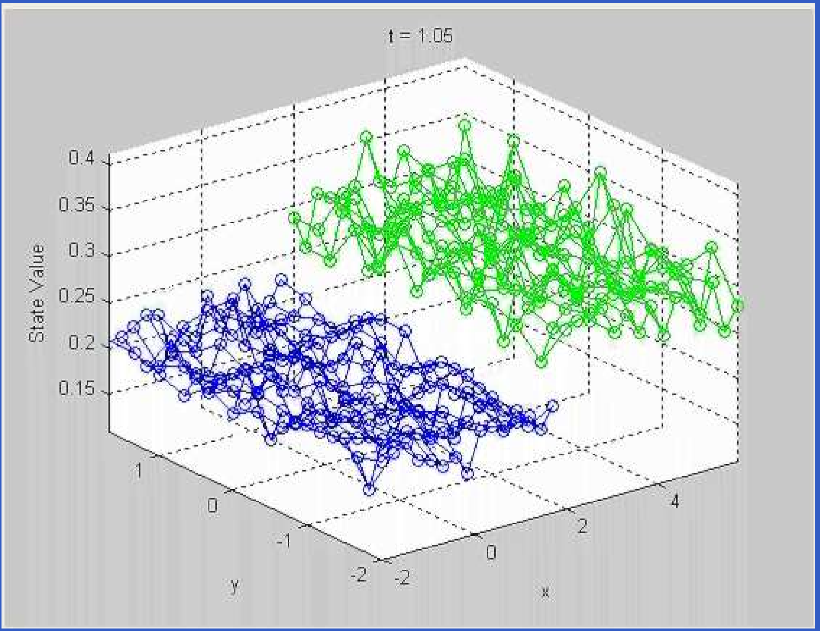
  - At $20th$ second two populations merge:

    Globally: $\omega_{p_i p_j}^{(N)} = |p_i - p_j|^{-0.5}$

Part IV – Separated and Linked Populations



Line



2-D System

Part V – Nonlinear MF Systems

- In the infinite population limit, a representative agent satisfies a controlled McKean-Vlasov Equation:

$$dx_t = f[x_t, u_t, \mu_t]dt + \sigma dw_t, \quad 0 \le t \le T$$

where $f[x, u, \mu_t] = \int_{\mathbb{R}} f(x, u, y)\mu_t(dy)$, with x_0, μ_0 given

$\mu_t(\cdot) =$ distribution of population states at $t \in [0, T]$.

- In the infinite population limit, individual Agents' Costs:

$$J(u, \mu) \triangleq E \int_0^T L[x_t, u_t, \mu_t]dt,$$

where $L[x, u, \mu_t] = \int_{\mathbb{R}} L(x, u, y)\mu_t(dy)$.

Part V – Mean Field and McK-V-HJB Theory

- Mean Field HJB equation (HMC, 2006):

$$-\frac{\partial V}{\partial t} = \inf_{u \in U} \left\{ f[x,u,\mu_t]\frac{\partial V}{\partial x} + L[x,u,\mu_t] \right\} + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x^2}$$

$$V(T,x) = 0, \qquad (t,x) \in [0,T) \times \mathbb{R}.$$

$\Rightarrow$ Optimal Control: $\quad u_t = \varphi(t,x|\mu_t), \quad (t,x) \in [0,T] \times \mathbb{R}.$

- Closed-loop McK-V equation:

$$dx_t = f[x_t, \varphi(t,x|\mu_\cdot), \mu_t]dt + \sigma dw_t, \quad 0 \le t \le T.$$

Part V – Mean Field and McK-V-HJB Theory

- Mean Field HJB equation (HMC, 2006):

$$-\frac{\partial V}{\partial t} = \inf_{u \in U} \left\{ f[x, u, \mu_t] \frac{\partial V}{\partial x} + L[x, u, \mu_t] \right\} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2}$$

$$V(T, x) = 0, \qquad (t, x) \in [0, T) \times \mathbb{R}.$$

$\Rightarrow$ Optimal Control: $\quad u_t = \varphi(t, x | \mu_t), \quad (t, x) \in [0, T] \times \mathbb{R}.$

- Closed-loop McK-V equation:

$$dx_t = f[x_t, \varphi(t, x | \mu_\cdot), \mu_t] dt + \sigma dw_t, \quad 0 \le t \le T.$$

Yielding Nash Certainty Equivalence Principle expressed in terms of McKean-Vlasov HJB Equation, hence achieving the highest Great Name Frequency possible for a Systems and Control Theory result.

Part VI – Adaptive NCE Theory:

Certainty Equivalence Stochastic Adaptive Control (SAC) replaces unknown parameters by their recursively generated estimates

Key Problem:

To show this results in asymptotically optimal system behaviour

Part VI – Adaptive NCE: Self & Popn. Identification

- Self-Identification Case: Individual parameters to be estimated:
  - Agent A_i estimates own dynamical parameters: $\theta_i^{\mathrm{T}} = (\mathbf{A}_i, \mathbf{B}_i)$
  - Parameter distribution of mass of agents is given:
    $F_\zeta = F_\zeta(\theta), \quad \theta \in \mathcal{A} \subset\subset \mathbb{R}^{n(n+m)}, \quad \zeta \in P \subset\subset \mathbb{R}^p$.
- Self + Population Identification Case: Individual & population distribution parameters to be estimated:
  A_i observes a random subset $\mathrm{Obs}_i(N)$ of all agents s.t.
  $$|\mathrm{Obs}_i(N)|/N \to 0 \text{ as } N \to \infty$$
  Agent A_i estimates:
  - Its own parameter θ_i:
  - Population distribution parameter ζ: $F_\zeta(\theta), \quad \theta \in \mathcal{A}, \quad \zeta \in P$

Part VI – SAC NCE Control Algorithm

- WLS Equations:

$$\hat{\theta}_i^{\mathrm{T}}(t) = [\hat{\mathbf{A}}_i, \hat{\mathbf{B}}_i](t), \qquad \psi_t^{\mathrm{T}} = [x_t^{\mathrm{T}}, u_t^{\mathrm{T}}],$$

$$d\hat{\theta}_i(t) = a(t)\Psi_t\psi(t)[dx_i^{\mathrm{T}}(t) - \psi_t^{\mathrm{T}}\hat{\theta}_i(t)dt] \qquad 1 \le i \le N$$

$$d\Psi_t = -a(t)\Psi_t\psi_t\psi_t^{\mathrm{T}}\Psi_t dt$$

- Long Range Average (LRA) Cost Function:

$$J_i(\hat{u}_i, \hat{u}_{-i})$$

$$= \lim_{T\to\infty} \sup \frac{1}{T}\int_0^T \left\{[x_i(t) - \Phi_i(t)]^{\mathrm{T}}\mathbf{Q}[x_i(t) - \Phi_i(t)] + \hat{u}_i^{\mathrm{T}}(t)\mathbf{R}\hat{u}_i(t)\right\} dt$$

$$1 \le i \le N, \quad a.s.$$

Part VI – SAC NCE Control Algorithm

- Solve the NCE Equations generating $x^*(t, F_\zeta)$. Note: F_ζ known
- Agent A_i's NCE Control:
  - $\hat{\Pi}_{i,t}$: Riccati Equation:

$$\hat{\mathbf{A}}_{i,t}^{\mathrm{T}}\hat{\mathbf{\Pi}}_{i,t} + \hat{\mathbf{\Pi}}_{i,t}\hat{\mathbf{A}}_{i,t} - \hat{\mathbf{\Pi}}_{i,t}\hat{\mathbf{B}}_{i,t}\mathbf{R}^{-1}\hat{\mathbf{B}}_{i,t}^{\mathrm{T}}\hat{\mathbf{\Pi}}_{i,t} + \mathbf{Q} = 0$$

  - $s_i(t, \hat{\theta}_{i,t})$: Mass Control Offset:

$$\frac{ds_i(t, \hat{\theta}_{i,t})}{dt} = (-\hat{\mathbf{A}}_{i,t}^{\mathrm{T}} + \hat{\mathbf{\Pi}}_i\hat{\mathbf{B}}_{i,t}\mathbf{R}^{-1}\hat{\mathbf{B}}_{i,t}^{\mathrm{T}})s_i(t, \hat{\theta}_{i,t}) + x^*(t, F_\zeta)$$

  - The control law from Certainty Equivalence Adaptive Control:

$$\hat{u}_i(t) = -\mathbf{R}^{-1}\hat{\mathbf{B}}_{i,t}^{\mathrm{T}}(\hat{\mathbf{\Pi}}_{i,t}x_i(t) + s_i(t, \hat{\theta}_{i,t})) + \xi_k\left[\epsilon_i(t) - \epsilon_i(k)\right]$$

Dither weighting: $\xi_k^2 = \frac{\log k}{\sqrt{k}}, \quad k \geq 1$ $\qquad \epsilon_i(t) =$ Wiener Process

Zero LRA cost dither after (Duncan, Guo, Pasik-Duncan; 1999)

Part VI – SAC NCE Theorem (Self-Identification Case)

Theorem (Kizilkale & PEC; after (LRA-NCE) Li & Zhang, 2008)

The NCE Adaptive Control Algorithm generates a set of controls

$$\hat{\mathcal{U}}^N_{nce} = \{\hat{u}_i; 1 \le i \le N\}, \;\; 1 \le N < \infty,$$

s.t.

(i) $\hat{\theta}_{i,t}(x_i^t) \to \theta_i$ w.p.1 as $t \to \infty, \;\; 1 \le i \le N$ (strong consistency)

(ii) All agent systems $S(A_i), \;\; 1 \le i \le N,$ are $\mathrm{LRA} - \mathrm{L}^2$ stable w.p.1

(iii) $\{\hat{\mathcal{U}}^N_{nce}; 1 \le N < \infty\}$ yields a (strong) ε-Nash equilibrium for all ε

(iv) Moreover $J_i(\hat{u}_i, \hat{u}_{-i}) = J_i(u_i^0, u_{-i}^0)$ w.p.1, $\;\; 1 \le i \le N$

Part VI – SAC NCE (Self + Population Identification Case)

Theorem (AK & PEC, 2009)

Assume each agent A_i:

(i) Observes a random subset $\mathrm{Obs}_i(N)$ of the total population N s.t. $|\mathrm{Obs}_i(N)|/N \to 0, \quad N \to \infty,$

(ii) Estimates own parameter θ_i via WLS

(iii) Estimates the population distribution parameter $\hat{\zeta}_i$ via RMLE

(iv) Computes $u_i^0(\hat{\theta}, \hat{\zeta})$ via the extended NCE eqns + dither where

$$\bar{x}(t) = \int_{\mathcal{A}} \bar{x}_\theta(t) dF_{\hat{\zeta}_i}(\theta)$$

Theorem (AK & PEC, 2009) ctn.

Under reasonable conditions on population dynamical parameter distribution $F_\zeta(.)$, as $t \to \infty$ and $N \to \infty$:

(i) $\hat{\zeta}_i(t) \to \zeta \in P$ a.s. and hence, $F_{\hat{\zeta}_i(t)}(\cdot) \to F_\zeta(\cdot)$ a.s. (weak convergence on P), $\quad 1 \leq i \leq N$

(ii) $\hat{\theta}_i(t, N) \to \theta_i$ a.s. $1 \leq i \leq N$

and the set of controls $\{\hat{\mathcal{U}}^N_{nce};\ \ 1 \leq N < \infty\}$ is s.t.

(iii) Each $S(A_i), 1 \leq i \leq N$, is an $\mathrm{LRA} - \mathrm{L}^2$ stable system.

(iv) $\{\hat{\mathcal{U}}^N_{nce}; 1 \leq N < \infty\}$ yields a (strong) ε-Nash equilibrium for all ε

(v) Moreover $\quad J_i(\hat{u}_i, \hat{u}_{-i}) = J_i(u^0_i, u^0_{-i}) \quad$ w.p.1, $\quad 1 \leq i \leq N$

Part VI – SAC NCE Animation

- 400 Agents
- System matrices $\{A_k\}, \{B_k\}, \quad 1 \leq k \leq 400$

$$A \triangleq \begin{bmatrix} -0.2 + a_{11} & -2 + a_{12} \\ 1 + a_{21} & 0 + a_{22} \end{bmatrix} \qquad B \triangleq \begin{bmatrix} 1 + b_1 \\ 0 + b_2 \end{bmatrix}$$

- Population dynamical parameter distribution a_{ij}'s and b_i's are independent.

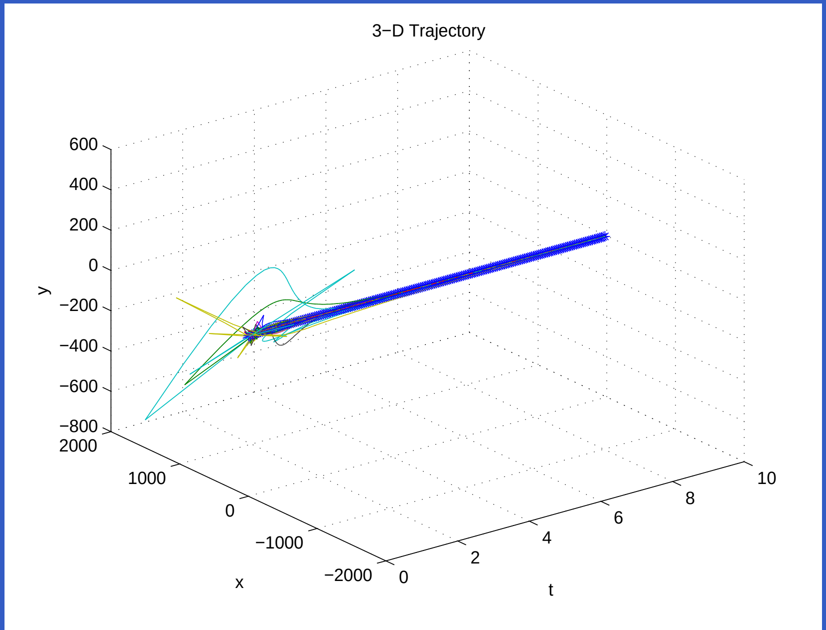
$$a_{ij} \sim N(0, 0.5) \qquad b_i \sim N(0, 0.5)$$

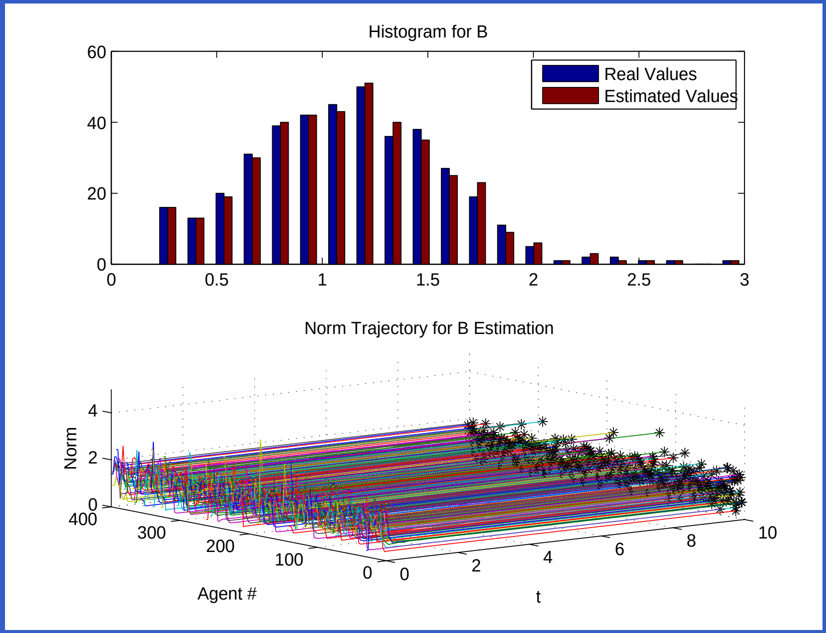
- Population distribution parameters:

$$\bar{a}_{11} = -0.2, \quad \sigma^2_{a_{11}} = 0.5, \quad \bar{b}_{11} = 1, \quad \sigma^2_{b_{11}} = 0.5 \quad \text{etc.}$$

Part VI – SAC NCE Animation

- All agents performing individual parameter and population distribution parameter estimation
- Each of 400 agents observing its own 20 randomly chosen agents' outputs and control inputs





Animation of Trajectories

Summary

- NCE Theory solves a class of decentralized decision-making problems with many competing agents.

- Asymptotic Nash Equilibria are generated by the NCE Equations.

- Key intuition:

  Single agent's control = feedback of stochastic local (rough) state + feedback of deterministic global (smooth) system behaviour

- NCE Theory extends to (i) localized problems, (ii) stochastic adaptive control, (iii) leader - follower systems.

Future Directions

- Further development of Minyi Huang's large and small players extension of NCE Theory
- Egoists and altruists version of NCE Theory
- Mean Field stochastic control of non-linear (McKean-Vlasov) systems
- Extension of SAC MF Theory in richer game theory contexts
- Development of MF Theory towards economic and biological applications
- Development of large scale cybernetics: Systems and control theory for competitive and cooperative systems

Thank You !

谢 谢