Mean Field Stochastic Control Systems

Peter E. Caines
McGill University

Mean Field Games
and Related Topics

Rome, 2011

Dipartimento di Matematica, SAPIENZA Università di Roma

Co-Authors

Minyi Huang

Roland Malhamé

Collaborators & Students

Mojtaba Nourian

Arman Kizilkale

Arthur Lazarte

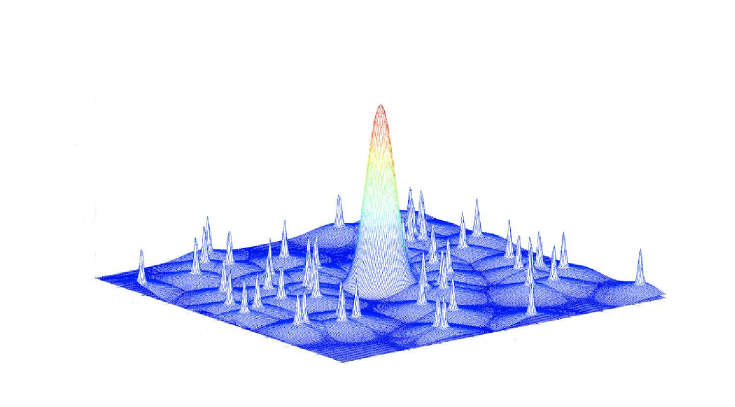
Zhongjing Ma

Overview

- Overall Objective:
    - Develop a theory of decentralized decision-making in stochastic dynamical systems with many competing or cooperating agents
- Outline:
    - A motivating control problem from code division multiple access (CDMA) uplink power control
    - Basic concepts of Mean Field (MF) control:
        - The Nash Certainty Equivalence - MF (NCE - MF) methodology
        - Main NCE results for Linear-Quadratic-Gaussian (LQG) systems
    - Adaptive NCE System Theory
    - Cucker-Smale Type Flocking: Stationary Solutions and Perturbation Analysis

Part 1 – CDMA Power Control

Base Station & Individual Agents



Part 1 – CDMA Power Control

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

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

  Transmitted power = channel attenuation × power

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

  (Charalambous, Menemenlis; 1999)

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

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

- How to optimize all the individual SIR's since it is self defeating for everyone to increase their power?
- Idea: Use large population properties of the system together with basic notions of game theory

Part 2 – Large Popn. Models with Game Theory Features

- Economic models: Cournot-Nash equilibria (Lambson)
- Advertising competition: game models (Erickson)
- Wireless network res. alloc.: (Alpcan et al., Altman, HCM)
- Admission control in communication networks: (Ma, MC)
- Public health: voluntary vaccination games (Bauch & Earn)
- Biology: stochastic PDE swarming models (Bertozzi et al.)
- Sociology: urban economics (Brock and Durlauf et al.)
- Renewable Energy: charging control of PEVs (Ma et al.)

Part 2 – 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, 2005 -, Adlakha, Johari & Goldsmith, 2008 -)
- Mean Field Games: Stochastic control of many agent systems with applications to finance (Lasry & Lions, 2006 -, Achdou, Cardaliaguet, Capuzzo-Dolcetta, Buckdahn, 2006 -)
- Mean Field Control of Oscillators: Controlled synchronization, chaotic motion via MF game control of populations of oscillators. Phase changes: NL-MF equation triple (The Illinois Four/Five: Yin/Yang, Mehta, Meyn, Shanbhag, 2009 -)
- Mean Field MDP Games on Networks: Exchangeability hypothesis; propagation of chaos in the popn. limit; evolutionary games. (Tembine et al., 2009 -)

Part 2 – Basic LQG Game Problem

- Massive game theoretic control systems: Large ensembles of partially regulated competing agents
- Fundamental issue: The relation between the actions of each individual agent and the resulting mass behavior

Individual Agent's Dynamics:

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

(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 2 – Basic LQG Game Problem

Individual Agent's Cost:

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

$$\nu \triangleq \gamma . (\tfrac{1}{N} \textstyle\sum_{k \neq i}^{N} x_k + \eta)$$

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 2 – Preliminary Optimal LQG Tracking

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

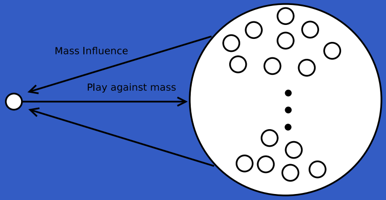
Optimal Tracking Control: $u_i = -\frac{b}{r}(\Pi_i x_i + s_i)$

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

Part 2 – 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 2 – LQG-NCE Equation Scheme

The Fundamental NCE Equation System

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

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

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

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

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

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

- 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 holds for all sufficiently small γ.

Part 2 – NCE Feedback Control

The MF (NCE) Control Law Specification

The Finite System of N Agents with Dynamics:

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

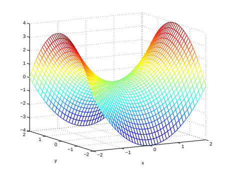
Let $u_{-i} \triangleq (u_1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_N)$; then

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

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

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

Part 2 – Nash Equilibrium



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

The Information Pattern:

$\mathcal{F}_i \triangleq \sigma(x_i(\tau); \tau \leq t)$ $\quad\quad$ $\mathcal{F}^N \triangleq \sigma(x_j(\tau); \tau \leq t, 1 \leq j \leq N)$

$\mathcal{F}_i$ adapted control: $\mathcal{U}_{loc,i}$ $\quad\quad$ $\mathcal{F}^N$ adapted control: $\mathcal{U}$

The Equilibria:

The set of controls $\mathcal{U}^0 = \{u_i^0;\ u_i^0$ adapted to $\mathcal{U}_{loc,i},\ 1 \leq i \leq N\}$ generates a (strong) Nash Equilibrium w.r.t. the costs $\{J_i; 1 \leq i \leq N\}$ and $\mathcal{U}$ if, for each i,

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

ϵ-Nash Equilibria:

Given $\epsilon > 0$, the set of controls $\mathcal{U}^0 = \{u_i^0; 1 \leq i \leq N\}$ generates a (strong) ϵ-Nash Equilibrium w.r.t. the costs $\{J_i; 1 \leq i \leq N\}$ and $\mathcal{U}$ if for each i,

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

Part 2 – NCE Control: First Main Result

Theorem: (MH, PEC, RPM, IEEE CDC 2003, IEEE TAC 2007)

Subject to technical conditions (s.t.t.c.), the NCE Equations have a unique solution yielding the set of NCE Control Laws

$$\mathcal{U}_0^N = \{u_i^0; 1 \le i \le N\}, \;\; 1 \le N < \infty, \text{ where}$$

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

which are s.t.

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

(ii) $\{\mathcal{U}_0^N; 1 \le N < \infty\}$ yields a (strong) ϵ-Nash equilibrium for all ϵ, i.e. $\forall \epsilon > 0 \;\; \exists N(\epsilon)$ s.t. $\forall N \ge N(\epsilon)$

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

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

□

Part 2 – 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.

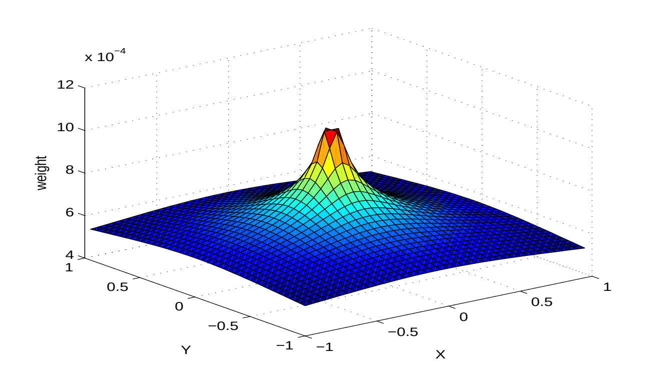
  Hence NCE Control is truly decentralized.

- It is a feature of this theory that the NCE control laws $\mathcal{U}^0$ result in statistically independent trajectories for all finite population sizes N.

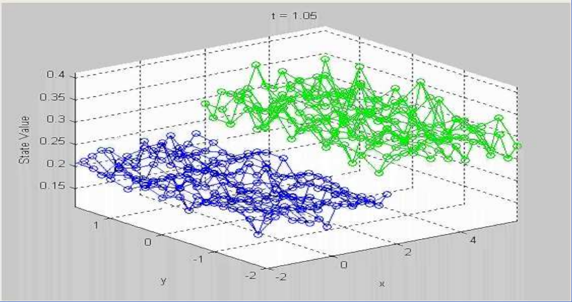
Part 3 – Localization of Influence

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 3 – Separated and Linked Populations



2-D System

- NL Individual Dynamics (Uniform Finite Population Case):

$$dx_i = \frac{1}{N}\sum_{j=1}^{N} f(x_i, u_i, x_j)dt + \sigma dw_i, \quad 1 \le i \le N$$

- The Finite Population Cost Function for the ith agent:

$$J_i(u_i, u_{-i}) \triangleq \mathbb{E}\int_0^T \Big[\frac{1}{N}\sum_{j=1}^{N} L(x_i, u_i, x_j)\Big]dt$$

$f(\cdot,\cdot,\cdot)$ and $L(\cdot,\cdot,\cdot)$ are nonlinear functions

S.t.t.c.:

- Infinite population limit: 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
$\mu_t(\cdot) =$ density of population states at $t \in [0, T]$.

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

- Mean Field HJB equation (HMC, Communications in Inf. and Systems 2006):

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

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

$$\frac{\partial \mu(t,x)}{\partial t} = -\frac{\partial \{ f[x,u,\mu]\mu(t,x) \}}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 \mu(t,x)}{\partial x^2}$$

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

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

- Mean Field HJB equation (HMC, Communications in Inf. and Systems 2006):

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

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

$$\frac{\partial \mu(t,x)}{\partial t} = -\frac{\partial \{ f[x,u,\mu] \mu(t,x) \}}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 \mu(t,x)}{\partial x^2}$$

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

- 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 (Fokker-Planck-Kolmogorov) Hamilton Jacobi Bellman Equation, hence achieving a Great Name Frequency optimum.

Theorem: (HMC, CIS 2006)

S.t.t.c., the McK-V (FPK) HJB Equations have a unique solution with the best response control given by

$$u_i^0 = \varphi(t, x|\mu_t), \;\; 1 \le i \le N.$$

Furthermore $\{\mathcal{U}_0^N; 1 \le N < \infty\}$ yields a (strong) ϵ-Nash equilibrium for all ϵ,
i.e. $\forall \epsilon > 0 \;\; \exists N(\epsilon)$ s.t. $\forall N \ge N(\epsilon)$

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

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

□

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

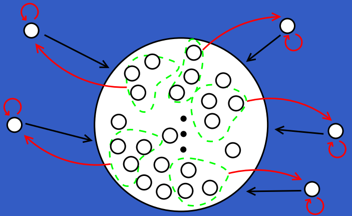
Key Problem:

To show this results in asymptotically optimal system behaviour in the ϵ-Nash sense

Part 5 – Adaptive NCE: Self & Popn. Ident.

Known:

- $\mathbf{Q}$
- $\mathbf{R}$



Observed:

- $x_i(t)$
- $u_i(t)$
- $\{x_j, u_j; j \in Obs_i(N)\}$

Estimated:

- $\hat{\mathbf{A}}_i$, $\hat{\mathbf{B}}_i$
- $F_{\hat{\zeta}}(\theta)$

- A_i observes a random subset $\mathrm{Obs}_i(N)$ of all agents s.t.
  $$|\mathrm{Obs}_i(N)| \to \infty, \ |\mathrm{Obs}_i(N)|/N \to 0 \text{ as } N \to \infty$$
- $\theta_i^{\mathrm{T}} = (\mathbf{A}_i, \mathbf{B}_i)$
- $F_\zeta = F_\zeta(\theta), \quad \theta \in \boldsymbol{\Theta} \subset\subset \mathbb{R}^{n(2n+m)}, \quad \zeta \in P \subset\subset \mathbb{R}^p$

Part 5 – NCE-SAC Cost Function

Each agent's Long Run Average (LRA) Cost Function:

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

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

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

Part 5 – NCE-SAC Control Algorithm

NCE-SAC Control Law for agent A_i, $t \geq 0$:

(I) Self Parameter Identification:
Solve the RWLS Equations for the dynamical parameters $[\hat{\mathbf{A}}_{i,t}, \hat{\mathbf{B}}_{i,t}]$

(II) Popn. Parameter Identification:

(a) Solve the RWLS equations for the dynamical parameters $\hat{\theta}_{i,t}^{[1:N_0]} = [\hat{\mathbf{A}}_{j,t},\, \hat{\mathbf{B}}_{j,t}]$, $j \in Obs_i(N)$

(b) Solve the MLE equation to estimate ζ^0 via $\hat{\zeta}_{i,t}^N = \arg\min_{\zeta \in P} L(\hat{\theta}_{i,t}^{[1:N_0]}; \zeta), \quad N_0 = |Obs_i(N)|$

(c) Solve the set of NCE Equations for all $\theta \in \boldsymbol{\Theta}$ generating $x^* \left(\tau, \hat{\zeta}_{i,t}^N\right), \tau \geq t$

(III) The control law from Certainty Equivalence Adaptive Control:

$$\hat{u}^0(t) = -\mathbf{R}^{-1}\hat{\mathbf{B}}_t^{\mathrm{T}}\left(\hat{\mathbf{\Pi}}_t x(t) + \hat{s}(t)\right) + \xi_k\left[\epsilon(t) - \epsilon(k)\right]$$

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

Theorem: (AK & PEC, 2010)

S.t.t.c., as $t \to \infty$ and $N \to \infty$:

(i) $\hat{\theta}_{i,t} \to \theta_i^0$ w.p.1, $1 \leq i \leq N$

(ii) $\hat{\zeta}_{i,t}^N \to \zeta^0 \in P$ w.p.1

and the set of controls $\{\hat{\mathcal{U}}_0^N;\ 1 \leq N < \infty\}$ is s.t.

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

(iv) $\{\hat{\mathcal{U}}_0^N; 1 \leq N < \infty\}$ yields a (strong) ϵ-Nash equilibrium for all ϵ, i.e. $\forall \epsilon > 0\ \exists N(\epsilon, \omega)$ s.t.

$$J_i(\hat{u}_i, \hat{u}_{-i}) - \epsilon \leq \inf_{u_i \in \mathcal{U}} J_i(u_i, \hat{u}_{-i}) \leq J_i(\hat{u}_i, \hat{u}_{-i}), \qquad w.p.1$$

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

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

□

Part 5 – NCE-SAC Simulation

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

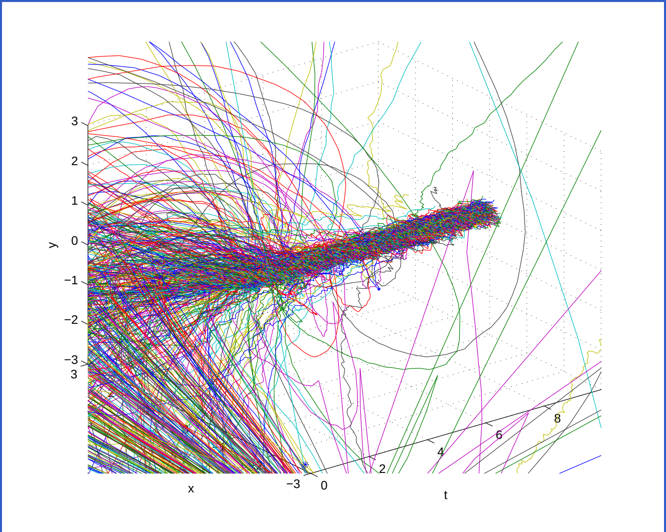
$$A \triangleq \begin{bmatrix} -0.2 + a_{11} & -2 + a_{12} \\ 1 + a_{21} & 0 + a_{22} \end{bmatrix} \quad B \triangleq \begin{bmatrix} 1 + b_1 \\ 0 + b_2 \end{bmatrix}$$

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

Part 5 – NCE-SAC Animation



Animation of Trajectories

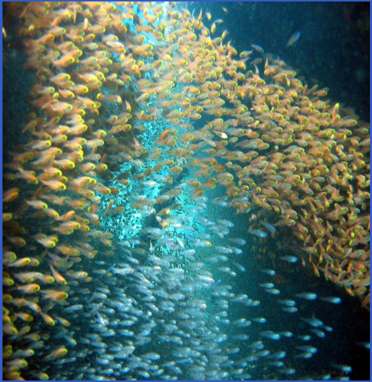
Part 6 – Mean Field Synthesis of Flocking Behaviour

Collective Motion: one of the most widespread phenomenon in nature.

Flocking of birds

Part 6 – Mean Field Synthesis of Flocking Behaviour

Collective Motion:



Schooling of fish

Part 6 – Mean Field Synthesis of Flocking Behaviour

Definition: A group of agents has a flocking behaviour if:

- agents' velocities converge to a common value (e.g., mean of initial velocities), i.e., consensus in velocity,
- the distance between agents remains bounded.

Flocking Models:

1. Microscopic:
   - Individual based (particle like) models (ODEs, SDEs);
   - Local communication with other agents;
   - Example: Cucker-Smale algorithm.
2. Macroscopic:
   - Infinite (continuum) population model;
   - Distribution functions in space-time (PDEs);
   - Example: C-S continuum and hydrodynamic models.

Part 6 – Mean Field Synthesis of Flocking Behaviour

Uncontrolled Cucker-Smale (C-S) Flocking Algorithm (IEEE TAC 2007):

$$\frac{dx_i(t)}{dt} = v_i(t), \qquad 1 \le i \le N,$$

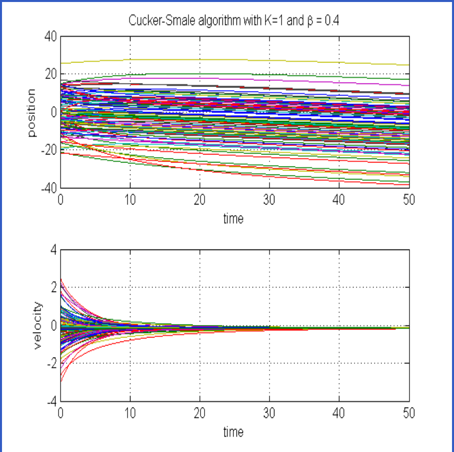
$$\frac{dv_i(t)}{dt} = \frac{1}{N}\sum_{j=1}^{N} a\big(\|x_i(t) - x_j(t)\|\big)\left(v_j(t) - v_i(t)\right),$$

$$a\big(\|x_i(\cdot) - x_j(\cdot)\|\big) \triangleq \frac{1}{(1 + \|x_i(\cdot) - x_j(\cdot)\|^2)^{\beta}}$$

- A special time-varying consensus algorithm (with communication rates $a_{ij}(\cdot)$).
- For $0 \le \beta \le \frac{1}{2}$ we have unconditional (i.e., regardless of initial configurations) flocking.

Part 6 – Mean Field Synthesis of Flocking Behaviour

Simulation: the positions and velocities of a group of 100 agents in the one dimensional C-S algorithm with $\beta = 0.4$.



Part 6 – Mean Field Synthesis of Flocking Behaviour

Continuum Model of the Uncontrolled C-S Algorithm (Ha & Tadmor, 2008, Carrillo et al., 2009):

Advection equation with velocity field $\xi(f)$

$$\frac{\partial f}{\partial t} + v.\nabla_x f = -\nabla_v.\left[\xi(f) f\right],$$

$$\xi(f)(x,v,t) \triangleq \int_{\mathbb{R}^{2n}} \frac{(v-w)}{(1+\|x-y\|^2)^\beta} f(y,w)dydw,$$

where $f(x,v,t)$ is the population density function of agents positioned at (x,t) with velocity v

- For $0 \leq \beta \leq \frac{1}{2}$ we have unconditional flocking.

Part 6 – Mean Field Synthesis of Flocking Behaviour

The flocking problem will now be analyzed using the Mean Field Stochastic Control approach:

- Flocking behaviour synthesized from optimization;
- Global (mass population) optimal control + local (individual) feedback with respect to mass behaviour;
- Nash equilibria between individuals.

For large population this theory reproduces the flocking behaviour of individuals under the (ad hoc) global feedback of the standard formulations

Part 6 – Mean Field Synthesis of Flocking Behaviour

Controlled Dynamic Mean Field Formulation of the C-S Algorithm:

- Dynamics :

$$dx_i(t) = v_i(t)dt, \qquad 1 \leq i \leq N$$
$$dv_i(t) = u_i(t)dt + Cdw_i(t),$$

written as

$$dz_i(t) = \big(Fz_i(t) + Gu_i(t)\big)dt + Ddw_i(t),$$

$$F \triangleq \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}, \quad G \triangleq \begin{pmatrix} 0 \\ I \end{pmatrix}, \quad D \triangleq \begin{pmatrix} 0 \\ C \end{pmatrix}$$

- Cost functions $(1 \leq i \leq N)$:

$$J_i^N \triangleq \limsup_{T\to\infty} \frac{1}{T}\int_0^T \left(\phi_i^N(x_i, v_i; x_{-i}, v_{-i}) + \|u_i\|^2\right) dt,$$

with the normalized cost-coupling:

$$\phi_i^{(N)}(x_i, v_i; x_{-i}, v_{-i}) \triangleq \left\| \frac{1}{\sum_{j=1}^N a(\|x_i - x_j\|)} \sum_{j=1}^N a(\|x_i - x_j\|)(v_j - v_i) \right\|^2$$

Part 6 – Mean Field Synthesis of Flocking Behaviour

MF Formulation: Individual vs Mass

- It is assumed that the generic agent's cost wrt mass converges (implied by various conditions):

$$\phi^{(N)}(x_i, v_i; x_{-i}, v_{-i}) \xrightarrow{N\to\infty} \phi^{\infty}(x, v, t)$$

where ϕ^{∞} only depends on $x = x_i$ and $v = v_i$.

- Replacing $\phi^{(N)}(x_i, v_i; x_{-i}, v_{-i})$ with $\phi^{\infty}(x, v, t)$ reduces the game model to a set of N independent optimal control problems.

- HJB (relative value function) equation: Agent (x, v)

$$\frac{\partial h}{\partial t} + \min_{u\in\mathcal{U}} \left\{ (Fz + Gu)\cdot \nabla_z h + u^T u + \phi^{\infty} + \frac{1}{2}\mathrm{Tr}(DD^T \triangle h) \right\} = \rho^o,$$

$$u^o = \arg\min_{u\in\mathcal{U}} \left\{ (Fz + Gu)\cdot \nabla_z h + u^T u + \phi^{\infty} + \frac{1}{2}\mathrm{Tr}(DD^T \triangle h) \right\},$$

where ρ^o is the optimal cost.

Part 6 – Mean Field Synthesis of Flocking Behaviour

The Nonlinear MF Triple of Equations (NCM, 2010 after Yin et. al., ACC 2010 and HMC, 2006):

$$\text{MF-HJB}: \ \partial_t h(z,t) + \left(Fz - \frac{1}{4} GG^T \nabla_z h(z,t) \right) \cdot \nabla_z h(z,t)$$
$$+ \phi^\infty(z,t) + \frac{1}{2} \mathrm{Tr}\big(DD^T \triangle h(z,t)\big) = \rho^o,$$
$$\text{MF-FPK}: \ \partial_t f(z,t) + \nabla_z \cdot \left(\Big(Fz - \frac{1}{2} GG^T \nabla_z h(z,t) \Big) f(z,t) \right)$$
$$= \frac{1}{2} \mathrm{Tr}\big(DD^T \triangle f(z,t)\big),$$
$$\text{MF-CC}: \quad \phi^\infty(z,t) = \left\| \frac{\int_{\mathbb{R}^{2n}} a(\|x - x'\|)(v' - v) f(x', v', t) dx' dv'}{\int_{\mathbb{R}^{2n}} a(\|x - x'\|) f(x', v', t) dx' dv'} \right\|^2,$$

Best Response: $u^o(\cdot) := -\frac{1}{2} G^T \nabla_z h(z, \cdot)$

Part 6 – Mean Field Synthesis of Flocking Behaviour

Theorem (Maxwellian Stationary Solution of the MF Triple of Equations) (NCM, 2011)):
Assume that the weight function $a(\|x\|)$ is integrable (i.e., $\int_{\mathbb{R}^n} a(\|x\|) < \infty$), and $\Sigma = CC^T > 0$, then the stationary of the MF Triple of equations is given by

$$h_\infty(v) = \phi_\infty(v) = \|v - \mu\|^2, \quad \rho^o = \mathrm{Tr}(CC^T),$$
$$f_\infty(v) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\Big(-\frac{1}{2}(v-\mu)^T\Sigma^{-1}(v-\mu)\Big),$$
$$u^o_\infty(v) = -(v-\mu),$$

where $\mu := \int_{\mathbb{R}^{2n}} vf(x,v,0)dvdx$ is the initial velocity population mean. □

Remark: The following weights satisfy the integrability condition:

- The C-S weights $a(\|x\|) = \frac{1}{(1+\|x\|^2)^\beta}$ for $\beta \geq 1$,
- The Gaussian weights $a(\|x\|) = \exp\big(-\alpha\|x\|^2\big)$ for $\alpha > 0$.

Part 6 – Mean Field Synthesis of Flocking Behaviour

Consensus in Velocity Property of the MF Control Laws:

Theorem (NCM, 2010): By applying the MF control laws,

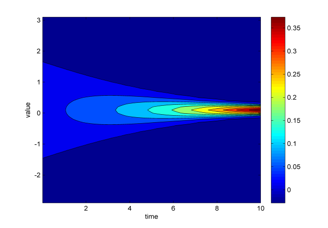
$$u_i^o(\cdot) = -\frac{1}{2}\nabla_v h_\infty(v)\big|_{v=v_i} = -(v_i(\cdot) - \mu),$$

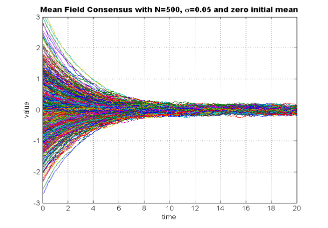
the agents in a finite N population system reach mean-consensus in velocity asymptotically as time goes to infinity, that is to say,

$$\lim_{t\to\infty} \|Ev_i(t) - Ev_j(t)\| = 0, \quad 1 \le i \ne j \le N$$

□

Simulation (Gaussian Initial Density): the continuum (left) and individual (right) models of a scalar MF consensus model ($\mu = 0$ and the noise intensity σ is 0.05)





Stability Analysis of the Stationary Solution for the Scalar Uniform Weights Case ($\beta = 0$):

Perturbation Analysis (based on the Guéant 2009, approach):

$$h_\epsilon(v,t) = h_\infty(v) + \epsilon\, \tilde{h}(v,t),$$
$$f_\epsilon(v,t) = f_\infty(v)\big(1 + \epsilon\, \tilde{f}(v,t)\big),$$
$$\phi_\epsilon^\infty(v,t) = \phi_\infty(v) + \epsilon\, \tilde{\phi}(v,t).$$

Theorem (NCM, 2010 after Guéant, 2009): The linearized MF equation system in the uniform weights case $\beta = 0$:

$$\partial_t \tilde{h}(v,t) = \mathcal{L}_v \tilde{h}(v,t) - \tilde{\phi}(v,t),$$
$$\partial_t \tilde{f}(v,t) = -\frac{1}{\sigma^2}\mathcal{L}_v \tilde{h}(v,t) - \mathcal{L}_v \tilde{f}(v,t),$$
$$\tilde{\phi}(v,t) = -2(v-\mu)\Big(\int_{\mathbb{R}} v \tilde{f}(v,t) f_\infty(v) dv\Big),$$

where the operator $\mathcal{L}_v := (v-\mu)\partial_v - \frac{\sigma^2}{2}\partial_{vv}^2$ has the countable family of Hermite polynomials $\{H_n : n \in \mathbb{N}_0\}$ as eigenfunctions. □

Stability Analysis of the Stationary Solution for the Scalar Uniform Weights Case ($\beta = 0$): Non-Gaussian Initial Densities

Theorem (NCM, 2010): In case $\tilde{f}(v,0) \in \text{span}\big(H_n(v) : n \geq 2\big)$ which gives rise to the non-Gaussian initial conditions:

$$f_\epsilon(v,0) = f_\infty(v)\Big(1 + \epsilon \sum_{n=2}^{\infty} k_n(0) H_n(v)\Big) \in L^2(\mathbb{R}, f_\infty(v)dv),$$

we have $$h_\epsilon(v,t),\ f_\epsilon(v,t) \in L^2(\mathbb{R}, f_\infty(v)dv), \qquad \forall t > 0,$$

and the stationary Gaussian solution is linearly asymptotically stable, that is to say,

$$\lim_{t\to\infty} \|h_\epsilon(v,t)\|_{L^2} = \lim_{t\to\infty} \|f_\epsilon(v,t)\|_{L^2} = 0$$

□

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) egoist-altruist, major agent-minor agent systems, (iv) leader-follower systems, (v) consensus and flocking behaviour.

Future Directions

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