Collaborators

Peter Caines
Minyi Huang
Arman Kizilkale
Mojtaba Nourian
Outline

- **Origins**: foundational papers, motivations and roots
- **Initial mathematical results for a continuum of agents:**
  - The LQG framework
  - The nonlinear stochastic framework
- **What it means for finite populations of interacting agents**: the $\epsilon$-Nash equilibria
- **Applications to the Energy systems area**:
  - Diffusion models (electric space heaters/air conditioners)
  - Jump Markov models (electric water heaters)
- **Conclusions and future directions**
I. Mean Field Games
Control engineering literature: 2003


Management science literature: 2005


Mathematical literature: 2006, 2007


Mean Field Games: The Modeling Setup

- A **large number** of interacting dynamic agents with individual cost/utility functions.

- The agents are not necessarily homogeneous but share a certain **degree of similarity** in dynamics and cost functions.

- Agents interactions are **anonymous, indifferent** to their particular **ordering**, and for a given agent are dictated by **empirical averages** of functions of only a **pair of states** at a time: the state of the agent and that of another agent in the mass.

- Agents are **rational**, and **share distributional information** on dynamical parameters, cost functions and initial state of the mass of agents. Agents with identical parameters are **interchangeable** in their decision making.
Crowd-following agents:

- **Nonmonopolistic pricing (Economics):** price is global consequence of actions of a large number of suppliers and consumers. Deviations from average can be costly.

- **Collective motion (Biology, Navigation):** In herds, fish schools, pursuit of self interest but need for cohesion. In navigating large groups of tiny robots for exploration purposes, autonomy and coherence must be balanced.

- **Societal dynamics:** the need to balance conformity with originality; analysis of emerging behaviors.

Agoraphobic agents: crowd dynamics in situations of panic evacuation.

Agents interacting through large interconnected systems:

- **Internet:** when congested, equalize bandwidth share of agents.
Motivations: Why Infinite Populations?

- **The case of N non-cooperative dynamic agents:** Finite dynamic games are notoriously difficult because actions of individual agents tend to create responses from all other agents. As $N \to \infty$, influence of single individual vanishes $\Rightarrow$ decoupling and hope for decentralization.

- **The case of centralized control of N interconnected dynamic agents (technological paradigm):** Deliberately structure as game with freedom over choice of utility function to achieve decentralized control as $N \to \infty$.

Intuition about large numbers making things easier already expressed in Von Neuman and Morgenstern (1944)!
Idealized economic behaviors occur with a continuum of players: Aumann (1964), perfect competition.
Why Statistical Mechanics?

- It is a theory that starts from microscopic dynamic descriptions of interacting particles to proceed with the derivation of macroscopic properties.
- It studies limiting behaviors when dealing with a continuum of agents.

What does one learn from Statistical Mechanics?

- Mathematical techniques for the study of the limits of empirical measures as one moves to a continuum of particles.
- The notions of “propagation of chaos” and phase changes.
- That a “generic” microscopic particle process is key to population aggregation.
Agent $i$ dynamics:

$$dx_i = \frac{1}{N} \sum_{j=1}^{N} f_{\theta_i}(x_i, u_i, x_j)dt + \sigma dw_i$$

Agent $i$ cost function:

$$J_i(x^o_i, x^o_{-i}, u_i, u_{-i}) := \mathbb{E}_{x^o_i, x^o_{-i}} \int_0^T \frac{1}{N} \sum_{j=1}^{N} L[x_i, u_i, x_j]dt, \quad T < \infty$$

Assumptions

- Noise processes and initial conditions independent.
- $\theta_i$ parameters indicate potential non homogeneity and have a limiting empirical distribution.
- Initial conditions have a limiting empirical distribution.
Agent $i$ Dynamics:

$$dx_i = f_{\theta_i}(x_i, u_i, \hat{\mu}_t^{(N)})dt + \sigma dw_i$$

Agent $i$ cost function

$$J_i(x^o_i, x^o_{-i}, u_i, u_{-i}) := \mathbb{E}_{|x_i^o, x_{-i}^o} \int_0^T L[x_i, u_i, \hat{\mu}_t^{(N)}]dt, \quad T < \infty$$

Assumptions: Same as before

Specifications: $\hat{\mu}_t^{(N)}$ is the empirical distribution of the agent states at time $t$. Its limit becomes the mean field.
Agent $i$ Dynamics:

\[ dx_i = (A\theta_i x_i + Bu_i + C\bar{x}^{(N)})dt + \sigma dw_i \]

Agent $i$ cost function

\[ J_i(x_i^\circ, x_{-i}^\circ, u_i, u_{-i}) := \mathbb{E}_{x_i^\circ, x_{-i}^\circ} \int_0^\infty e^{-\rho t} \left\{ |x_i - \Phi(\bar{x}^{(N)})|^2_Q + |u_i|^2_R \right\} dt \]

Specifications:

- $\bar{x}^{(N)} := \frac{1}{N} \sum_{j=1}^N x_j$, $\Phi(\bar{x}^{(N)}) := \Gamma \bar{x}^{(N)} + \eta$, $Q \geq 0$, $R > 0$

and $z|_Q^2 := z^\top Q z$.

- Cost functions could be modified to include cross terms
- $\theta_i$ covers non homogeneous dynamics case
- $\Phi(\cdot)$ could be an arbitrary nonlinear function
Asymmetric individual-mass influences: Because of the $1/N$ factor, as $N$ goes to infinity, an isolated individual action will not modify mass behavior. However, individuals all share the influence of the mass through the mass mean trajectory $m(t)$.

The Nash certainty equivalence (NCE) principle as a guessing device (HCM 2007): Treat $m(t)$ as a deterministic given trajectory, and look for optimal response of the individual = individual must solve an optimal tracking problem.

A necessary fixed point property: if $m(t)$ should exist as a deterministic trajectory characteristic of a Nash equilibrium in an infinite population, then it is not sustainable unless it is collectively replicated as the time average of all agent states when optimally responding to $\Phi(m(\cdot))$ (when $\Phi(\cdot)$ is affine).
Scalar LQG case: An Illustration of the NCE Equations

Computation of optimal control law for agent with parameter $a$ and mass trajectory $m(t)$

$$u_a(t) = -\frac{b}{r} (\Pi_a x_a(t) + s_a(t)) \text{ with } \rho \Pi_a = 2a \Pi_a - \frac{b^2}{r} \Pi_a^2 + 1 \quad \text{(Riccati)}$$

$$\rho s_a(t) = \frac{ds_a(t)}{dt} + a s_a(t) - \frac{b^2}{r} \Pi_a s_a(t) - \Phi(m(t)) \quad \text{(Tracking equation)}$$

Collective replication requirement for $m(t)$ under the optimal feedback strategy

$$dx_a(t) = \left( a - \frac{b^2}{r} \Pi_a \right) x_a(t) - \frac{b^2}{r} s_a(t) \ dt + \sigma dw(t) \quad \text{(the generic process!)}$$

$$\frac{d\bar{x}_a(t)}{dt} = \left( a - \frac{b^2}{r} \Pi_a \right) \bar{x}_a(t) - \frac{b^2}{r} s_a(t) \quad \text{(the mean state conditional on $a$)}$$

$$m(t) = \int_{\mathcal{A}} \bar{x}_a(t) dF(a) \quad \text{(the unconditional mean state)}$$
Asymmetric individual-mass influences: Because of the $1/N$ factor, as $N$ goes to infinity, an isolated individual action will not modify mass behavior. However, individuals now share the influence of the mass through the evolution of not only the mean mass trajectory, but that of the “entire” population limiting empirical distribution $\mu_t(x)$, $t \geq 0$, i.e., a flow of measures!

Let $m_t(x)$, $t \geq 0$, be the family of Radon-Nykodim derivatives (probability density functions) associated with these measures.
The Nash certainty equivalence (NCE) principle as a guessing device: Treat $m_t, t \geq 0$, as a “deterministic” given flow of pdfs and look for optimal response of the individual = Individual must solve a “backward propagating $m_t$ dependent HJB equation” to produce a local state feedback strategy which depends on the assumed entire flow of pdfs.

A necessary fixed point property: If $m_t, t \geq 0$, should exist as a deterministic flow of pdfs characteristic of a Nash equilibrium in an infinite population, then it is not sustainable unless it is collectively replicated as the Radon-Nykodim derivative of the limiting flow of empirical measures of the infinite population when optimally responding in the context of the assumed flow of pdfs $m_t, t \geq 0$. Require that $m_t$ satisfy the “forward Kolmogorov (or Fokker-Planck) equations” associated with the $m_t, t \geq 0$, dependent HJB based closed loop optimally controlled generic process.
Computation of optimal control law for agent (under homogeneous dynamics) with assumed mass pdf flow \( m_t, t \geq 0 \)

\[
- \frac{\partial V}{\partial t} = \inf_{u \in U} \left[ f(x, u, m_t) \frac{\partial V}{\partial x} + L(x, u, m_t) \right] + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} \quad \text{(HJB equation)}
\]

\( V(x, T) = 0 \) \quad \text{(Backwards propagation)} \quad (t, x) \in [0, T] \times \mathbb{R}

\( u^*_m(t, x) := \psi(t, x, m_\tau : \tau \in [t, T]) \) \quad \text{(Candidate optimal control)}

Collective replication requirement for \( m_t, t \geq 0 \), under the optimal feedback strategy

\[
dx_t = f(x_t, u^*_m(t, x), m_t)dt + \sigma dw_t \quad \text{(The generic closed-loop process)}
\]

\[
\frac{\partial m_t}{\partial t} = - \frac{\partial [f(x_t, u^*_m(t, x), m_t)m_t]}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 m_t}{\partial x^2} \quad \text{(Fokker-Planck equation)}
\]

\( m_t(x)|_{t=0} = m_0(x) \) \quad \text{(Forward propagation!)}
Existence and Uniqueness: The HCM’07 Approach for LQG Case

MFG (NCE) LQG Equations

\[
\rho s_a(t) = \frac{ds_a(t)}{dt} + as_a(t) - \frac{b^2}{r} \Pi_a s_a(t) - \Phi(m(t)) \quad \text{(Tracking equation)}
\]

\[
\frac{d\bar{x}_a(t)}{dt} = \left(a - \frac{b^2}{r} \Pi_a\right) \bar{x}_a(t) - \frac{b^2}{r} s_a(t) \quad \text{(the mean state conditional on } a)\]

\[
m(t) = \int_A \bar{x}_a(t) dF(a) \quad \text{(the unconditional mean state)}
\]

**Idea:** Look at the map that goes from \( m(t) \) in the tracking differential equation, back to \( m(t) \) in the unconditional mean state equation, and develop sufficient conditions for this map to be a contraction operator from the space of bounded continuous functions of time on itself. Guarantee at the same time existence and uniqueness (Theorems 4.3, 4.4 in Huang, Caines and Malhamé IEEE TAC, 2007)

See also ergodic analysis by Li-Zhang (2008), M. Bardi (2012).
Existence and Uniqueness: The General Case (HMC’06)

MFG finite horizon equations

\[-\frac{\partial V}{\partial t} = \inf_{u \in U} \left[ f(x, u, m_t) \frac{\partial V}{\partial x} + L(x, u, m_t) \right] + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} \]  \hspace{1cm} (HJB equation)

\[\frac{\partial m_t}{\partial t} = -\frac{\partial}{\partial x} \left[ f(x_t, u^*_m(t, x), m_t) \right] m_t + \frac{\sigma^2}{2} \frac{\partial^2 m_t}{\partial x^2} \]  \hspace{1cm} (Fokker-Planck equation)

\[V(x, T) = 0 \quad m_t(x)|_{t=0} = m_0(x)\]

Idea (Huang, Mallamé, Caines, CIS 2006)

- Starting with an arbitrary flow of measures with weak but sufficient “continuity” properties, and under assumptions of strict convexity and adequate smoothness of Hamiltonian in HJB equation, show that control law exists uniquely and is Lipschitz continuous.

- Using McKean-Vlasov equations theory and standard assumptions on system dynamics, show existence of unique flow of measures for FKP equation for which the assumed continuity properties persist.

- Consider combined HJB-FKP operator from measure flows, back to measure flows. Under sufficient conditions it is a contraction and existence/uniqueness are guaranteed.
Existence and Uniqueness: The General Case (LL’06)


- **Existence**: Pretty much the same two steps analysis of starting from a suitably continuous arbitrary flow of measures but also living on a convex set feeding into the HJB equation. Existence and uniqueness of optimal control solution of the HJB equation is based on transformation to heat equation for dynamics affine in control and cost quadratic in control.

Properties of Itô differential equations are used to conclude that the corresponding FPK equation has a unique solution. The combined operator is shown to be continuous. Schauder’s fixed point theorem is used for existence. Can handle more general Hamiltonians.

- **Uniqueness**: Separable Hamiltonian and agoraphobic agents

\[
H \left( x, m_t, \frac{\partial V}{\partial x} \right) := S \left( x, \frac{\partial V}{\partial x} \right) + T(x, m_t)
\]

\[
\int_{\mathbb{R}} (T(x, m_1) - T(x, m_2)) d(m_1 - m_2)(x) > 0, \forall m_1, m_2 \in P_1, m_1 \neq m_2
\]
What It Means for Games with $N$ Large but Finite

- **An approach directly inspired by the methods of statistical mechanics (LL’06,’07):**
  - Start studying Nash equilibria for finite agent systems.
  - Obtain MFG equations as characteristic of limiting behavior of finite agent systems as $N \to \infty$ and study existence and uniqueness properties.
  - Under “sufficient regularity” of the relevant functions, and uniqueness of solutions to MFG equations, show that difference of performance relative to Nash equilibrium of infinite population can be made arbitrarily small by increasing $N$ sufficiently.

- **A more direct approach (HMC’06, HCM’07):**
  - Through intuitive analysis, obtain form of MFG equations and carry out corresponding existence and uniqueness analysis.
  - Characterize the family of decentralized Nash equilibrium inducing control laws for infinite population.
  - Apply infinite population based control policies to all agents but one in a finite game, and study the limits on its cost improvement.
Discussion and Nash Equilibrium Approximation Concepts

- **Statistical mechanics versus large population games:**
  In statistical mechanics, no control on dynamics $\Rightarrow$ Only an approximation problem $\Rightarrow$ Essential to move from finite to infinite. In large population games, one looks for Nash equilibrium inducing feedback control strategies $\Rightarrow$ Given any set of feedback strategies, possible to study directly distance from Nash equilibrium.

- **Nash equilibrium approximation concepts:** $\epsilon$-Nash equilibrium

  $$J_i(u_i^*, u_{-i}^*) - \epsilon \leq \inf_{u_i} J_i(u_i, u_{-i}^*) \leq J_i(u_i^*, u_{-i}^*)$$

  where $\epsilon \rightarrow 0$ as $N \rightarrow \infty$, with

  - **Full information:** $u_i$ depends on $(t, x_1, x_2, \ldots, x_N)$
  - **Local information:** $u_i$ depends on $(t, x_i)$. 


Results for MFG Induced Decentralized Control Laws

- **Full information \( \epsilon\)-Nash property holds for:**
  - **Infinite horizon exponentially discounted LQG case** with linear or nonlinear coupling in the cost and dynamics (HMC 2006, HCM 2007)
  - **Finite horizon nonlinear case** with dynamic interactions of agents with mean field of the form \([f(x,u)g(m) + h(m)]\), and arbitrary individual running cost \(L(x,u,m)\) (HMC 2006)
  - **Ergodic MFG**, i.e. considering time average cost, motion occurring on a toroidal structure, and all functions being periodic, and under sufficient conditions for uniqueness (LL 2006, Cardaliaguet MFG notes 2012)

- **Local information \( \epsilon\)-Nash property holds for:**
  - **Finite horizon general nonlinear case** (HMC 2006, LL 2006, Cardaliaguet MFG notes 2012)
Individual dynamics:

\[ dx_i = (Ax_i + Bu_i)dt + \sigma dw_i, \quad 1 \leq i \leq N. \]

Cost for the \( i \)th leader:

\[ J_i(x_i^o, x_{-i}^o, u_i, u_{-i}) = E \int_0^{\infty} e^{-\rho t} \left\{ \left| x_i - \left( \lambda h + (1 - \lambda) \frac{1}{N} \sum_{k=1}^{N} x_k \right) \right|^2 + |u_i|^2_R \right\} dt \]

Cost for the \( i \)th follower:

\[ J_i(x_i^o, x_{-i}^o, u_i, u_{-i}) = E \int_0^{\infty} e^{-\rho t} \left\{ \left| x_i - \frac{1}{N_1} \sum_{k \in \mathcal{L}} x_k \right|^2 + |u_i|^2_R \right\} dt \]

\( \mathcal{L} \): The set of leaders with cardinal \( N_1 \)
\( h(\cdot) \) is intended for a certain reference trajectory assumed to be known to both leaders and followers (the case of reference trajectory unknown to the followers: NCMH, IEEE TAC 2012).

Leader-follower architecture is of interest in animal behaviour analysis and also in group motion of multiple autonomous robots.

Direct tracking of \( h \) may result in poor cohesiveness during transient phase due to different initial conditions.

In many problems of interest (e.g., swarming, flocking of animals or schooling of fish), group cohesiveness is important at all stages (Couzin et. al. Nature, Feb. 2005).
Assumption: There exists $\alpha \in (0, 1)$ such that $\lim_{N \to \infty} \frac{N_1}{N} = \alpha$, i.e., the proportion of leaders is fixed.

Let the mean field centroid trajectory for leaders and followers be $\bar{x}_L(\cdot)$ and $\bar{x}_F(\cdot)$

MFG (NCE) LQG equations for the leaders

\[
\rho s_L(t) = \frac{ds_L(t)}{dt} + A^\top s_L(t) - \Pi B R^{-1} B^\top s_L(t) - x^*_L(t)
\]

\[
\frac{d\bar{x}_L(t)}{dt} = (A - B R^{-1} B^\top \Pi)\bar{x}_L(t) - B R^{-1} B^\top s_L(t)
\]

\[
x^*_L(t) = \lambda h(t) + (1 - \lambda) (\alpha \bar{x}_L(t) + (1 - \alpha)\bar{x}_F(t))
\]

- Individual control action for the leaders

\[
u_i^o = -R^{-1} B^\top (\Pi x_i + s_L)
\]
MFG and a Model of Fish Schooling Dynamics

MFG (NCE) LQG equations for the followers

\[ \rho s_F(t) = \frac{ds_F(t)}{dt} + A^\top s_F(t) - \Pi BR^{-1}B^\top s_F(t) - \bar{x}_L(t) \]

\[ \frac{d\bar{x}_F(t)}{dt} = (A - BR^{-1}B^\top \Pi)\bar{x}_F(t) - BR^{-1}B^\top s_F(t) \]

- Individual control action for the followers

\[ u_o^i = -R^{-1}B^\top (\Pi x_i + s_F) \]
Numerical Example

- 200 leaders and 300 followers
- System matrices $A = \begin{bmatrix} 0.2 & -0.3 \\ -0.4 & 0.2 \end{bmatrix}$, $B = I$, $\sigma = 2I$, $R = 0.5I$
- Reference trajectory $h(t) = \begin{bmatrix} 10 + 15 \sin(2t) \\ 10 + 20 \cos(2t) \end{bmatrix}$, $0 \leq t \leq 10$
- Gaussian initial distribution $N(0, 10I)$
Numerical Example

$\rho = 0.2$ and $\lambda = 0.5$

$\rho = 2$ and $\lambda = 0.5$
Numerical Example

\[ \rho = 0.1 \text{ and } \lambda = 0.1 \]

\[ \rho = 0.1 \text{ and } \lambda = 0.1 \]

\[ \rho = 0.1 \text{ and } \lambda = 0.9 \]

\[ \rho = 0.1 \text{ and } \lambda = 0.9 \]
Numerical Example

State trajectories of the leaders and the followers.

Animation of Trajectories
Towards mean field LQG cooperative games: Decentralized socially optimal control laws (HCM, IEEE TAC July 2012)

Agent $i$ dynamics:

$$dx_i = \left(A_{\theta_i} x_i + B u_i + C \bar{x}^{(N)}(N)\right) dt + \sigma dw_i$$

Social cost to be optimized:

$$J(x^\circ, u) := \sum_{i=1}^{N} J_i(x_i^\circ, x_{-i}^\circ, u_i, u_{-i})$$

$$= \sum_{i=1}^{N} E_{|x_i^\circ, x_{-i}^\circ} \int_0^\infty e^{-\rho t} \left\{ \left| x_i - \Phi \left( \bar{x}^{(N)} \right) \right|^2_Q + |u_i|^2_R \right\} dt$$

- As $N \to \infty$, small individual agent actions are reflected in the individual costs of all other agents (one way coupling of individual and mass property of non-cooperative MFGs is lost!).

- Consider person by person by person optimization as in team theory together with MFG collective replication requirement.

- Individual asymptotic optimal control laws are still decentralized and are full information $\epsilon$-optimal!
Further Research II


- **Towards oligopolies within MFG:**
  


  **Nonlinear case:** N. Şen and P.E. Caines, “Mean Field Games with Partially Observed Major Player and Stochastic Mean Field”, IEEE CDC, 2014.
Nonlinear MFG models:

Applications in Smart Grids:
Numerical Analysis of MFG Equations


Book


Survey


Encyclopedia Article


Introductory Blogs

T. Tao, “Mean Field Games”, http://terrytao.wordpress.com/2010/01/07/mean-field-equations/
L.N. Hoang, “The New Big Fish Called Mean Field Game Theory”, http://www.science4all.org/le-nguyen-hoang/mean-field-games/
MFG’s are a convenient approach for crystallizing the interaction of an individual agent with a large population of similar agents which individually carry a vanishing decision weight.

In general, they lead to $\epsilon$-Nash or $\epsilon$-optimal decentralized feedback control strategies with the degree of error typically decaying as $1/\sqrt{N}$.

Decentralization is possible only if the necessary a priori data is available i.e. typically, initial agent states empirical distributions as well as dynamic parameters empirical distributions (identification and on line learning issues (LQG case: A.C. Kizilkale, P.E. Caines, IEEE TAC 2013)).
Summary (cnt)

- MFGs exploit to the fullest extent the predictability stemming from the law of large numbers under conditions of propagation of chaos. They do so by constructing a generic stochastic agent process.

- With LQG-MFGs you almost get a free lunch because the mass is aggregated into a single agent. Many applications are expected; already appearing in the smart grids area!

- Numerical methods are bound to become a very important issue for the viability of nonlinear MFG analysis in general.
II. Applications to Energy Systems
A higher level of penetration of renewable sources of energy in the energy mix of power systems (wind or solar) is synonymous with greater variability.

Energy storage becomes an essential asset to compensate generation/load mismatch.

Fundamental idea: Use the energy storage from electrical sources naturally present in the power system at customer sites based on mutually beneficial agreements (electric water heaters, electric space heating, electric space cooling).
Challenges and Previous Work

- **Challenges:** Literally millions of control points to model, monitor and control; severe computational requirements; large communication costs.

- **Past approaches:** Direct load control. Send the same interruption/reconnection signals to large collections of devices.

Aggregate modeling scheme:

- Develop elemental stochastic load model of individual load behavior.
- Build aggregate load model by developing ensemble statistics of the devices, much as in the statistical mechanics framework.
An Example: A Diffusion Model of Heating/Cooling Loads

A hybrid state stochastic system

**(Continuous State):**

\[ C_a \frac{dx_{in}}{dt} = -U_a (x_{in} - x_{out}) dt + Q_h m_t b_t dt + \sigma dw_t \]

divide by \( C_a \) and obtain

\[ dx_{in} = a (x_{in} - x_{out}) dt + Q' h m_t b_t dt + \sigma' dw_t \]

**(Discrete State):**

\[ m_{t+\Delta t} = m_t + \pi(x_{in}; x_+, x_-) \]

\[ \pi(x_{in}, m; x_+, x_-) = \begin{cases} 
0 & x_- < x_{in} < x_+ \\
-m & x_{in} \geq x_+ \\
1 - m & x_{in} \leq x_- 
\end{cases} \]

- \( m \) the operating state of the device (1 for “on” or 0 for “off”)
- \( b \) the state of the power supply (1 for “on” or 0 for “off”).
The Coupled Fokker-Planck Equations

The resulting coupled Fokker-Planck equation model describing the evolution of temperature distributions within controlled residences

\[ T_{\lambda,t}^k[f] = \frac{\partial f}{\partial t} - \frac{\partial}{\partial \lambda} \left[ (a(\lambda - x_a(t)) - k b(t) R) f \right] - \frac{\sigma^2}{2} \frac{\partial^2}{\partial \lambda^2} f, \quad k = 0, 1 \]
The optimal control problem becomes one of controlling PDEs using on-off signals.

A fraction of customers is inevitably penalized.

The smaller this fraction, the less effective the control is.
Ongoing improvements

- PDE based randomized controls locally implemented (Totu, Winieski 2014)
- State estimation and control of electric loads to manage real-time energy imbalance (Mathieu, Koch, Callaway 2013)
- Mean field related ideas (Meyn, Barooah, Bušić, Ehren 2013)
- Mean field game based methods (Kizilkale, Malhamé 2013, Ma, Callaway and Hiskens 2013)
Implementation Principles

1. Each controller has to be situated locally: decentralized

2. Data exchange should be kept at minimum both with the central authority and among users

3. User disturbance should be kept at minimum
Envisioned Overall Architecture: The case of a single central authority

- Forecasts of wind and solar generation
- Aggregate models of energy storage capable devices
- Mathematical Programming
- Energy/temperature schedule of large aggregates of devices
  - electrical space heaters
  - electrical water heaters
  - etc.
- Uncontrollable component of load

Aggregate models of energy storage capable devices

Forecasts of wind and solar generation

Mathematical Programming

Energy/temperature schedule of large aggregates of devices
  - electrical space heaters
  - electrical water heaters
  - etc.

Uncontrollable component of load
Mean Field Games: Why?

Linear Quadratic Mean Field Rendez-vous Problem

\[ J_i(u^i, u^{-i}) = \mathbb{E} \int_0^\infty e^{-\delta t} \left\{ \left[ x_t^i - \gamma (\bar{x}^N + \eta) \right]^2 q + (u_t^i)^2 r \right\} dt, \quad 1 \leq i \leq N \]
Mean Field Games: The Reasons?

Two fundamental reasons:

- Games are a natural device for enforcing decentralization.
- The large numbers involved induce decoupling effects which allow the law of large numbers to kick in.

Practical benefits:

- The resulting control laws can be computed in an open loop manner by individual devices thus significantly reducing communication requirements.
- Control implementation is local unlike direct control, thus permitting local enforcement of comfort and safety constraints.
II.I. Non-cooperative Collective Target Tracking Mean Field Control for Space Heaters
Elemental space heating model: two changes

1. Thermostat controlled heating element is replaced by a continuous controller.

\[
dx_t^\text{in} = \frac{1}{C_a}[-U_a(x_t^\text{in} - x_t^\text{out}) + Q_h(t)]dt + \sigma dw_t
\]

\[\equiv\]

\[
dx_t = [-a(x_t - x_t^\text{out}) + bu_t]dt + \sigma dw_t
\]

2. The control input is redefined so that no control effort is required on average to remain at initial temperature.

\[
dx_t = [-a(x_t - x_t^\text{out}) + b(u_t + u^\text{free})]dt + \sigma dw_t
\]

where \[u^\text{free} \triangleq a(x_t^\text{out} - x_0^\text{i}).\]

- Change 2 is made so that diversity is preserved in the water heater population while mean population temperature tracks the target.

- With no control effort, the temperature stays at \(x_0\). We do not penalize the control effort that is used to stay at the initial temperature at the start of the control law horizon.
Constant Level Tracking Problem Setup

- **Initial Mean Temperature**
- **Target Mean Trajectory** = $y$
- **High Comfort Line** = $h$
- **Low Comfort Line** = $z$

![Graph showing the Constant Level Tracking Problem Setup](image-url)
Classical LQG Target Tracking

\[ J_i^N(u^i, u^{-i}) = \mathbb{E} \int_0^{\infty} e^{-\delta t} \left[ (x^i_t - y)^2 q + b(u^i_t)^2 r \right] dt, \quad 1 \leq i \leq N \]

\( x^i \): temperature

\( y \): tracking target

\( u^i \): control

Agents Applying LQG Tracking
A Novel Integral Control Based Cost Function

\[ J_{i}^{N}(u^{i}, u^{-i}) = \mathbb{E} \int_{0}^{\infty} e^{-\delta t} \left[ (x_t^{i} - z)^2 q_{t}^{y} + (x_t^{i} - x_0^{i})^2 q_{0}^{x} + b(u_t^{i})^2 r \right] dt \]

- \( x^{i} \): temperature
- \( z \): lower comfort bound
- \( u^{i} \): control

- Integral controller embedded in mean-target deviation coefficient:
  \( q_{t}^{y}, t \in [0, T] \), calculated as the following integrated error signal:

\[ q_{t}^{y} = \left| \lambda \int_{0}^{t} (\bar{x}_{\tau}^{N} - y) d\tau \right| \]

- \( \bar{x}^{N} \): mean temperature of the population
- \( y \): mean target

\[ q_{0}^{x} \]
Mean Field Based Collective Target Tracking

- Novelty is that the mean field effect is mediated by the quadratic cost function parameters under the form of an integral error as compared to prevailing mean field theory where the mean field effect is on the tracking signal.
Define the function space $G \subset C_b[0, \infty)$, where for function $f \in G$, $f(0) = \bar{x}_0$ and $z \leq f(t) \leq \bar{x}_0$ for all $t \geq 0$.

$\theta \in \Theta$ corresponds to all sources of heterogeneity including initial conditions.

**Theorem**: Schauder’s fixed point theorem guarantees the existence of a fixed point to the MF Equation System on space $G$. 
Successful experiment with a naive iterative algorithm. Note that lower comfort bound is 17, whereas the mean tracking target is 21.
Unsuccessful experiment with a naive iterative algorithm (different parameters). Note that lower comfort bound is 17, whereas the mean tracking target is 21.
From the experimentation, one concludes that the cases of interest correspond to convergence within the region \([x_0, y]\); also one observes monotonicity of the temperature trajectories in the early part of their behavior (pre crossing target \(y\)). This suggest to us the idea of using a so-called restricted operator whereby anytime the mean state hits the target \(y\), it is frozen at \(y\).
Operator Definitions

- Define the function space $G^r \subset G$, where for functions $f \in G^r$, $f(0) = \bar{x}_0$ and $y \leq f(t) \leq \bar{x}_0$ for all $t \geq 0$.

- Define the operator $\Delta(\bar{x}; \lambda) : C_b[0, \infty) \to C[0, \infty)$:

  \[
  q_t^y = \Delta(\bar{x}_t; \lambda) \overset{\triangle}{=} \lambda \int_0^t (\bar{x}_\tau - y) d\tau
  \]

- Define $T : C[0, \infty) \to C[0, \infty)$:

  \[
  \bar{x}_t = (T q)(t) \overset{\triangle}{=} \begin{cases} 
  -\frac{d\pi_t}{dt} &= (-2a - \delta)\pi_t - b^2 r^{-1} \pi_t^2 + q_t^y + q_{x_0}^y, \\
  -\frac{ds_t}{dt} &= (-a - \delta - b^2 \pi_t r^{-1})s_t + a\bar{x}_0 \pi_t - q_t^y z - q_{x_0}^y \bar{x}_0, \\
  \frac{d\bar{x}_t}{dt} &= (-a - b^2 \pi_t r^{-1})\bar{x}_t - b^2 r^{-1} s_t + a\bar{x}_0
  \end{cases}
  \]

- Hence, write the MF Eq. System for Collective Target Tracking as

  \[
  \bar{x}_t = (T \circ \Delta)(\bar{x})(t)
  = (\mathcal{M}\bar{x})(t).
  \]
Restricted Operator

Define $T_h \in (\mathbb{R}^+ \cup \infty)$ as the first time that $\bar{x}$ hits $y$.

**Restricted Operator**

Define $\mathcal{T}^r : \mathbb{C}[0, \infty) \to \mathbb{C}_b[0, \infty)$:

$$\mathcal{T}^r \triangleq \begin{cases} \mathcal{T}(q_t^y) & \text{for } t \in [0, T_h) \\ y & \text{for } t \in [T_h, \infty) \end{cases}$$

Define $\mathcal{M}^r \triangleq (\mathcal{T}^r \circ \Delta) : \mathbb{C}_b[0, \infty) \to \mathbb{C}_b[0, \infty)$.
A Numerical Algorithm for the Restricted Operator

- \( k = 0 \)
- while \( \|\bar{x} - \bar{x}^{old}\|_{\infty} > \epsilon_1 \), do
  - \( \bar{x}^{old}(2) = \bar{x}^{old} \)
  - \( x^{old} = \bar{x} \)
  - if \( \text{mod}(k, 2) == 1 \) then
    - \( q^y = \Delta(\bar{x}; \lambda) \)
    - if \( q^y_{\infty} > (q^y_{\infty})^* \) then
      - \( \lambda = \lambda \times \frac{1}{1 + \epsilon_2} \)
      - \( q^y = \Delta(\bar{x}; \lambda) \)
    - end if
  - elseif \( |\bar{x} - \bar{x}^{old}(2)| == 0 \) then
    - \( q^y = \Delta^r(\bar{x}; \lambda) \)
  - else
    - \( q^y = \Delta(\bar{x}; \lambda) \)
  - end if
  - \( \bar{x} = T^r(q^y) \)
  - \( k = k + 1 \)
- end while
- return \( \bar{x} \).

Define \((q^y_{\infty})^*\)

\[ \Delta(\bar{x}) \text{ of } \bar{x} \text{ for } \bar{x}_t \rightarrow y \text{ as } t \rightarrow \infty. \]

\[ (q^y_{\infty})^* = \frac{a(a + \delta)r + b^2q^{x_0}}{b^2} \left( \frac{\bar{x}_0 - y}{y - z} \right) \]
Coefficients can be tuned to guarantee successful convergence at the cost of slowing collective dynamics.

The algorithm calculates a $\lambda^*$ such that there exists a desirable fixed point trajectory for the operator $M = T \circ \Delta(x; \lambda) : G \to G$.

The numerical algorithm always converges to a smooth fixed point trajectory.
\(\epsilon\)-Nash Theorem

**Theorem** Under technical conditions the collective target tracking MF stochastic control law generates a set of controls 
\(U_{col}^N \triangleq \{(u^i)^\circ; 1 \leq i \leq N\}, 1 \leq N < \infty\), with 
\[ (u^i_t)^\circ = -b/r(\pi_t^ix_t^i + s_t^i), \quad t \geq 0, \]
such that

(i) all agent system trajectories \(x^i, 1 \leq i \leq N\), are stable in the sense that 
\[ \mathbb{E} \int_0^\infty e^{-\delta t} \|x^i_t\|^2 dt < \infty; \]

(ii) \(\{U_{col}^N; 1 \leq N < \infty\}\) yields an \(\epsilon\)-Nash equilibrium for all \(\epsilon > 0\); there exists \(N(\epsilon)\) such that for all \(N \geq N(\epsilon)\)

\[ J_i^N ((u^i)^\circ, (u^{-i})^\circ) - \epsilon \leq \inf_{u^i \in U^N_g} J_i^N (u^i, (u^{-i})^\circ) \leq J_i^N ((u^i)^\circ, (u^{-i})^\circ). \]
Simulations

- 200 heaters with maximum power: 10kW
- 2 experiments:
  - increase 0.5 °C mean temperature,
  - decrease 0.5 °C mean temperature,

over a 10 hours control horizon

- case 1: the central authority provides target, local controllers apply LQG tracking
- case 2: the central authority provides the target, local controllers apply collective target tracking mean field
Energy Release: LQG Tracking

Agents Applying LQG Tracking
Energy Release: Collective Target Tracking MF Control

Agents Applying Collective Target Tracking MF Control
Energy Release: LQG Tracking vs Collective Target Tracking MFC
Accelerated Engineering Solution: No Control until agent’s temperature reaches its individual steady state:

\[(x^i_\infty)^* = x^i_0 - \frac{b^2(q^y_\infty)^*(x^i_0 - z)}{ar(a + \delta) + (q^y_\infty)^*b^2 + q^{x_0}b^2}.\]
Aggregate Power Relief Curve

Aggregate Power Consumption
Energy Accumulation: Homogeneous vs Heterogeneous Populations

**Accelerated Engineering Solution:** Max Control until agent's temperature reaches its individual steady state:

\[
(x^i_\infty)^* = x^i_0 - \frac{b^2(q^y_\infty)^*(x^i_0 - h)}{ar(a + \delta) + (q^y_\infty)^*b^2 + qx_0b^2}.
\]

Agents Applying Collective Target Tracking MF Control

- First group: higher initial point, second group: lower initial point
- Second group's control penalty coefficient \( r \) is lower than the second group
Energy Accumulation: Homogeneous vs Heterogeneous Populations

First group: higher initial temperature, second group: lower initial temperature
experiment 1: same control penalty coefficient $r$ for both groups
experiment 2: second group’s control penalty coefficient $r$ is lower than the first group
II.II. Non-cooperative Collective Target Tracking Mean Field Control for Water Heaters
Water Heater Stratification Model

\[ M_l C_{pf} \frac{dx_l, t}{dt} = U A_l (x_{env} - x_{l,t}) + \dot{m}_L C_{pf} (x_{(l+1), t} - x_{l,t}) + \dot{Q}_l u_{l,t}, \]

\( t \geq 0, \quad l \neq n \)

\[ M_l C_{pf} \frac{dx_l, t}{dt} = U A_l (x_{env} - x_{l,t}) + \dot{m}_L C_{pf} (x_{L} - x_{l,t}) + \dot{Q}_l u_{l,t}, \]

\( t \geq 0, \quad l = n \)

Elemental Agent Dynamics

- $\dot{m}_L$: modeled as a (stochastic) jump process.

- Physical model:

$$
M_l C_{pf} \frac{dx_{l,t}}{dt} = UA_l (x_{env} - x_{l,t}) + \dot{m}_L C_{pf} (x_{(l+1),t} - x_{l,t}) + \dot{Q}_l u_{l,t},
$$

$$
t \geq 0, \quad l \neq n
$$

$$
M_l C_{pf} \frac{dx_{l,t}}{dt} = UA_l (x_{env} - x_{l,t}) + \dot{m}_L C_{pf} (x_{L} - x_{l,t}) + \dot{Q}_l u_{l,t},
$$

$$
t \geq 0, \quad l = n
$$

- We write it as:

$$
\frac{dx_t}{dt} = A^{\theta_t} x_t + B u_t + c^{\theta_t}, \quad t \geq 0.
$$

- $\theta_t$, $t \geq 0$, is a continuous time Markov chain taking values in $\Theta = \{1, 2, \ldots, p\}$ with infinitesimal generator matrix $\Lambda$. Each discrete value is associated with a type of event (showers, dishwashers, etc ... )
A Single Agent’s Temperature Trajectory Applying Markovian Jump LQ Tracking
Constant Level Tracking Problem Setup

- High comfort line = h
- Initial mean temperature
- Target mean trajectory = y
- Population direction = z
- Low comfort line = l

Graph showing temperature over time with different comfort levels and population direction.
Redefined Dynamics

Dynamics for a population of $N$ water heaters:

$$\frac{dx^i_t}{dt} = A^i \theta^i_t x^i_t + Bu^i_t + c^i \theta^i_t, \quad t \geq 0, \quad 1 \leq i \leq N$$

The control input is redefined so that no control effort is required on average to remain at initial temperature.

$$\frac{dx^i_t}{dt} = A^i \theta^i_t x^i_t + B(u^i_t + u^i_{t, free}) + c^i \theta^i_t, \quad t \geq 0, \quad 1 \leq i \leq N,$$

where

$$u^i_{t, free} = \sum_{l=1}^{n} UA_l(x^i_{t,0} - x^{env}) + \mathbb{E} \sum_{j=1}^{n} \zeta^j(t) \mu^L_t(j) C^{pL}(x_{1,t} - x^L_t)$$

$$\zeta(t) = [\zeta^1(t), ..., \zeta^p(t)]$$ is the probability distribution of the Markov chain.
Integral Control Based Cost Function

Cost functions:

\[ J^N_i(u^i, u^{-i}) = \mathbb{E} \int_0^T \left[ (Hx_t^i - z)^2 q^y_t + (Hx_t^i - Hx^i_0)^2 q^x_0 + \|u^i_t\|_R^2 \right] dt \]

\[ + (Hx^i_T - z)^2 q^y_T + (Hx^i_T - Hx^i_0)^2 q^x_0 \]

\( x^i \) temperature
\( z \) lower comfort bound
\( u^i \) control
\( H \) \([1/n, \ldots, 1/n]\)

Integral controller embedded in mean-target deviation coefficient:
\( q^y_t, \ t \in [0, T] \), calculated as the following integrated error signal:

\[ q^y_t = \left| \lambda \int_0^t (H\overline{x}^N_t - y) dt \right| \]

\( \overline{x}^N \) mean temperature of the population
\( y \) mean target
For a given \( \bar{x}_t \) and thus \( q^y_t \), \( t \in [0, T] \), compute optimal agent response: [W. M. Wonham, 1971]

- each agent \( A_i, 1 \leq i \leq N \), obtains the positive solution to the coupled set of Riccati equations

\[
- \frac{d\Pi^j_t}{dt} = \Pi^j_t A^j + A^j \top \Pi^j_t
\]

\[
- \Pi^j_t B R^{-1} B \top \Pi^j_t + \sum_{k=1}^{p} \lambda_{jk}(t) \Pi^k_t + (q^y_t + q^{x_0}) H \top H,
\]

where

\[
\Pi^j_T = (q^y_T + q^{x_0}) H \top H, \quad 1 \leq j \leq p
\]
for a given target signal $z$, the individual $i$th agent offset function is generated by the coupled differential equations

$$-\frac{ds_{i,t}^j}{dt} = (A^j - B R^{-1} B^\top \Pi_t^j)^\top s_{i,t}^j - q_t^y H^\top z - q^{x_0} H^\top x_0^i + \Pi_t^j c_i^j$$

$$+ \sum_{k=1}^p \lambda_{jk}(t)s_{i,t}^k,$$

where

$$s_{i,T}^j = -[q_T^y H^\top z + q^{x_0} H^\top x_0^i], \quad 1 \leq j \leq p$$

the optimal tracking control law is given by

$$u_{i,t}^o = -\sum_{j=1}^p I[\theta_{i,t} = j] R^{-1} B^\top (\Pi_t^j x_{i,t} + s_{i,t}^j), \quad t \geq 0.$$
Fixed Point Equation System: Step 2

Under best response to posited \( \bar{x}_t \), agents mean must replicate \( \bar{x}_t \):

\[
q_i^y = \left| \lambda \int_0^t (H\bar{x}_\tau - y) d\tau \right|
\]

\[
-\frac{d\Pi_t^j}{dt} = \Pi_t^j A^j + A^j \Pi_t^j - \Pi_t^j BR^{-1} B^\top \Pi_t^j
\]

\[
+ \sum_{k=1}^p \lambda_{jk} \Pi_t^k + (q_i^y + q^{x0}) H^\top H, \quad \Pi_T^j = (q_i^y + q^{x0}) H^\top H, \quad 1 \leq j \leq p,
\]

\[
-\frac{ds_t^j}{dt} = (A^j - BR^{-1} B^\top \Pi_t^j)^\top s_T^j - q_i^y H^\top z - q^{x0} H^\top \bar{x}_0
\]

\[
+ \Pi_t^j c^j + \sum_{k=1}^p \lambda_{jk} s_t^k, \quad s_T^j = -[q_i^y H^\top z + q^{x0} H^\top \bar{x}_0], \quad 1 \leq j \leq p,
\]

\[
\frac{d\bar{x}_t^j}{dt} = (A^j - BR^{-1} B^\top \Pi_t^j)\bar{x}_t^j + \sum_{k=1}^p \lambda_{kj} \bar{x}_t^k + \zeta_t^j c^j - \zeta_t^j BR^{-1} B^\top s_t^j, \quad 1 \leq j \leq p
\]

\[
\bar{x}_t = \sum_{j=1}^p \bar{x}_t^j,
\]

One recalls

- \( \bar{x}_t^j = E(\bar{x}_t I_{\theta_t=j}) \)
- \( \Lambda = \{\lambda_{i'j'} | i', j' = 1, \ldots, p\} \) is the infin. gen. of the MC
- \( \zeta_t = [\zeta_t^1, \ldots, \zeta_t^p] \) is the prob. dist. of the MC
Fixed Point Analysis

Define the set $\mathcal{G}$: all functions $f \in C_b[0, T], f(0) = \bar{x}_0$ and $z \leq f(t) \leq \bar{x}_0, t \in [0, T]$.

**Lemma:** Under $\|\cdot\|_\infty$, $\mathcal{G}$ is closed in $C_b[0, T]$; therefore it is a complete metric space.

**Theorem:** The following assumption guarantees the existence of a unique fixed point for the map $\mathcal{M} : \mathcal{G} \to \mathcal{G}$, due to the Banach fixed point theorem.

**Contraction Assumption**

$$\frac{2b_1}{a_1 \min_j \sqrt{\inf_{0 \leq t \leq T, \bar{x} \in \mathcal{G}} [\lambda_{\min}(\Pi_j^t)]}} < 1$$

where

$$a_1 = \min_j \inf_{0 \leq t \leq T, \bar{x} \in \mathcal{G}} \left[ \lambda_{\min}(Q_t^j + \Pi_t^j BR^{-1} B^\top \Pi_t^j) \right] \sup_{0 \leq t \leq T, \bar{x} \in \mathcal{G}} \left[ \lambda_{\sup}(\Pi_t^j) \right],$$

$$b_1 = \max_j \sqrt{\sup_{0 \leq t \leq T, \bar{x} \in \mathcal{G}} \left[ \lambda_{\max}(\Pi_t^j) \right] \|BR^{-1}B^\top\| (\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4),$$

$$\gamma_1 = (\max_j \kappa_j^3 \lambda) \left[ (\sqrt{p/n} x_0 \bar{x}_0) + (\sqrt{p/n} \kappa_1 z) \right],$$

$$\gamma_2 = \lambda \sqrt{p/n} z \sup_{\bar{x} \in \mathcal{G}} \left\| \Psi^{\bar{x}}(t, T) \right\| T,$$

$$\gamma_3 = (\max_j \kappa_j^2 \lambda) (\sqrt{p/n} x_0 \bar{x}_0) (\sqrt{p/n} \kappa_1 z) (c_j \sup_{0 \leq t \leq T, \bar{x} \in \mathcal{G}} \left[ \lambda_{\max}(\Pi_t) \right]),$$

$$\gamma_4 = (\sqrt{p/n} z + \sqrt{n} \max_j \|c_j\| \kappa_2^2 \lambda) \int_{T} \sup_{\bar{x} \in \mathcal{G}} \left\| \Psi^{\bar{x}}(t, \tau) \right\| \tau d\tau,$$

where

$$\frac{d\Psi^{\bar{x}}(t, \tau)}{dt} = -[G^\top + \Lambda \otimes I] \Psi^{\bar{x}}(t, \tau),$$

$$G = \text{diag}(G_1, \ldots, G_p), \text{ and } G_j = A_j - BR^{-1}B^\top \Pi_j^j.$$
Collective Target Tracking MJ MF Stochastic Control Theorem

Under technical conditions the Collective Target Tracking MJ MF Equations have a unique solution which induces a set of controls 
\[ \mathcal{U}_\text{col}^N \triangleq \{(u^i)^0; 1 \leq i \leq N\}, 1 \leq N < \infty, \]
with
\[
u_t^0 = -\sum_{j=1}^{p} I[\theta_t=j] R^{-1} B^\top (\Pi_t^j x_t + s_t^j), \quad t \geq 0,
\]
such that

1. all agent system trajectories \( x^i, 1 \leq i \leq N \), are second order stable;
2. \( \{\mathcal{U}_\text{col}^N; 1 \leq N < \infty\} \) yields an \( \epsilon \)-Nash equilibrium for all \( \epsilon > 0 \).
Control Architecture

Figure: Control Architecture
Simulations

- 200 water heaters (60 gallons): 2 stratification layers
- Two elements with total maximum elemental power of 4.5kW
- Initial mean: 55°C
- 2 experiments:
  - Increase 2 °C mean temperature,
  - Decrease 2 °C mean temperature,
- Over a 6 hours control horizon
- Constant water extraction rate: 0.05 l/sec
- Time invariant 2 state Markov chain:
  - Arrival rate: 0.5
  - Departure rate: 7
  - Consequently average water consumption is 288 l/day

The central authority provides the target, local controllers apply collective target tracking mean field
Fixed Point Iterations

![Graph showing iterations over time with temperatures and iterations indicated.](image-url)
Energy Release: Collective Target Tracking Markovian Jump MF Control

Agents Applying Collective Target Tracking Markovian Jump MF Control:
All Agents Following the Low Comfort Level Signal
Aggregate Power Relief Curve

Finite Population Agg. Control
As $N \to \infty$ (adj. to 200 agents)

Aggregate Power Consumption
Agents Applying Collective Target Tracking Markovian Jump MF Control: All Agents Following the High Comfort Level Signal
Aggregate Power Accumulation Curve

Finite Population Agg. Control
As $N \rightarrow \infty$ (adj. to 200 agents)

Aggregate Power Consumption
Energy Accumulation: Heterogeneous Populations

Agents Applying Collective Target Tracking Markovian Jump MF Control
First group: higher initial temperature, second group: lower initial temperature
Second group's control penalty coefficient $R$ is lower than the first group
Energy Accumulation: Homogeneous vs Heterogeneous Populations

First group: higher initial temperature, second group: lower initial temperature

experiment 1: same control penalty coefficient $R$ for both groups

experiment 2: second group's control penalty coefficient $R$ is lower than the first group
Time Varying Target and Time Varying Water Consumption Profile

Agents Applying Collective Target Tracking Markovian Jump MF Control
Conclusions / Future Work

- Mean field control is a **natural** approach for load management in a smart grid context.
- It exploits the **predictability** of large number averages to produce **decentralized controls** with near centralized optimality properties.
- It preserves system **diversity** while minimizing communications requirements.
- It is a **flexible tool** for shaping control effort among devices.

Weakness:

- It **overly relies** on a **correct statistical** description of the underlying driving stochastic processes as well as the random distribution of device parameters.

Future work:

- Develop online device model parameter identification and adaptation algorithms.
- Consider time varying collective target tracking problems.
- Better address the impact of local constraints on global target generation.
- Investigate cooperative mean field control solutions.