Mean Field Games in Energy Systems Applications

Roland P. Malhamé

Department of Electrical Engineering
École Polytechnique de Montréal and GERAD

IPAM
Mean Field Games and Applications workshop
May, 2020

Collaborators

Arman Kizilkale

Rabih Salhab

Collaborators (Cont.)

Quentin Lenet

Overview

1. Introduction
2. Non-cooperative Collective Target Tracking Mean Field Control for Space Heaters
3. The Inverse-Nash approach
4. Non-cooperative Collective Target Tracking Mean Field Control for Water Heaters
5. Conclusions / Future Work

Motivation

- 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 (Malhamé-Chong, TAC 1985)

- Continuous State:

$$C_a dx_t^{in} = -U_a(x_t^{in} - x^{out})dt + Q_h m_t b_t dt + \sigma dw_t$$

divide by C_a and obtain

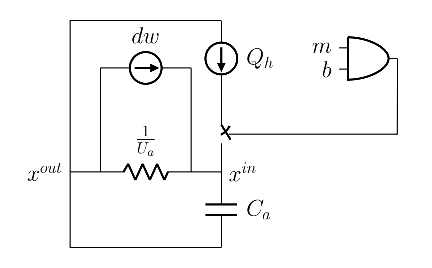
$$dx_t^{in} = a(x_t^{in} - x^{out})dt + Q'_h m_t b_t dt + \sigma' dw_t$$

- Discrete State: $m_{t+\Delta t} = m_t + \pi(x_t^{in}; x_+, x_-)$

$$\pi(x^{in}, m; x_+, x_-) = \begin{array}{ll} 0 & x_- < x^{in} < x_+ \\ -m & x^{in} \geq x_+ \\ 1-m & x^{in} \leq x_- \end{array}$$

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").

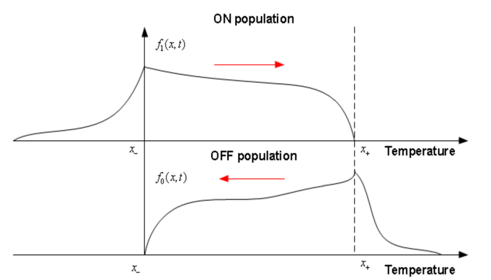


An Example: A Diffusion Model of Heating/Cooling Loads

The Coupled Fokker-Planck Equations

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

$$T^k_{\lambda,t}[f] = \frac{\partial f}{\partial t} - \frac{\partial}{\partial \lambda}[(a(\lambda - x_a(t)) - kb(t)R)f] - \frac{\sigma^2}{2}\frac{\partial^2}{\partial \lambda^2} f, \quad k = 0, 1$$



Fokker-Planck Equation Simulation

Model discretized and extended to second-order dynamics (Zhang *et al., 2013)*

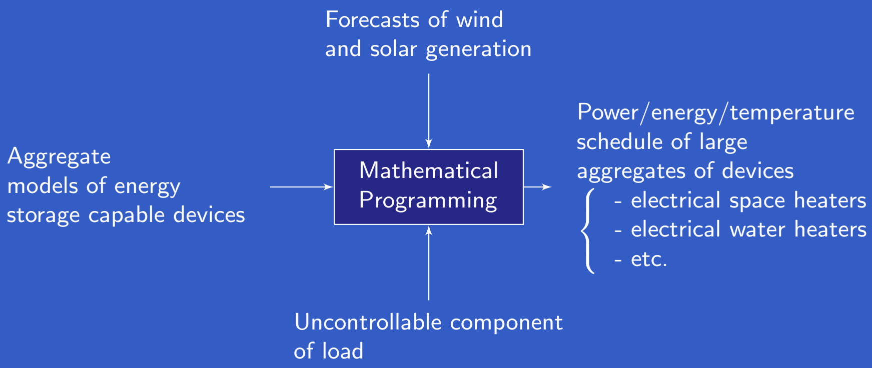
Remarks

- 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.

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



Prescriptive Linear Quadratic Mean Field Games: Why?

Linear Quadratic Mean Field Rendez-vous Problem

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

Prescriptive Linear Quadratic 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 simple enough to be computed locally 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.

1 Introduction

2 Non-cooperative Collective Target Tracking Mean Field Control for Space Heaters

3 The Inverse-Nash approach

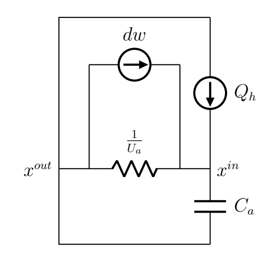
4 Non-cooperative Collective Target Tracking Mean Field Control for Water Heaters

5 Conclusions / Future Work

Elemental space heating model: two changes

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

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



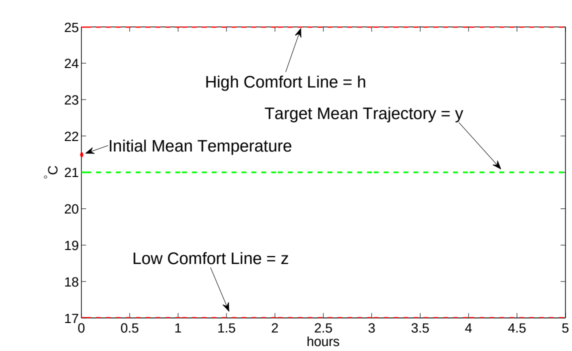
2. The control input is redefined so that no control effort is required on average to remain at initial temperature.

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

where $u^{free} \triangleq a(x^{out} - x_0^i)$.

- 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.

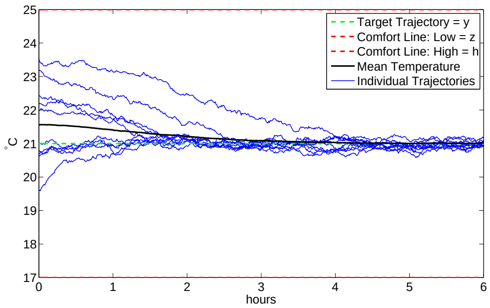
Constant Level Tracking Problem Setup



Classical LQG Target Tracking

$$J_i^N(u^i, u^{-i}) = \mathbb{E}\int_0^\infty e^{-\delta t}\left[\frac{1}{2}q(x_t^i - y)^2 + \frac{1}{2}r(u_t^i)^2\right]dt, \quad 1 \le i \le N$$

x^i : temperature
y : tracking target
u^i : control



Agents Applying LQG Tracking

A Novel Integral Control Based Cost Function

$$J_i^N(u^i, u^{-i}) = \mathbb{E} \int_0^\infty e^{-\delta t} \left[\frac{1}{2} q_t (x_t^i - z)^2 + \frac{1}{2} q^{x_0} (x_t^i - x_0^i)^2 + \frac{1}{2} r (u_t^i)^2 \right] dt$$

x^i : temperature
z : lower comfort bound
u^i : control

- Integral controller embedded in mean-target deviation coefficient: q_t, $t \in [0, \infty)$, calculated as the following integrated error signal:

$$q_t = \left| \int_0^t g(\bar{x}_\tau^N - y) d\tau \right|$$

$\bar{x}^N$: mean temperature of the population
y : mean target
g : Function $\mathbb{R} \mapsto \mathbb{R}$

Examples of mean-target deviation coefficients q_t

- Symmetric penalty $g(x) = \lambda x$, $\lambda > 0$:

$$q_t = \left| \lambda \int_0^t (\bar{x}_\tau^N - y) d\tau \right|$$

- Asymmetric penalty $g(x) = \lambda(1 - e^{\beta x})$, $\lambda > 0$, $\beta > 0$:

$$q_t = \left| \lambda \int_0^t \left(1 - e^{\beta(\bar{x}_\tau^N - y)}\right) d\tau \right|$$

Mean Field Based Collective Target Tracking

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

Continuum of Heaters and Generic Heater's Best Response

Under a continuum of heaters assumption, a generic heater's best response to q_t is the optimal control law the following LQG problem,

$$J(u) = \mathbb{E} \int_0^\infty e^{-\delta t} \left[\frac{1}{2} q_t (x_t - z)^2 + \frac{1}{2} q^{x_0} (x_t - x_0)^2 + \frac{1}{2} r (u_t)^2 \right] dt$$

s.t. $dx_t = -a(x_t - x_0) + bu_t - \sigma dw_t$

q_t is assumed known here.

Mean Field Equations

MF Equation System on $[0, \infty)$

$$
\begin{aligned}
-\frac{d\pi_t}{dt} &= (-2a-\delta)\pi_t - b^2 r^{-1}(\pi_t)^2 + q_t + q^{x_0} \\
-\frac{ds_t}{dt} &= (-a-\delta-b^2\pi_t r^{-1})s_t + (a\pi_t - q^{x_0})(\bar{x}_0 - z) \\
\frac{d\bar{x}_t}{dt} &= (-a - b^2\pi_t r^{-1})\bar{x}_t - b^2 r^{-1}(s_t - \pi_t z) + a\bar{x}_0 \\
q_t &= \left| \int_0^t g(\bar{x}_\tau - y) d\tau \right|
\end{aligned}
$$

- $\bar{x}_t$: Mean state of the infinite population when it optimally responds to q_t.
- $q_t = \left| \int_0^t g(\bar{x}_\tau - y) d\tau \right|$: Consistency requirement.
- A solution $\bar{x}$ is a fixed point of $\mathcal{M} := \mathcal{T} \circ \Delta$, where $\Delta(\bar{x}_t) = \left| \int_0^t g(\bar{x}_\tau - y) d\tau \right|$ and $\mathcal{T}(q_t)$ is the mean state of the infinite population when it optimally responds to q_t.

Preliminary Obervations

$\mathcal{T}(q_t)$ is the optimal state of the LQR problem:

$$J(u) = \int_0^\infty e^{-\delta t} \left(\frac{1}{2} q_t (x_t - z)^2 + \frac{1}{2} q^{x_0} (x_t - \bar{x}_0)^2 + \frac{1}{2} r u_t^2 \right) dt$$

$$\text{s.t. } \frac{d}{dt} x = -a(x - \bar{x}_0) + bu, \qquad x(0) = \bar{x}_0.$$

This implies two important results:

- A fixed point $\bar{x}$ of $\mathcal{M}$ (if it exists) must belong to $\mathcal{G} = \{x \in C[0, \infty), \quad z \leq x_t \leq y, \forall t \geq 0\}$.
- $\mathcal{M}$ is discontinuous when defined on $\mathcal{G}$, where $\mathcal{G}$ inherits its topology from the Banach space of bounded continuous functions $(C_b[0, \infty), ||_\infty)$.

Fixed Point Theorem

Instead of $(C_b[0,\infty), ||_\infty)$, we assume that $\mathcal{G}$ inherits its topology from the Banach space $(C_k[0,\infty), ||_k)$, where

- $C_k[0,\infty) = \left\{ x : x \in C[0,\infty) \text{ such that } \sup_{t\geq 0} |e^{-kt} x_t| < \infty \right\}$.
- $|x|_k = \sup_{t\geq 0} |e^{-kt} x_t|$.

Main Theorem

For $0 < k < a + \delta$, the map $\mathcal{M} : \mathcal{G} \subset (C_k[0,\infty), ||_k) \to \mathcal{G} \subset (C_k[0,\infty), ||_k)$ is continuous and Schauder's fixed point theorem guarantees the existence of at least one fixed point path $\bar{x}_t$.

ϵ-Nash Theorem

ϵ-Nash Theorem

Under technical conditions the collective target tracking MF stochastic control law generates a set of control laws $\{(u^i)^\circ; 1 \leq N < \infty\}$, where $(u^i)^\circ$ are defined for a fixed point $\bar{x}$ of $\mathcal{M}$, that constitute an ϵ-Nash equilibrium in the sense that, for all $\epsilon > 0$, there exists $N(\epsilon)$ such that for all $N \geq N(\epsilon)$

$$J_i^N \left((u^i)^\circ, (u^{-i})^\circ\right) - \epsilon \leq \inf_{u^i \in \mathcal{U}} J_i^N \left(u^i, (u^{-i})^\circ\right) \leq J_i^N \left((u^i)^\circ, (u^{-i})^\circ\right).$$

Desirable Fixed Point Algorithm

Goal: develop an algorithm to compute a fixed point mean trajectory $\bar{x}$ that converges to the target mean temperature y.

We consider only the cases $g(x) = \lambda x$ and $g(x) = \lambda(1 - e^{\beta x})$. The main idea of the algorithm:

1. Construct a family of mean trajectories indexed by λ, $\{\bar{x}(\lambda)\}_\lambda$, such that $\lim_{t\to\infty} \bar{x}'_t(\lambda) = \lim_{t\to\infty} \mathcal{M}_\lambda\big(\bar{x}(\lambda)\big)_t = y$.
2. Find a λ_* that minimizes $\|\bar{x}(\lambda) - \mathcal{M}_\lambda\big(\bar{x}(\lambda)\big)\|_{L_2}$.

Point 2) implies that $\bar{x}$ is approximately a fixed point.

Point 1) implies that $\bar{x}$ is a desirable fixed point (the mean trajectory of the population when it optimally responds to $\bar{x}$ converges to the target mean temperature y).

Desirable Fixed Point Algorithm (Cont.)

How to generate the family $\{\bar{x}(\lambda)\}_\lambda$?

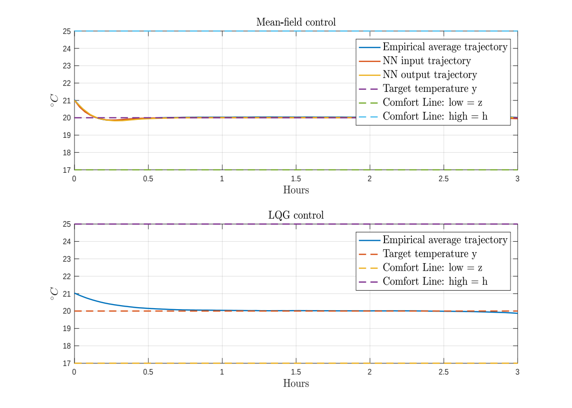
1. Define $q^*_\infty := \frac{[a(a+\delta)r+q^{x_0}b^2]}{b^2}\left(\frac{\bar{x}_0-y}{y-z}\right)$, such that if the population optimally responds to a q_t with $\lim_{t\to\infty} q_t = q^*_\infty$, then its mean temperature converges to y.
2. $$q_t = \begin{cases} nq^*_\infty, & \text{for } t \in [0, t_0] \\ q^*_\infty, & \text{for } t > t_0 \end{cases} \text{ generates } \bar{x}^s = \mathcal{T}(q_t).$$
   Similarly, we get $\bar{x}^i$ by replacing n by a larger number N. $\bar{x}^s \geq \bar{x}^i$, and they both converge to y.
3. Define $\lambda_s = q^*_\infty / \lim_{t\to\infty} \Delta_1(\bar{x}^s)$. Thus, $q^s_t = \Delta_{\lambda_s}(\bar{x}^s)$ converges to q^*_∞. Similarly, define λ_i.
4. Fix $\lambda_s \leq \lambda \leq \lambda_i$. If $\bar{x}^i \leq \bar{x} \leq \bar{x}' \leq \bar{x}^s$, then $\lim_{t\to\infty} \Delta_\lambda(\bar{x}) \leq \lim_{t\to\infty} \Delta_\lambda(\bar{x}')$, $q^*_\infty \in [\lim_{t\to\infty} \Delta_\lambda(\bar{x}^i), \lim_{t\to\infty} \Delta_\lambda(\bar{x}^s)]$
5. Use dichotomy to find a convex combination $\bar{x}(\lambda)$ of $\bar{x}^s$ and $\bar{x}^i$, such that $\lim_{t\to\infty} \Delta_\lambda(\bar{x}(\lambda)) = q^*_\infty$.

Simulations

- 200 heaters with initial temperatures drawn from a Gaussian distribution with mean 21°C and variance 1.
- decrease 1 °C the mean temperature over a 3 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.

Simulations (Cont.)

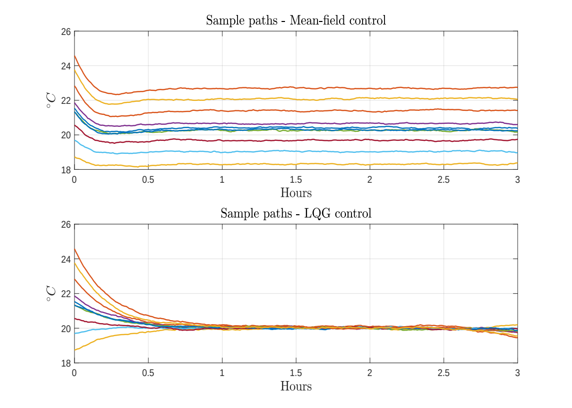
Collective tracking mean field based on $g(x) = \lambda x$:



For a large number of heaters (here, 200 heaters), the empirical average is approximately equal to the Near Nash (NN) output trajectory as a result of the law of large numbers.

Simulations (Cont.)

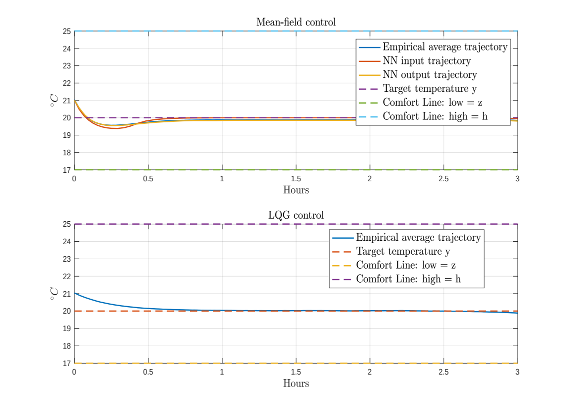
Collective tracking mean field based on $g(x) = \lambda x$:



Sample paths in the mean-field (linear g) and LQG control cases

Simulations (Cont.)

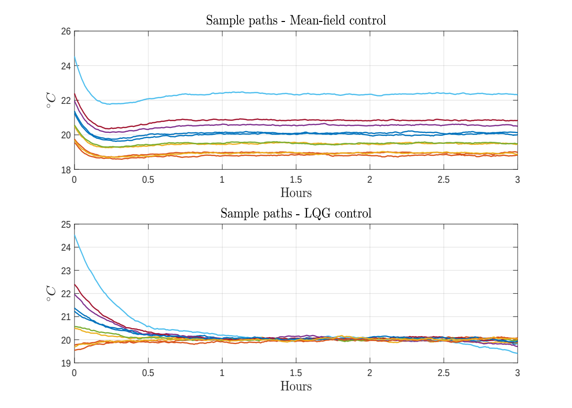
Collective tracking mean field based on $g(x) = \lambda(1 - e^{\beta x})$:



For a large number of heaters (here, 200 heaters), the empirical average is approximately equal to the Near Nash (NN) output trajectory as a result of the law of large numbers.

Simulations (Cont.)

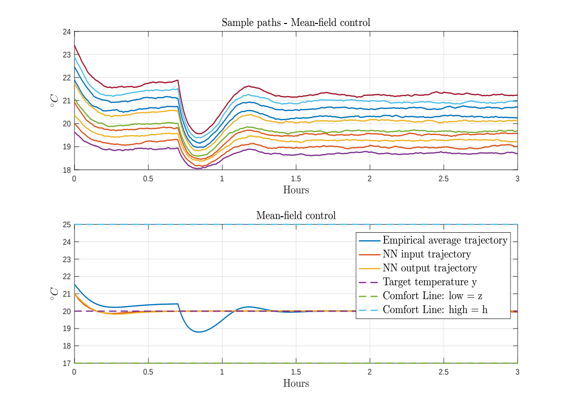
Collective tracking mean field based on $g(x) = \lambda(1 - e^{\beta x})$:



Sample paths in the mean-field (exponential g) and LQG control cases

Robustness

Error in the estimation of the initial mean and ambient temperature: Actual initial mean = 21.5 °C, estimated initial mean = 21 °C, actual ambient temperature = -11 °C, and estimated ambient temperature = -10 °C



1 Introduction

2 Non-cooperative Collective Target Tracking Mean Field Control for Space Heaters

3 The Inverse-Nash approach

4 Non-cooperative Collective Target Tracking Mean Field Control for Water Heaters

5 Conclusions / Future Work

Why Inverse-Nash ?

Shortcomings of iterative methods

- In practice converging towards an equilibrium can take a long time.
- It is often the case that one seeks not an energy increase or decrease target but rather a global *power* target.

Solution : Invert the approach

- Determine the shape of a would be desirable Nash-Equilibrium mean temperature trajectory, i.e. corresponding to some target mean power consumption.
- Compute the mean-field cost coefficient trajectory q_t from the numerical solving of the mean-field equations and an analytical form of the mean dynamics.

A finite horizon cost function

$$J^N(u^i, u^{-i}) = E\int_0^T \left[\frac{q_t}{2}(x_t^i - z)^2 + \frac{q_{x_0}}{2}(x_t^i - x_0^i)^2 + \frac{r}{2}(u_t^i)^2\right] dt$$

- q_t is the mean-field cost coefficient. Instead of defining it as an integral error-tracking coefficient, we will compute it numerically.

Expected mean dynamics

Average power target

We define $\overline{Q_h}^{target} = \frac{1}{N} Q_{h,tot}^{target}$

Mean dynamics

For an average power target $\overline{Q}_h$:

$$\dot{\overline{x}}_t^{target} = -a(\overline{x}_0 - \overline{x}_\infty)e^{-at}$$

$$\overline{x}_t^{target} = \overline{x}_\infty + (\overline{x}_0 - \overline{x}_\infty)e^{-at}$$

Where $\overline{x}_\infty = x_{out} + \frac{\overline{Q}_h^{target}}{U_a}$

Control equations

Individual equations

$$(u_t^i)^* = -\frac{b}{r}(\pi_t(x_t^i - z) + \beta_t^i)$$

$$\dot{\pi}_t = \frac{b^2}{r}(\pi_t)^2 + 2a\pi_t - q_{x_0} - q_t^y$$

$$\dot{\beta}_t^i = (a + \frac{b^2}{r}\pi_t)\beta_t^i - (a\pi_t - q_{x_0})(x_0^i - z)$$

Collective equations

$$\dot{\beta}_t = (a + \frac{b^2}{r}\pi_t)\beta_t - (a\pi_t - q_{x_0})(\overline{x}_0 - z)$$

$$\dot{\overline{x}}_t = -(a + \frac{b^2}{r}\pi_t)\overline{x}_t - \frac{b^2}{r}(\beta_t - \pi_t z) + a\overline{x}_0$$

Boundary conditions

To solve the control equations we define boundary conditions. We assume that if T is high enough, we can aproximate the horizon as steady-state.

Values at $t = T$

$$q_T^y = \frac{a^2 r + q_{x_0} b^2}{b^2} \left(\frac{\overline{x}_0 - \overline{x}_\infty}{\overline{x}_\infty - z} \right)$$

$$\pi_T = \frac{r}{b^2} \left(-a + \sqrt{a^2 + b^2 r^{-1} (q_{x_0} + q_T^y)} \right)$$

$$\beta_T = \frac{(a\pi_T - q_{x_0})(\overline{x}_0 - z)}{a + \pi_T \frac{b^2}{r}}$$

Solving the problem for q_t

- β_t is the solution of the following differential equation

$$\dot{\beta}_t = -\beta_t^2 \frac{b^2}{r(\overline{x}_t - z)} + \beta_t \left[a - \frac{a(\overline{x}_t - 2\overline{x}_0 + z) + \dot{\overline{x}}_t}{\overline{x}_t - z} \right] + (\overline{x}_0 - z) \left[q_{x_0} + \frac{ar(a(\overline{x}_t - x_0) + \dot{\overline{x}}_t)}{b^2(\overline{x}_t - z)} \right]$$

- We then get π_t thanks to the following equation

$$\pi_t = -\frac{\beta_t}{\overline{x}_t - z} - r\frac{(a(\overline{x}_t - \overline{x}_0) + \dot{\overline{x}}_t)}{b^2(\overline{x}_t - z)}$$

With $\overline{x}_t = \overline{x}_t^{target}$ as previously defined.

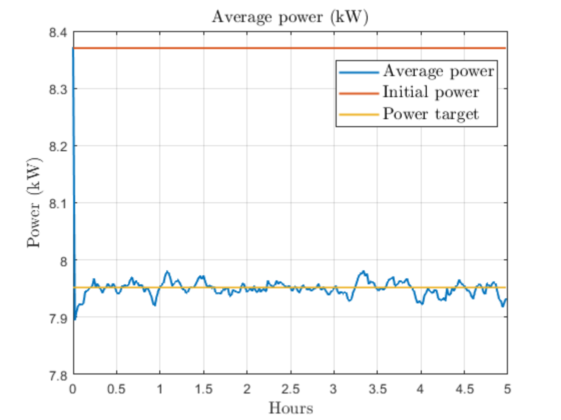
- We can finally numerically compute q_t

Simulations

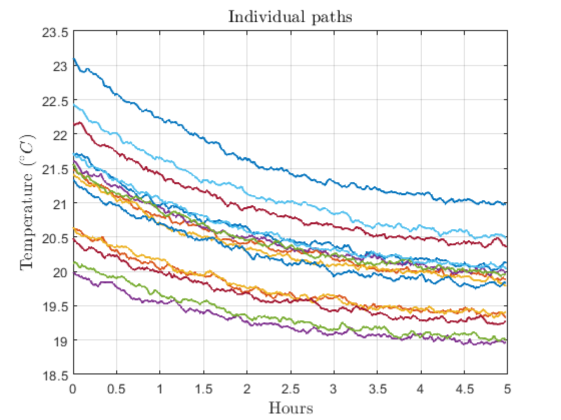
- 200 heaters with initial temperatures drawn from a Gaussian distribution with mean 21°C and variance 1.
- Decrease the total heating power by 5% and follow the evolution on a 5 hour interval.
- The central authority provides target, local controllers apply Inverse-Nash control.

Simulations (cont.)

Average heating power



Individual behaviours under Inverse-Nash control



1 Introduction

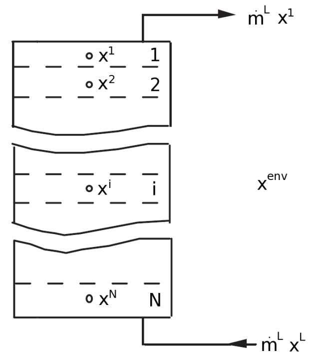
2 Non-cooperative Collective Target Tracking Mean Field Control for Space Heaters

3 The Inverse-Nash approach

4 Non-cooperative Collective Target Tracking Mean Field Control for Water Heaters

5 Conclusions / Future Work

Water Heater Stratification Model



x_l	temperature of the lth segment
u_l	control action at the lth segment
$\dot{m}_L$	fluid mass flow rate to the load
$\dot{Q}_l$	rate of energy input by the heating element
x^{env}	temperature of the environment
x^L	temperature of the inlet fluid
M_l	mass of the fluid in the lth segment
A_l	surface area of the lth segment
C^{pf}	specific heat of the fluid
U	loss coefficient betwen the tank and its env.

$$M_l C^{pf} \frac{dx_{l,t}}{dt} = UA_l(x^{env} - x_{l,t}) + \dot{m}_t^L C^{pf}(x_{(l+1),t} - x_{l,t}) + \dot{Q}_l u_{l,t}, \quad t \geq 0, \quad l \neq n$$

$$M_l C^{pf} \frac{dx_{l,t}}{dt} = UA_l(x^{env} - x_{l,t}) + \dot{m}_t^L C^{pf}(x_t^L - x_{l,t}) + \dot{Q}_l u_{l,t}, \quad t \geq 0, \quad l = n$$

S. Klein, "A design procedure for solar heating systems," Ph.D. dissertation, Department of Chemical Engineering, University of Wisconsin-Madison, 1976.

Elemental Agent Dynamics

- $\dot{m}^L$: modeled as a (stochastic) jump process.
- Physical model:

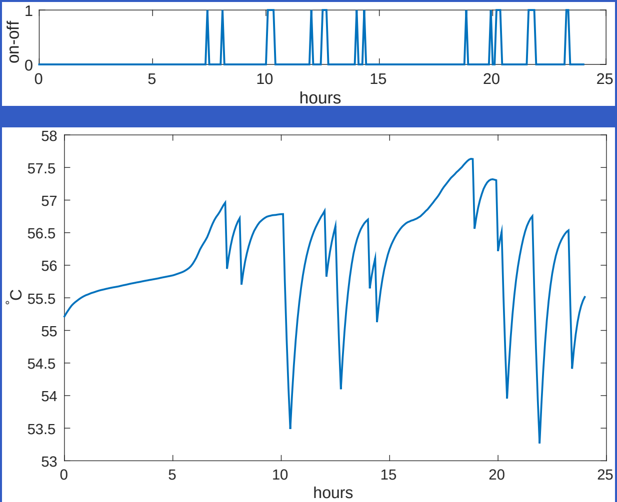
$$M_l C^{pf} \frac{dx_{l,t}}{dt} = UA_l(x^{env} - x_{l,t}) + \dot{m}_t^L C^{pf}(x_{(l+1),t} - x_{l,t}) + \dot{Q}_l u_{l,t}, \quad t \geq 0, \quad l \neq n$$
$$M_l C^{pf} \frac{dx_{l,t}}{dt} = UA_l(x^{env} - x_{l,t}) + \dot{m}_t^L C^{pf}(x_t^L - x_{l,t}) + \dot{Q}_l u_{l,t}, \quad t \geq 0, \quad l = n$$

- We write it as:

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

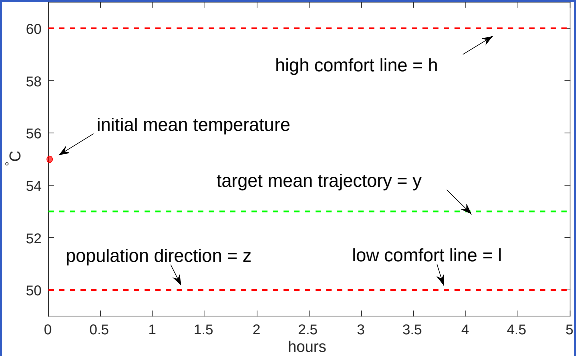
- ρ_t, $t \geq 0$, is a continuous time Markov chain taking values in $L = \{1, 2, ..., p\}$ with infinitesimal generator matrix $\Lambda(t)$. Each discrete value is associated with a type of event (showers, dishwashers, etc ...)

Simulations



A Single Agent's Temperature Trajectory Applying Markovian Jump LQ Tracking

Constant Level Tracking Problem Setup



Redefined Dynamics

Dynamics for a population of N water heaters:

$$\frac{dx_t^i}{dt} = A'^{\rho_t^i} x_t^i + B u_t^i + c'^{\rho_t^i}, \quad t \geq 0, \quad 1 \leq i \leq N$$

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

$$\frac{dx_t^i}{dt} = A^{\rho_t^i} x_t^i + B(u_t^i + u_t^{i,free}) + c^{\rho_t^i}, \quad t \geq 0, \quad 1 \leq i \leq N,$$

where

$$\sum_{l=1}^{n} B u_l^{free} = \sum_{l=1}^{n} UA_l(x_{l,0} - x^{env}) + \mathbb{E} \sum_{j=1}^{p} \zeta^j(t) \dot{m}_t^L(j) C^{pf}(x_{1,t} - x_t^L)$$

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

Integral Control Based Cost Function

Cost functions:

$$J_i^N(u^i, u^{-i}) = \mathbb{E} \int_0^\infty e^{-\delta t} \left[(Hx_t^i - z)^2 q_t^y + (Hx_t^i - Hx_0^i)^2 q^{x_0} + \|u_t^i\|_R^2 \right] dt$$

x^i	temperature
z	lower comfort bound
u^i	control
H	$[1/n, ..., 1/n]$

Integral controller embedded in mean-target deviation coefficient: q_t^y, $t \in [0, T]$, calculated as the following integrated error signal:

$$q_t^y = \left| \lambda \int_0^t (H\bar{x}_t^N - y) dt \right|$$

$\bar{x}^N$	mean temperature of the population
y	mean target

Synthesis of the Mean Field Control Law: Step 1

For a given $\bar{x}_t$ and thus q_t^y, $t \in [0, T]$, compute optimal agent response: [W. M. Wonham, 1971]

- each agent $\mathcal{A}_i, 1 \leq i \leq N$, obtains the positive solution to the coupled set of Riccati equations

$$-\frac{d\Pi_t^j}{dt} = \Pi_t^j \left(A^j - \frac{\delta}{2} I\right) + \left(A^j - \frac{\delta}{2} I\right)^\top \Pi_t^j$$

$$-\Pi_t^j B R^{-1} B^\top \Pi_t^j + \sum_{k=1}^{p} \lambda_{jk}(t) \Pi_t^k + (q_t^y + q^{x_0}) H^\top H, \quad 1 \leq j \leq p$$

Synthesis of the Mean Field Control Law: Step 1

- for a given target signal z, the individual ith agent offset function is generated by the coupled differential equations

$$-\frac{ds^j_{i,t}}{dt} = (A^j - \delta I - BR^{-1}B^\top \Pi^j_t)^\top s^j_{i,t} - q^y_t H^\top z - q^{x_0} H^\top x^i_0 + \Pi^j_t c^j_i + \sum_{k=1}^{p} \lambda_{jk}(t) s^k_{i,t}, \quad 1 \leq j \leq p$$

- the optimal tracking control law is given by

$$u^\circ_{i,t} = -\sum_{j=1}^{p} I_{[l_{i,t}=j]} R^{-1} B^\top (\Pi^j_t x_{i,t} + s^j_{i,t}), \quad t \geq 0.$$

Fixed Point Equation System: Step 2

Step 2: Add consistency requirements. Under best response to posited $\bar{x}_t$ (or q_t), agents mean must replicate $\bar{x}_t$.

$$-\frac{ds^j_{\theta,t}}{dt} = (A^j - \delta I - BR^{-1}B^\top \Pi^j_t)^\top s^j_{\theta,t} - q^y_t H^\top z - q^{x_0} H^\top x^\theta_0 + \Pi^j_t c^j + \sum_{k=1}^{p} \lambda_{jk}(t) s^k_{\theta,t}$$

$$\frac{d\bar{x}^j_{\theta,t}}{dt} = (A^j - BR^{-1}B^\top \Pi^j_t)\bar{x}^j_{\theta,t} + \sum_{k=1}^{p} \lambda_{kj}(t)\bar{x}^j_{\theta,t} + \zeta^j(t)c^j - \zeta^j(t)BR^{-1}B^\top s^j_{\theta,t}$$

$$\frac{d\zeta(t)}{dt} = \zeta(t)\Lambda(t)^\top$$

$$\bar{x}_{\theta,t} = \sum_{j=1}^{p} \bar{x}^j_{\theta,t}$$

$$\bar{x}_t = \int_\Theta \bar{x}_{\theta,t} dF^\theta$$

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

where

- $\bar{x}^j_{\theta,t} = \mathbb{E}(\bar{x}_{\theta,t} I_{[\theta_t = j]})$
- $\Lambda(t) = \{\lambda_{ij}(t), i, j = 1, ..., p\}$ is the infinitesimal generator of the Markov chain
- $\zeta(t) = [\zeta^1(t), ..., \zeta^p(t)]$ is the probability distribution of the Markov chain

Control Architecture

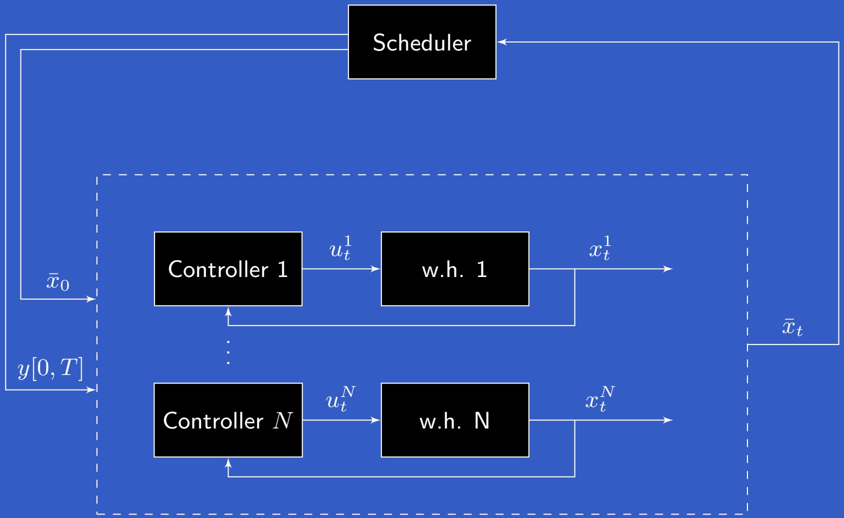


Figure: Control Architecture

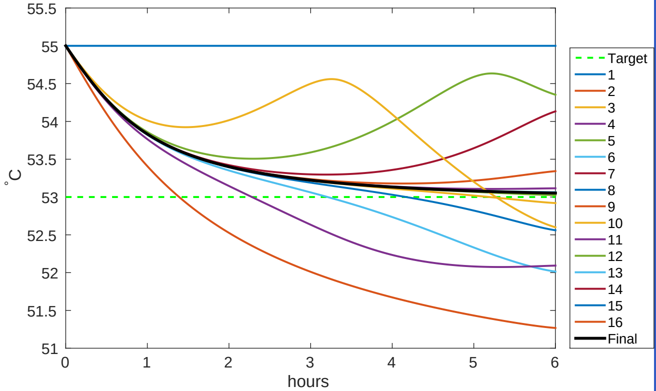
Simulations

- 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

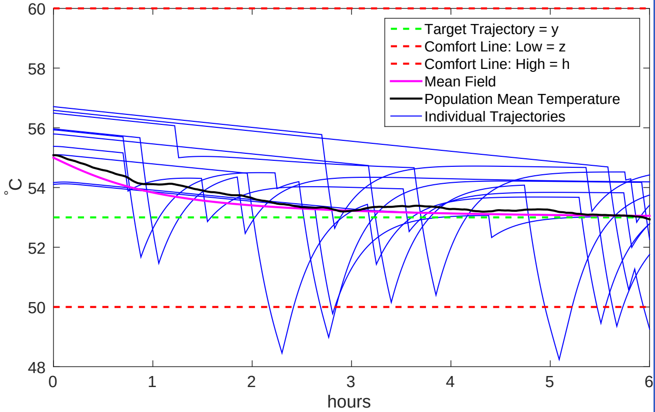
  The central authority provides the target, local controllers apply collective target tracking mean field

Fixed Point Iterations



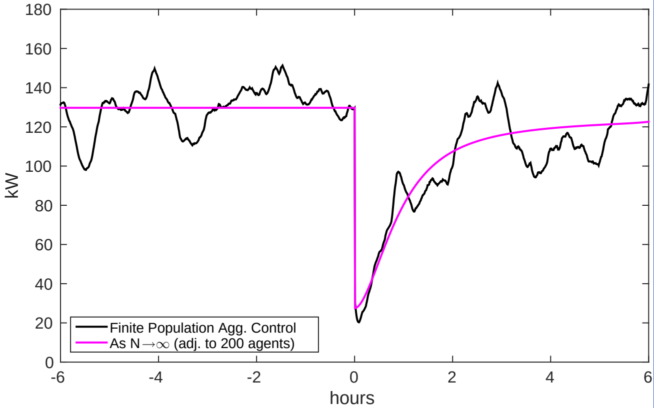
Iterations

Energy Release: Collective Target Tracking Markovian Jump MF Control



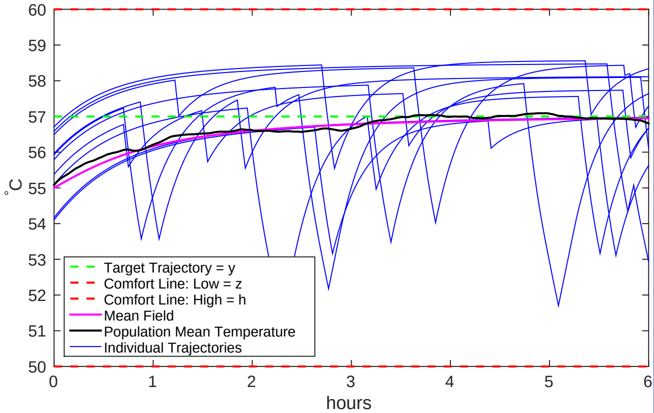
Agents Applying Collective Target Tracking Markovian Jump MF Control:
All Agents Following the Low Comfort Level Signal

Aggregate Power Relief Curve



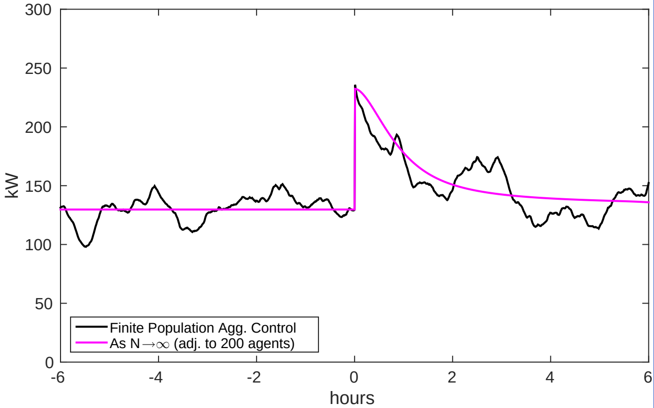
Aggregate Power Consumption

Energy Accumulation: Collective Target Tracking Markovian Jump MFC



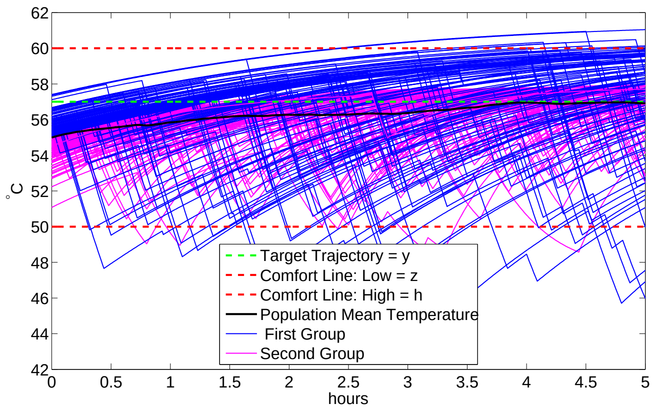
Agents Applying Collective Target Tracking Markovian Jump MF Control:
All Agents Following the High Comfort Level Signal

Aggregate Power Accumulation Curve



Aggregate Power Consumption

Energy Accumulation: Heterogeneous Populations

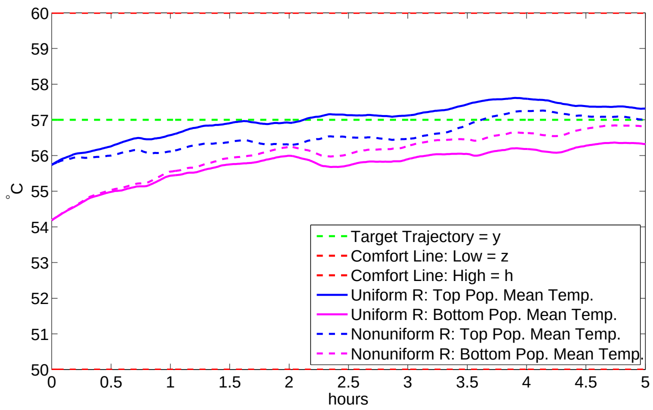


Agents Applying Collective Target Tracking Markovian Jump MF Control

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

1 Introduction

2 Non-cooperative Collective Target Tracking Mean Field Control for Space Heaters

3 The Inverse-Nash approach

4 Non-cooperative Collective Target Tracking Mean Field Control for Water Heaters

5 Conclusions / Future Work

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 arbitrary collective target tracking problems using the inverse Nash approach.
- Better address the impact of local constraints on global target generation.
- Investigate cooperative mean field control solutions.

Thank you