Control of Aggregate Electric Water Heating Loads via Mean Field Games Based Methods

Roland P. Malhamé

work with Arman C. Kizilkale

Department of Electrical Engineering
École Polytechnique de Montréal and GERAD

Advances in Modelling and Control for Power Systems of the Future
Palaiseau, Sep 2018

Practical context: The changing role of load management

- Initial idea: Using deferrable grid electric loads, mainly thermal loads (electric space heating, air conditioners, refrigerators) to shave electric demand peaks and fill valleys
  - deferred generation and transmission system expansion, efficiency of electricity production
- Currently: Strong push for renewable sources
  - Intermittency
  - A control architecture to convert dispersed grid thermal loads into an effective energy storage potential
- Challenges: Literally millions of devices to (in theory) observe and actuate
  - staggering communication load, staggering computations, challenge to keep customers happy

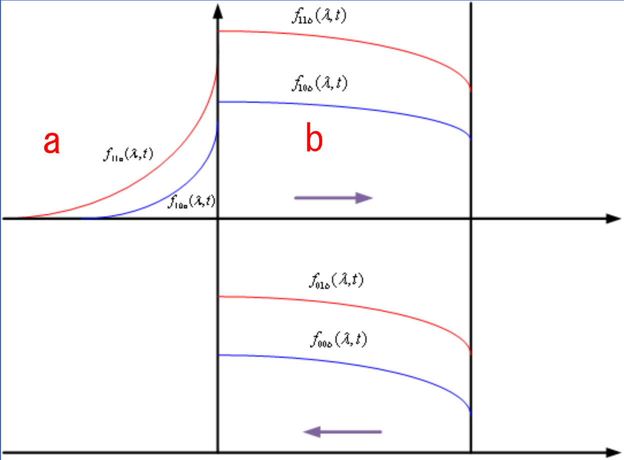
Past Approach: Homogeneous deterministic direct load control

- Idea: Send the same interruption/reconnection signal to very large numbers of homogeneous electric devices attached to thermostatically controlled loads.
- Analytical tools (statistical mechanics): Microscopic level stochastic models under uniform control + Law of large numbers
  - Ensemble statistics (PDE description)
- An example (Laurent - Malhamé 1994): Elemental hybrid state (continuous temperature, discrete thermostat) stochastic model of water heaters with Kolmogorov equations based wave model at the aggregate level.
- Main problem: Uniform controls applied to devices in potentially very different states.
  - Too careful $\rightarrow$ too conservative.
  - Too daring $\rightarrow$ a fraction of customers unhappy.

Improvements since

- Higher order PDE model extensions: (Zhao, Zhang 2017)
- PDE based randomized controls locally implemented (Totu, Winieski 2017)
- PDE based identification approaches (Moura, Bendsten, Ruiz 2014)
- Aggregation and state estimation based control (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)

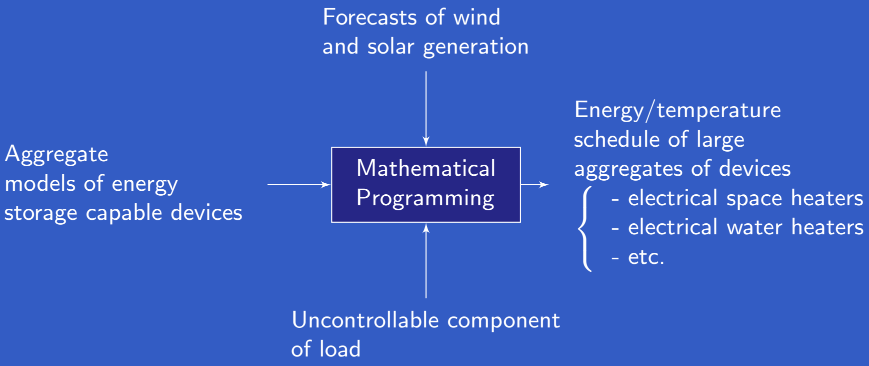
Aggregate PDE model of electric water heaters (1994)



Implementation principles

1. Decentralization: Each controller has to be situated locally because comfort and safety constraints can be locally secured.
2. Parsimony of communications: Communications should be kept at minimum both with the central authority and among users.
3. Minimal intrusiveness of controls: Deviations from uncontrolled behavior should be minimized.

Envisioned Overall Architecture: The case of a single central authority



MFG = (Statistical Mechanics + Optimal Control Theory)

Linear Quadratic Mean Field Rendez-vous Problem

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

MFG's: A Fast Growing Literature

Controls and PDEs:

Huang M. Y., Malhame R. P. and Caines P. E. (2006), “Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. Communications in Information and Systems. Vol. 6, No. 3, pp. 221-252.

Lasry J. M. and Lions, P. L. (2006), “Jeux à champ moyen. I-Le cas stationnaire”. Comptes Rendus Mathematique. Vol. 343, No. 9, pp. 619-625.

Lasry J. M. and Lions, P. L. (2006), “Jeux à champ moyen. II-Horizon fini et contrôle optimal”. Comptes Rendus Mathematique. Vol. 343, No. 10, pp. 679-684.

Huang M. Y., Caines P. E. and Malhamé R. P. (2007), “Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria”. IEEE Tans. on Automatic Control, Vol. 52, No. 9, pp. 1560-1571.

Books:

Bensoussan A., Frehse J., Yam P. (2013), "Mean Field Games and Mean Field Type Control Theory" Carmona R.,

Delarue F. (2018), "Probabilistic Theory of Mean Field Games with Applications I and II"

Mean Field Games: The Reasons?

Two fundamental reasons:

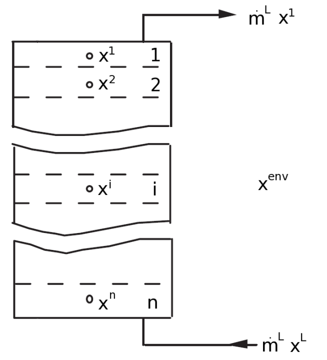
- Games are a natural device for enforcing decentralization.
- The large numbers involved induce decoupling effects which allow the law of large numbers to kick in.

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.

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

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 to
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	lateral surface area of the lth segment
C^{pf}	specific heat of the fluid
U	loss coefficient between the tank and its enviror

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

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

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:

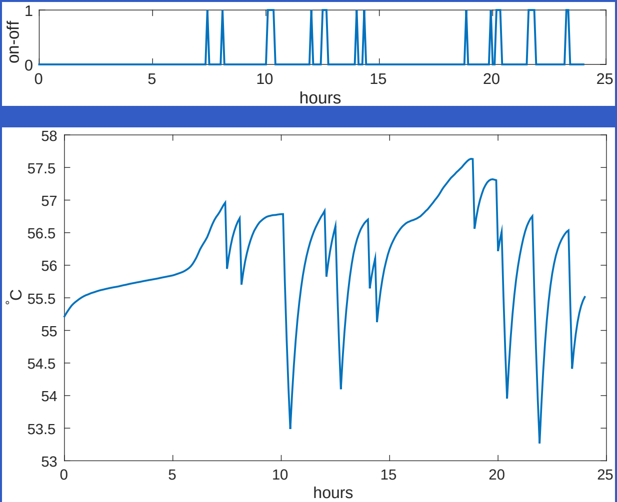
$$M_l C^{pf} \frac{dx_{l,t}}{dt} = UA_l(x^{env} - x_{l,t}) + \dot{m}_t^L C^{pf}(x_{(l+1),t} - x_{l,t}) + \dot{Q}_l u_{l,t},$$
$$t \geq 0, \quad l \neq n$$
$$M_l C^{pf} \frac{dx_{l,t}}{dt} = UA_l(x^{env} - x_{l,t}) + \dot{m}_t^L C^{pf}(x_t^L - x_{l,t}) + \dot{Q}_l u_{l,t},$$
$$t \geq 0, \quad l = n$$

- We write it as:

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

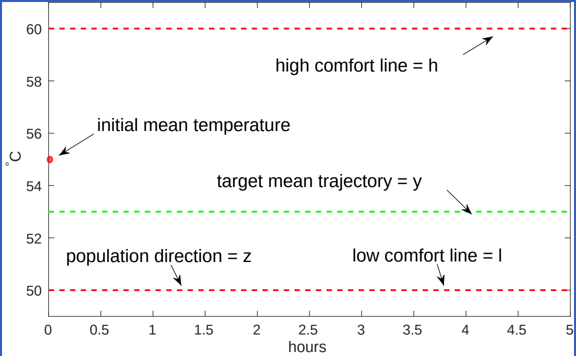
- θ_t, $t \geq 0$, is a continuous time Markov chain taking values in $\Theta = \{1, 2, ..., p\}$ with infinitesimal generator matrix Λ. Each discrete value is associated with a type of event (showers, dishwashers, etc ...)

Sample Trajectory



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:

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

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

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

where

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

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

Integral Control Based Cost Function

Cost functions:

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

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:

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

$\bar{x}^N$	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

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

where

$$\Pi_T^j = (q_T^y + q^{x_0}) H^\top H, \quad 1 \leq j \leq p$$

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

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

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

$$u_{i,t}^\circ = -\sum_{j=1}^{p} I_{[\theta_{i,t}=j]} R^{-1} B^\top (\Pi_t^j x_{i,t} + s_{i,t}^j), \quad t \geq 0.$$

Fixed Point Equation System: Step 2

Under best response to posited $\bar{x}_t$, agents mean must replicate $\bar{x}_t$.

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

$$-\frac{d\Pi_t^j}{dt} = \Pi_t^j A^j + A^{j\top}\Pi_t^j - \Pi_t^j B R^{-1} B^\top \Pi_t^j$$
$$+ \sum_{k=1}^p \lambda_{jk}\Pi_t^k + (q_t^y + q^{x_0})H^\top H, \quad \Pi_T^j = (q_T^y + q^{x_0})H^\top H, \quad 1 \le j \le p,$$

$$-\frac{ds_t^j}{dt} = (A^j - BR^{-1}B^\top \Pi_t^j)^\top s_t^j - q_t^y H^\top z - q^{x_0} H^\top \bar{x}_0$$
$$+ \Pi_t^j c^j + \sum_{k=1}^p \lambda_{jk} s_t^k, \quad s_T^j = -[q_T^y H^\top z + q^{x_0} H^\top \bar{x}_0], \quad 1 \le j \le p,$$

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

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

$$\frac{d\zeta_t}{dt} = \zeta_t \Lambda^\top.$$

One recalls

- $\bar{x}_t^j = E(\bar{x}_t I_{[\theta_t = j]})$
- $\Lambda = \{\lambda_{i'j'}, i', j' = 1, ..., p\}$ 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 func. s.t. $f \in \mathbf{C}_b[0,T], f(0) = \bar{x}_0$ and $z \leq f(t) \leq \bar{x}_0, t \in [0,T]$.

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

Theorem: The assumption below 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

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

where

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

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

$$\gamma_1 = (\max_j \kappa_3^j \lambda)\left[(\sqrt{p/n}q^{x_0}\bar{x}_0) + (\sqrt{p/n}\kappa_1 z)\right],$$

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

$$\gamma_3 = (\max_j \kappa_2^j \lambda)(\sqrt{p/n}q^{x_0}\bar{x}_0)(\sqrt{p/n}\kappa_1 z)(c_j \sup_{(0\leq t\leq T,\bar{x}\in\mathcal{G})}[\lambda_{max}(\Pi_t)]),$$

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

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

$G = diag(G_1,\ldots,G_p)$, and $G_j = A_j - BR^{-1}B^\top\Pi^j$.

ϵ-Nash Theorem

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 $\mathcal{U}_{col}^N \triangleq \{(u^i)^0; 1 \leq i \leq N\}$, $1 \leq N < \infty$, with

$$u_t^\circ = -\sum_{j=1}^{p} I_{[\theta_t = j]} R^{-1} B^\top (\Pi_t^j x_t + s_t^j), \quad t \geq 0,$$

such that

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

Control Architecture

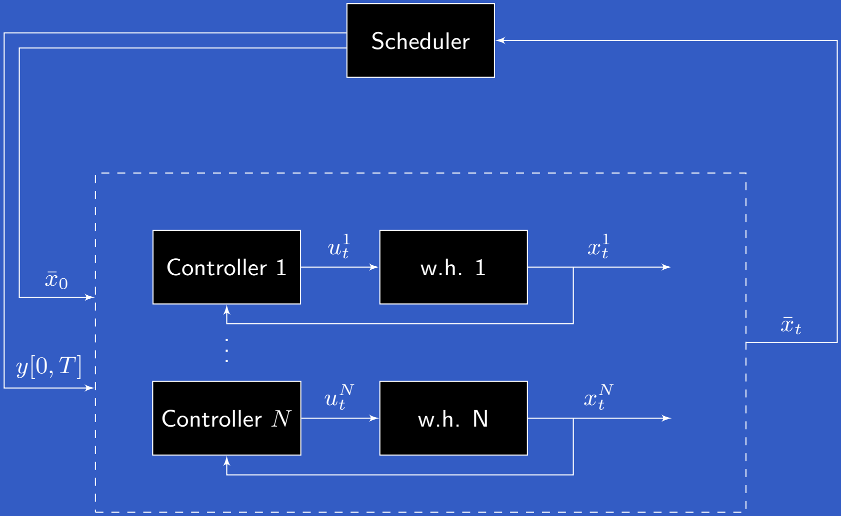


Figure: Control Architecture

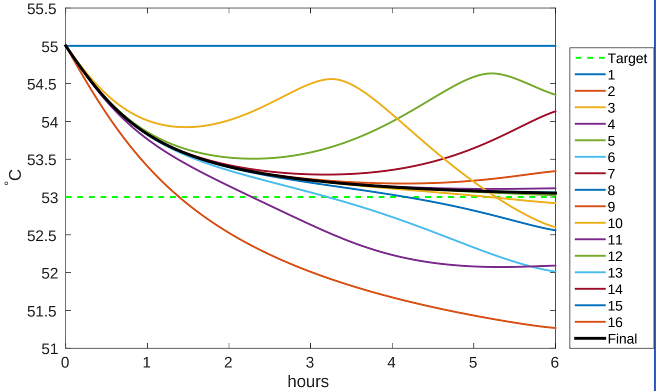
Simulations

- 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,

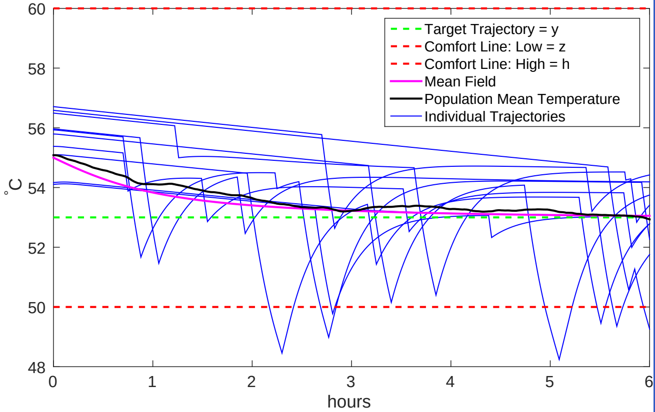
  over a 6 hours control horizon
- constant water extraction rate: 0.05 l/sec
- time invariant 2 state Markov chain:
    - arrival rate: 0.5
    - departure rate: 7
    - consequently average water consumption is 288 l/day

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

Fixed Point Iterations

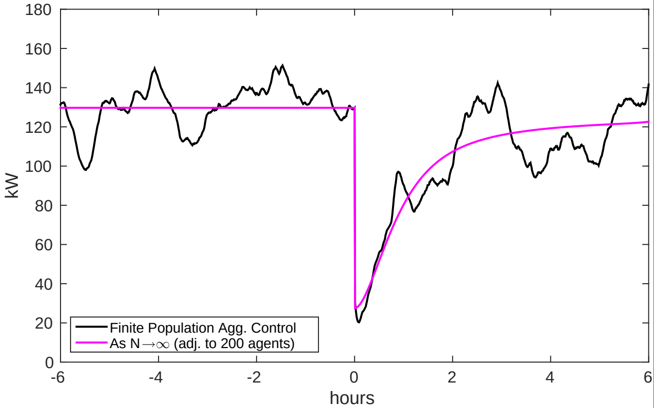


Iterations



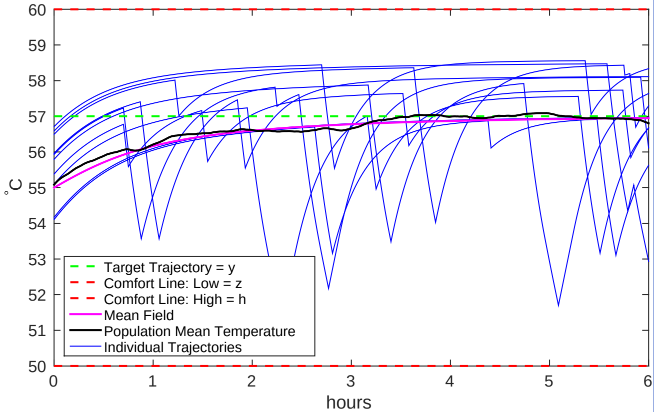
Agents Applying Collective Target Tracking Markovian Jump MF Control:
All Agents Following the Low Comfort Level Signal

Aggregate Power Relief Curve



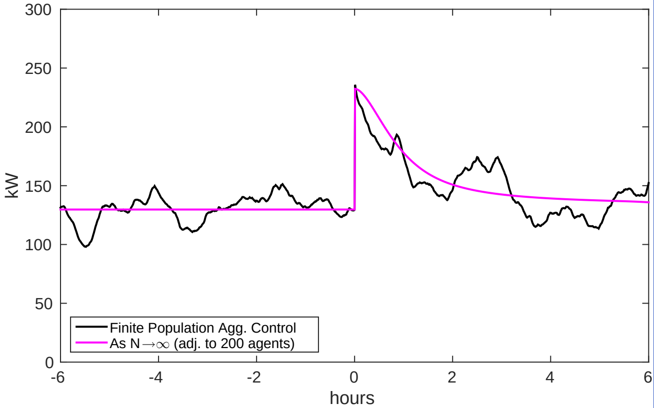
Aggregate Power Consumption

Energy Accumulation: Collective Target Tracking Markovian Jump MFC



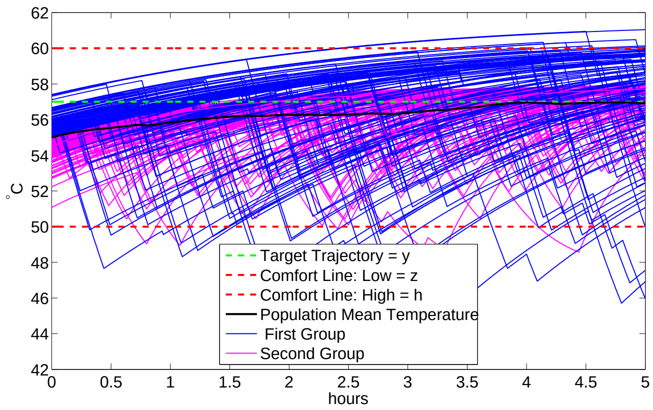
Agents Applying Collective Target Tracking Markovian Jump MF Control:
All Agents Following the High Comfort Level Signal

Aggregate Power Accumulation Curve



Aggregate Power Consumption

Energy Accumulation: Heterogeneous Populations

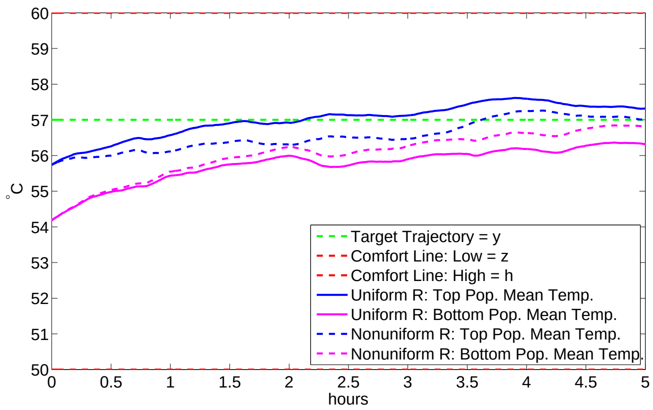


Agents Applying Collective Target Tracking Markovian Jump MF Control

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

Conclusions / Future Work

- Mean field games based 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.