Les Cahiers du GERAD ISSN: 0711–2440

Load shaping via grid wide coordination of heating-cooling electric loads: A mean field games based approach

A.C. Kizilkale
R.P. Malhamé

G–2015–68

July 2015

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Load shaping via grid wide coordination of heating-cooling electric loads: A mean field games based approach

Arman C. Kizilkale
Roland P. Malhamé

GERAD & Department of Electrical Engineering, Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7

arman.kizilkale@polymtl.ca
roland.malhame@polymtl.ca

July 2015

Les Cahiers du GERAD

G–2015–68

Abstract: Pressure on ancillary reserves in power systems has significantly mounted due to the recent generalized increase of the fraction of (highly fluctuating) wind and solar energy sources in grid generation mixes. Dedicated energy storage devices have seen their role reaffirmed as potentially low carbon print, if expensive tools, for smoothing the resulting generation/demand imbalances. However, a hitherto under utilized, relatively inexpensive energy storage alternative is that formed by the tiny energy wells of electric origin attached to millions of individual customer electric thermal loads. A hierarchical mean field games approach is proposed for shaping their collective load, whereby the top level sets system optimal mean aggregate temperature target trajectories. In turn based on a local state and a mean field dependent cost function, each individual load develops a decentralized local control law such that the aggregate load can meet the set targets. This control law is to be followed only as long as local comfort and safety constraints are secured, thus guaranteeing acceptability by customers. The corresponding mathematical theory is developed and numerical results are reported.

Acknowledgments: The authors gratefully acknowledge the support of Natural Resources Canada.

1 Introduction

The advances in smart grid systems could enable users to track the electricity price signals periodically. At first sight, one would expect that an important fraction of consumers would react by lowering their consumption when prices are high, and displacing the power consuming jobs towards cheaper power hours, thus achieving system wide load relief at peak hours. However, individual responses without a proper dynamic optimization perspective could lead to difficult to predict oscillations in the aggregate load, and in fact possibly amplified albeit delayed peaks.

Reducing peak load is a challenging task for system operators. One implementation alternative aimed at achieving peak load reduction, given prior agreements with customers, would be to employ a centralized controller which, given forecasts of the uncontrolled portions of the demand, as well as aggregate demand models of the controllable loads, would generate within safety and comfort customer constraints, a power usage schedule for participating customers. However, such a standard implementation runs into scalability problems, and there would be a need for high levels of data exchanges between the control authority and the devices. Past research on aggregate modeling (see [1], [2] for example) leads to models which can anticipate controlled device states statistical distribution with minimal observation. These classes of models have been leveraged in [3], [4] as well as extended in [5]. However limitations are still present concerning heterogeneity of load parameters and communication/estimation needs.

Several reasons favor renewable energy (wind, solar, etc.) as a crucial option for the power grid; the most notable are absence of carbon emissions, cost independence from the highly volatile and mostly increasing oil prices, and the potential deferral of costs of carbon emission taxes. Nevertheless, renewable sources such as wind and solar tend to be structurally problematic; they lack continuous availability due to factors out of the power system direct control. This fact leads to more need for rapidly dispatchable production which is typically costly and polluting. In our current line of research we further explore the potential of energy storage in dispersed devices naturally present in power systems, whether associated with electric water heaters, electric space heaters, air conditioners, batteries of plug-in electric vehicles etc., as a tool to mitigate renewable generation variability and reduce peak load. In this particular case a demand dispatch mechanism is implemented to employ the free storage in space heaters such that (i) intermittent renewable penetration can be increased in the grid by using this storage capacity, and (ii) load peaks can be mitigated by smoothing the aggregate load. The envisioned control architecture is hybrid: (i) centralized in terms of target trajectory generation for adequately partitioned groups of energy storage capable electric devices, so as to preserve overall optimality characteristics, (ii) decentralized at the implementation level so as to locally enforce safety and comfort constraints, as well as to minimize communication requirements. More specifically, we mention the following implementation principles, and argue that a class of decentralized control schemes based on a mean field game (MFG) setup (see [6], [7–10]) can actually meet all the requirements. Note that while we refer to a central authority in our discussion, the proposed control scheme could be applied by so-called "aggregators" [11] which manage large groups of energy storing devices as a storage or "negative load" resource on electricity markets.

(i) Each controller has to be situated locally.

As mentioned earlier, a completely centralized control architecture micromanaging every individual device to be controlled requires significant communication requirements as well as a very large computational power. Moreover, in the event of a loss of communication, users might face difficulties in the sense of comfort and safety. When the controller is situated locally, these worries are void since the controller can locally enforce comfort and safety constraints. Even if the communication with the central authority is lost, it is able to respect the safety and comfort requirements of the user. As it turns out, MFG controllers are thoroughly decentralized.

(ii) Data exchanges should be kept minimum both with the central authority and among users.

The control architecture must be able to operate with a highly reduced volume of data exchanges since, given the very large number of agents involved, a requirement of constant flows of information between agents and the central authority, as well as amongst agents themselves, besides confidentiality issues, would create scalability problems. In this context, the MFG based controllers turn the large numbers issue from a hurdle into an asset, by relying on the predictability stemming from the law of large numbers. Note that it is precisely the law of large numbers which is largely behind the successes of statistical mechanics, a mathematical theory at the basis of MFG. MFG based controllers can indeed operate by relying only on an agent's local observations and for example, for the case of thermal heating-cooling loads, hereon considered for illustration purposes as electric space heaters, the shared information on the initial electrically heated spaces population mean temperature, as well the mean target temperature trajectory over an adequate control horizon, as precomputed by the central authority using deterministic macroscopic aggregate models.

(iii) User disturbance should be kept at minimum.

Since the centralized authority is solely interested in the aggregate consumption, individual trajectories do not need to necessarily follow the (mean) targets set by the authority. On the contrary, in fact in the case of a population of thermal loads, it is desirable to shape the mean temperature of the population with least disturbance; ideally without customers even noticing the effects of the control actions on their comfort level. Also, it is important to maintain some measure of fairness among the users when it comes to sharing the control effort. We shall show that the developments in this work address these issues.

We introduce *collective target tracking mean field control*, where the presence of large numbers of space heating electric devices is employed to develop a decentralized mean field control based approach to the problem of these devices following a desired mean trajectory. The proposed solution deviates from the classical linear quadratic tracking formulation or a typical mean field rendez-vous problem (see [9]) which would have each element track the desired mean temperature thus introducing unnecessary control actions. Instead, our proposed solution enforces collective mean temperature tracking while minimizing temperature disturbance on individual devices. The mean field effect is mediated by the *weights* in a commonly shared quadratic cost function. The advantage of a mean field based control approach is that when provided with the mean temperature target trajectory as well as the initial mean temperature in the controlled group, the devices will generate their own control locally, and thus enforce their safety and comfort constraints locally as well. As a result, communication requirements become drastically reduced. Note however, that provision is made that devices would send their temperature readings according to a Poisson process with low intensity. The results are aggregated at the central level thus providing at all times some feedback on the current temperature distribution in the population; this is to avoid a risky mode of operation which would be completely open loop, with prediction errors increasing over time.

After a brief review of the main results of mean field control of linear quadratic Gaussian agents, a diffusion model of elemental space heating/cooling loads is considered. Contributions consist of (i) a formulation of the control problem whereby the objective is for, in general, a heterogeneous group of devices to follow a desired mean target temperature, (ii) the description of our collective target tracking mean field control algorithm, (iii) the corresponding system of mean field equations, (iv) a fixed point analysis, (v) an ϵ-Nash theorem for the population of agents, and numerical simulations for a population of space heaters aimed at illustrating the methodology.

For dynamic large population games where the agents are coupled through their cost functions and dynamics through a state averaging function, the *mean field* framework [6–9] provides decentralized strategies that yield Nash equilibria in the asymptotic limit of an infinite population. The control laws use only the local information of each agent on its own state and own dynamical parameters, while the mass effect is calculated offline using statistical information. These laws yield approximate equilibria when applied in the finite population case, and must be periodically readjusted after long intervals of time using on line aggregate measurements. This is because of prediction errors buildup, due to finite albeit large numbers and potentially non-stationary elemental devices stochastic models.

Using dispersed storage for accommodating renewable sources is a growing area of research. An aggregate model for a large number of pool pumps using mean field limits is developed in [12–14]. Dynamic pricing for controlling the load of aggregates of large commercial buildings is analyzed in [15], and domestic heating systems are employed as heat buffers in [16]. A decentralized charging control strategy for large populations of plug-in electric vehicles (PEVs) using the mean field methodology is presented in [17].

The rest of the paper is organized as follows. In Section 2 we briefly review linear quadratic Gaussian (LQG) mean field theory. In Section 3 we introduce the model that will be used throughout the paper and propose our collective target tracking mean field formulation. In Section 4 we present a fixed point analysis for the equation system characterizing the limiting mean field. In Section 5 we introduce a numerical algorithm that guarantees convergence, and in Section 6 we develop our ϵ-Nash Theorem indicating that an approximate Nash Equilibrium is attained. Lastly, in Section 7, we provide simulation results together with comparisons to a prevailing target tracking control formulation.

The following notation is defined; the set of nonnegative real numbers is denoted by $\mathbb{R}_+$. For vectors $x, y \in \mathbb{R}^n$, we use the notation $x \leq y$ if $x_i \leq y_i$ for all $1 \leq i \leq n$, and $x < y$ if $x_i \leq y_i$ for all $1 \leq i \leq n$, and there exists at least one $i, 1 \leq i \leq n$, such that $x_i < y_i$. The norm $\|\cdot\|$ denotes the 2-norm of vectors and matrices, and $\|x\|_Q^2 \triangleq x^\top Qx$. The set $\mathbf{C}[0,\infty)$ denotes the family of all continuous functions on $[0,\infty)$, $\mathbf{C}_b[0,\infty) = \{x : x \in \mathbf{C}, \sup_{t\geq 0}\|x_t\| < \infty\}$ denotes the family of all bounded continuous functions, and for any $x \in \mathbf{C}_b$, $\|\cdot\|_\infty$ denotes the supremum norm: $\|x\|_\infty \triangleq \sup_{t\geq 0}\|x_t\|$. For the trace operator the notation $\mathrm{tr}(X)$ is employed, and $X^\top$ denotes the transpose of a matrix X.

2 Background on linear quadratic gaussian mean field control

A review of the LQG multi-agent heterogeneous population mean state tracking problem

A large population of N stochastic dynamic agents is considered where agents are stochastically independent, but which shall be cost coupled and such that the individual dynamics are defined by

$$dx_t^i = (A^i x_t^i + B^i u_t^i + c^i)dt + D dw_t^i, \quad t \geq 0, \tag{1}$$

$1 \leq i \leq N$, where for agent $\mathcal{A}_i$, $x^i \in \mathbb{R}^n$ is the state, $u^i \in \mathbb{R}^m$ is the control input; $w^i \in \mathbb{R}^r$ is a standard Wiener process on a sufficiently large underlying probability space $(\Omega, \mathcal{F}, P)$ such that w^i is progressively measurable with respect to $\mathcal{F}^{w^i} \triangleq \{\mathcal{F}_t^{w^i}; t \geq 0\}$. We denote the population average state by $x^{(N)} = (1/N)\sum_{i=1}^N x^i$. As N goes to infinity, asymptotic decoupling of the agents suggests $x^{(N)}$ converging to some a priori unknown trajectory x^*.

The cost function for agent $\mathcal{A}_i, 1 \leq i \leq N$, is given by

$$J_i(u^i, u^{-i}) = \mathbb{E}\int_0^\infty e^{-\delta t}\Big[\|x_t^i - m_t\|_{Q_t}^2 + \|u_t^i\|_R^2\Big]dt, \tag{2}$$

where the cost-coupling is assumed to be in the form of an arbitrary averaging function $m_t \triangleq m(x_t^{(N)} + \eta)$, $\eta \in \mathbb{R}^n$ and δ is a strictly positive discount factor. The term u^i is the control input of the agent $\mathcal{A}_i$ and u^{-i} denotes the control inputs of the complementary set of agents $\mathcal{A}_{-i} = \{\mathcal{A}_j : j \neq i, 1 \leq j \leq N\}$.

Each agent $\mathcal{A}_i, 1 \leq i \leq N$, solves the Riccati equation

$$-\frac{d\Pi_t^i}{dt} = \Pi_t^i\left(A^i - \frac{\delta}{2}I\right) + \left(A^i - \frac{\delta}{2}I\right)^\top \Pi_t^i - \Pi_t^i B^i R^{-1} {B^i}^\top \Pi_t^i + Q_t, \tag{3}$$

for $t \in [0,\infty)$. First define the function class
$\mathbf{C}_{\delta/2}[0,\infty) = \{x : x \in \mathbf{C}[0,\infty), \sup_{t\geq 0}(\|x_t\| e^{-(\delta'/2)t}) < \infty\}$ for some $\delta' \in [0,\delta)$. For a given posited *mass tracking signal* $x^* \in \mathbf{C}_{\delta/2}[0,\infty)$ the mass offset function s^i is generated by the differential equation

$$-\frac{ds_t^i}{dt} = (A^i - \delta I - B^i R^{-1} {B^i}^\top \Pi_t^i)^\top s_t^i + \Pi_t^i c^i - Q_t x_t^*, \tag{4}$$

for $t \in [0,\infty)$.

We introduce the following assumptions.

A1: The processes w_t^i, $t \geq 0$, $1 \leq i \leq N$, are mutually independent and independent of the initial conditions, and $\sup_{i\geq 1}[\mathrm{tr}\Sigma + \mathbb{E}\|x_0^i\|^2] < \infty$, where $\mathbb{E}[w^i {w^i}^\top] = \Sigma$, $1 \leq i \leq N$. □

A2: Γ is a compact set such that for each $\gamma^i \triangleq [A^i, B^i, c^i] \in \Gamma$, (i) the pair $[A^i - (\delta/2)I, B]$ is stabilizable, (ii) $[Q_t^{1/2}, A^i - (\delta/2)I]$, $t \in [0,\infty)$, is detectable, and (iii) $\|Q\|_\infty < \infty$. □

Then, under **A1**-**A2**, (i) there exists a unique solution s^i for (4) in the class $\mathbf{C}_{\delta/2}[0,\infty)$, and (ii) the optimal tracking control law [18] is given by

$$u^{i*}(x_t^i, x_t^*) = -R^{-1}{B^i}^\top(\Pi_t^i x_t^i + s_t^i), \quad t \geq 0, \tag{5}$$

where $u^{i*}(\cdot)$ solves $\inf J_i(u^i, x^*)$, which is defined below by an abuse of notation:

$$\inf J_i(u^{i*}, x^*) \triangleq \inf\left[\mathbb{E}\int_0^\infty e^{-\delta t}[\|x_t^i - x_t^*\|_{Q_t}^2 + \|u_t^i\|_R^2]dt\right].$$

Note that $x^* \in \mathbf{C}_{\delta/2}[0,\infty)$ is assumed to be fixed although unknown. For that x^* to be sustainable, it must be collectively replicated by the agents implementing their best responses to that signal. Thus, system (4)–(5) must be complemented by a fixed point requirement leading to the mean field equation system to be defined later.

We first define the empirical distribution associated with the first N agents: $F_N^\gamma = (1/N)\sum_{i=1}^N \mathbb{I}_{(\gamma^i<\gamma)}$, $\gamma \in \mathbb{R}^{n(n+m+1)}$, where $\{\gamma^i, 1 \leq i \leq N\}$ is a set of random matrices on $(\Omega, \mathcal{F}, P)$ with the common probability distribution F^γ. Then for the basic rendez-vous type MF control problem, the following assumption is adopted.

A3: The cost-coupling is assumed of the form: $m(\cdot) \triangleq m(x^{(N)} + \eta)$, $\eta \in \mathbb{R}^n$, where the function $m(\cdot)$ is Lipschitz continuous on $\mathbb{R}^n$ with a Lipschitz constant $\lambda > 0$; i.e. $\|m(x) - m(y)\| \leq \lambda\|x - y\|$ for all $x, y \in \mathbb{R}^n$. □

Each agent solves the equation system below to calculate the mass tracking signal x_t^*, $t \in [0,\infty)$, offline, for an infinite population of agents.

Definition 2.1 *Mean Field (MF) Equation System on $t \in [0,\infty)$:*

$$\begin{aligned}
-\frac{ds_t^\gamma}{dt} &= (A^\gamma - \delta I - B^\gamma R^{-1}{B^\gamma}^\top \Pi_t^\gamma)^\top s_t^\gamma + \Pi_t^\gamma c^\gamma - Q_t x_t^*,\\
\frac{d\bar{x}_t^\gamma}{dt} &= (A^\gamma - B^\gamma R^{-1}{B^\gamma}^\top \Pi_t^\gamma)\bar{x}_t^\gamma - B^\gamma R^{-1}{B^\gamma}^\top s_t^\gamma + c^\gamma,\\
\bar{x}_t &= \int_\Gamma \bar{x}_t^\gamma dF^\gamma,\\
x_t^* &= m(\bar{x}_t + \eta), \quad t \in [0,\infty).
\end{aligned} \tag{6}$$

□

Let us introduce the following assumption before analysing (6).

A4: $\|Q\|_\infty\|R^{-1}\|\lambda\int_{\gamma\in\Gamma}\|B^\gamma\|^2(\int_0^\infty\|e^{A_*^\gamma(t)t}\|dt)^2 dF^\gamma < 1$, where $A_*^\gamma(t) \triangleq A^\gamma - B^\gamma R^{-1}{B^\gamma}^\top \Pi_t^\gamma$. □

Lemma 2.1 *Under **A1**–**A4** the MF Equation System (6) admits a unique bounded solution.*

The proof is similar to the proof of Theorem 4.4 in [9], and is therefore omitted.

Note in the above, the potential heterogeneity of the agent dynamic parameters is captured by parameter γ considered as a random vector on a compact set (recall **A2** and see [9] for further details).

The Global Observation Control Set $\mathcal{U}_g^N$*:* For the optimality analysis, we first introduce the global observation control set. The set of control inputs $\mathcal{U}_g^N$ consists of all feedback controls adapted to $\mathcal{F}_t^N, t \geq 0$, where $\mathcal{F}_t^N$ is the σ-field generated by the set $\{x_\tau^j : 0 \leq \tau \leq t, 1 \leq j \leq N\}$.

The Local Observation Control Set $\mathcal{U}_{l,i}$*:* The local observation control set of agent $\mathcal{A}_i$ is the set of control inputs $\mathcal{U}_{l,i}$ which consists of the feedback controls adapted to the set $\mathcal{F}_{i,t}, t \geq 0$, where the σ-field $\mathcal{F}_{i,t}$ is generated by $(x_\tau^i;\, 0 \leq \tau \leq t)$.

Theorem 2.2 *MF Stochastic Control Theorem (following [9])*

For systems (1) with cost function (2) let ***A1****–****A4*** *hold;*

(i) the MF equations (6) have a unique solution; which induces a family of decentralized feedback control policies, $\mathcal{U}_{MF}^N \triangleq \{(u^i)^\circ; 1 \leq i \leq N\}$, $1 \leq N < \infty$, *described by (5) such that*

(ii) all agent system trajectories x^i, $1 \leq i \leq N$, *are stable in the sense that*

$$\mathbb{E}\int_0^\infty e^{-\delta t}\|x_t^i\|^2 dt < \infty;$$

(iii) $\{\mathcal{U}_{MF}^N; 1 \leq N < \infty\}$ *yields an* ϵ*-Nash equilibrium for all* $\epsilon > 0$*, i.e., for all* $\epsilon > 0$*, there exists* $N(\epsilon)$ *such that for all* $N \geq N(\epsilon)$

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

Proof.

(i) Property (i) follows from Lemma 2.1.

(ii) Note that the closed loop system is given by

$$dx_t^i = (A^i - B^i R^{-1} {B^i}^\top \Pi_t^i) x_t^i dt - B^i R^{-1} {B^i}^\top s_t^i dt + c^i dt + D dw_t^i, \quad t \geq 0.$$

First of all (a) **A2** ensures that $A^i - B^i R^{-1} {B^i}^\top \Pi_t^i - (\delta/2)I$ is Hurwitz, and (b) **A1** ensures that x_0^i is independent of $w^i(.)$ and $\mathbb{E}\|x_0\|^2 < \infty$. Moreover, $s^i \in \mathbf{C}_{\delta/2}[0,\infty)$; therefore, Lemma A.4 ensures that

$$\mathbb{E}\int_0^\infty e^{-\delta t}\|x_t^i\|^2 dt < \infty.$$

(iii) The proof is similar to the proof of Theorem 5.6 in [9], and is therefore omitted.

□

In essence Theorem 2.2 states that the MF equation system produces a set of decentralized control policies for each agent, which collectively become arbitrarily close in performance to a Nash equilibrium in the space of feedback strategies, provided the number of agents increases sufficiently.

3 Electric space heater models

In the following, we first introduce the model for space heating dynamics that will be employed throughout the paper. Subsequently, a *collective target tracking mean field* control model is defined together with the individual control actions and the corresponding mean field system of equations is developed.

We employ a one dimensional equivalent thermal parameter (ETP) model (see Figure 1) [19] to describe the thermal dynamics of a single household, which is written as

$$dx_t^{in} = \frac{1}{C_a}[-U_a(x_t^{in} - x^{out}) + Q_h(t)]dt + \sigma dw_t, \quad t \geq 0, \tag{7}$$

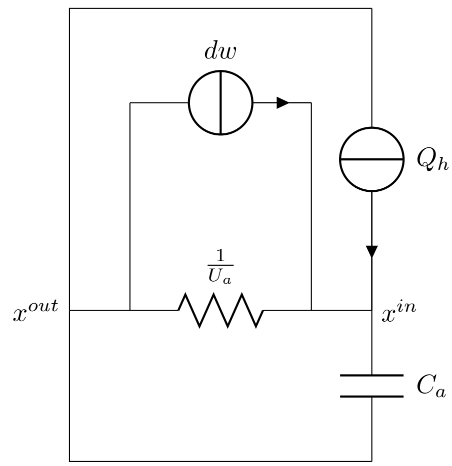

Figure 1: Equivalent thermal parameter (ETP) model of a household

where x^{in} is the air temperature inside the household, x^{out} is the outside ambient temperature, C_a is the thermal mass of air inside, U_a is conductance of the walls and Q_h is the heat flux from the heater. Note that w_t, $t \geq 0$, is a standard Wiener process defined on $(\Omega, \mathcal{F}, P)$ to reflect the noise on the system caused by random processes of heat gain and loss due to customer activity within the dwellings, and σ is its volatility term.

For brevity of notation, the system for heater $\mathcal{A}_i$, $1 \leq i \leq N$, is equivalently written as

$$dx_t^i = [-a^i(x_t^i - x^{out,i}) + b^i u_t^i]dt + \sigma dw_t^i, \quad t \geq 0, \quad 1 \leq i \leq N, \tag{8}$$

where $x^i \triangleq x^{in,i}$ and $u^i \triangleq Q_h^i$ for $1 \leq i \leq N$; $a^i = U_a^i / C_a^i$ and $b^i = {C_a^i}^{-1}$. Note that this model is similar to the model given in [1], where the thermostat control is exchanged with a linear control.

Recalling the implementation principle (iii), we consider that users "naturally" would like their devices to stay at their initial temperatures (normally attained via thermostatic action and before the intervention of the power utility control center). Thus, we have reformulated the control effort as the signal required to make them deviate from that initial temperature; more precisely, the control effort to stay at the initial temperature is considered free and will not be penalized by the cost function to be defined. The corresponding dynamical equation is written as

$$dx_t^i = [-a^i(x_t^i - x^{out,i}) + b^i(u_t^i + u^{free,i})]dt + \sigma dw_t^i, \quad t \geq 0, \tag{9}$$

where $u^{free,i} \triangleq {b^i}^{-1} a^i (x_0^i - x^{out,i})$.

A5: The initial states x_0^i are bounded from above and below by comfort levels; i.e., $l \leq x_0^i \leq h$, $1 \leq i \leq N$. Dynamical parameters a^i, b^i and $x^{out,i}$ together with x_0^i are parameterized by $\theta^i \in \Theta$, i.e., a^{θ^i}, b^{θ^i}, x^{out,θ^i}, $x_0^{\theta^i}$; where $\{\theta^i, 1 \leq i \leq N\}$ is a set of random real numbers on $(\Omega, \mathcal{F}, P)$ with the probability distribution F^θ, reflecting a possibly heterogeneous population of devices. □

3.1 Benchmark: Classical linear quadratic gaussian (LQG) tracking model

Following the results of a global optimization analysis, and for simplicity here, it is assumed that the central authority wants the mean temperature of a particular population to track some *constant* target temperature signal y. In the classical linear quadratic Gaussian (LQG) tracking formulation each agent's cost function is defined as

$$J_i(u^i) = \mathbb{E}\int_0^\infty e^{-\delta t}\left[(x_t^i - y)^2 q + (u_t^i)^2 r\right] dt. \tag{10}$$

The problem with this approach is that each agent minimizes its own cost function and tracks the same signal y. Even though the central authority is only interested in aggregate behaviour and *mean* temperature, this control approach causes *all agents* to track the target signal.

3.2 Collective target tracking mean field model

The currently prevailing mean field LQG control formulation provides the requirements (i) and (ii) in Section 1, but fails short of objective (iii). The reason is that mean field theory is based on a noncooperative dynamic game approach and prevailing cost functions penalize only individual's Euclidean distance from the mean field signal (which could be a convex combination of agents' mean and target trajectory; see [20]) together with the squared norm of the control effort. In this work, instead of each agent trying to track a mean field signal, the individual cost structures are formulated such that ultimately, it is only the mean of the population trajectories that tracks a desired signal. In the proposed method, the novelty is that the mean field effect is mediated by the *quadratic cost function parameters* under the form of an integral error. The resulting concept will be called *collective target tracking mean field*.

We employ the dynamics for the heaters given in (8). The infinite horizon discounted cost function for agent $\mathcal{A}_i, 1 \leq i \leq N$, is defined as follows:

$$J_i(u^i, u^{-i}) = \mathbb{E}\int_0^\infty e^{-\delta t}\left[(x_t^i - z)^2 q_t^y + (x_t^i - x_0^i)^2 q^{x_0} + (u_t^i)^2 r\right] dt, \tag{11}$$

where z, is a direction assigned to each agent in the population and each agent's deviation from this direction is penalized by the deviation penalty coefficient q_t^y, $t \in [0, \infty)$, which captures the mean field information and is calculated as the following *integrated error signal*:

$$q_t^y = \left|\lambda \int_0^t (x_\tau^{(N)} - y) d\tau\right|, \quad t \geq 0, \tag{12}$$

$\lambda > 0$, where y is the main control center dictated mean target constant level, and $x^{(N)} \triangleq (1/N)\sum_{i=1}^N x^i$. Moreover, agents are also penalized with respect to their deviation from their individual initial points, which acts as another layer of heterogeneity for the population.

The justification for the above cost function is that by pointing individual agents towards what is considered as the minimum (or maximum) comfort temperature z, it dictates a global decrease (or increase) in their individual temperatures. This pressure persists as long as the differential between the mean temperature and the mean target y is high. The role of the integral controller is to mechanically compute the *right* level of penalty coefficient $q_t^y, t \in [0, \infty)$, which, in the steady-state, should maintain the mean population temperature at y. When this happens, individual agents reach themselves their steady states (in general different from y and closer to their initial diversified states than classical LQG tracking would dictate).

In order to derive the limiting infinite population MF equation system we start this time assuming a given (albeit initially unknown) cost penalty trajectory $q^y \in \mathbf{C}_b[0, \infty)$ and constant q^{x_0}. Given q^y and q^{x_0}, individual agents $\mathcal{A}_i$, $1 \leq i \leq N$, solve a classical target tracking LQG problem [18] with time varying cost coefficient with Riccati gain π^i and offset term s^i evolving as follows:

$$-\frac{d\pi_t^i}{dt} = (-2a^i - \delta)\pi_t^i - {b^i}^2 r^{-1}(\pi_t^i)^2 + q_t^y + q^{x_0}, \quad t \geq 0, \tag{13}$$

$$-\frac{ds_t^i}{dt} = (-a^i - \delta - {b^i}^2 \pi_t^i r^{-1}) s_t^i + a^i x_0^i \pi_t^i - q_t^y z - q^{x_0} x_0^i, \quad t \geq 0. \tag{14}$$

Then, the optimal tracking control law is given by

$$(u_t^i)^\circ = -b^i r^{-1}(\pi_t^i x_t^i + s_t^i), \quad t \geq 0. \tag{15}$$

The calculation of the unknown q_t^y, $t \geq 0$, is obtained by requiring that q_t^y, $t \geq 0$, be such that when individual agents implement their associated best responses, they must collectively replicate the posited q_t^y, $t \geq 0$, trajectory. This fixed point requirement leads to the specification below of the collective target tracking MF equation system.

Definition 3.1 *Collective Target Tracking (CTT) MF Equation System on* $t \in [0, \infty)$*:*

$$
\begin{aligned}
q_t^y &= \left| \lambda \int_0^t (\bar{x}_\tau - y) d\tau \right|, \\
-\frac{d\pi_t^\theta}{dt} &= (-2a^\theta - \delta)\pi_t^\theta - {b^\theta}^2 r^{-1} {\pi_t^\theta}^2 + q_t^y + q^{x_0}, \\
-\frac{ds_t^\theta}{dt} &= (-a^\theta - \delta - {b^\theta}^2 \pi_t^\theta r^{-1}) s_t^\theta + a^\theta x_0^\theta \pi_t^\theta - q_t^y z - q^{x_0} x_0^\theta, \\
\frac{d\bar{x}_t^\theta}{dt} &= (-a^\theta - {b^\theta}^2 \pi_t^\theta r^{-1}) \bar{x}_t^\theta - {b^\theta}^2 r^{-1} s_t^\theta + a^\theta x_0^\theta, \\
\bar{x}_t &= \int_\Theta \bar{x}_t^\theta dF^\theta.
\end{aligned}
\tag{16}
$$

□

The calculation of the collective target tracking (CTT) MF equation system (16) is performed offline locally by each agent with statistical information F^θ, available at the start of the control horizon. In theory, the control scheme is fully decentralized; i.e., no communication takes place among the controllers throughout the horizon. In practice however, because of the anticipated prediction error accumulations over time, we intend to readjust periodically the control laws over long intervals based on aggregate system measurements (see Figure 2).

Note that the MF Equations for this model are significantly different from (4.6)-(4.9) in [9] or (6) in Section 2. Indeed both systems are amenable to analysis within a linear systems framework and uniqueness of the fixed point is obtained via a reasonably verifiable contraction condition. In contrast, system (16) is fundamentally nonlinear (because of the form of q_t^y, $t \geq 0$) and special arguments have to be developed for its analysis.

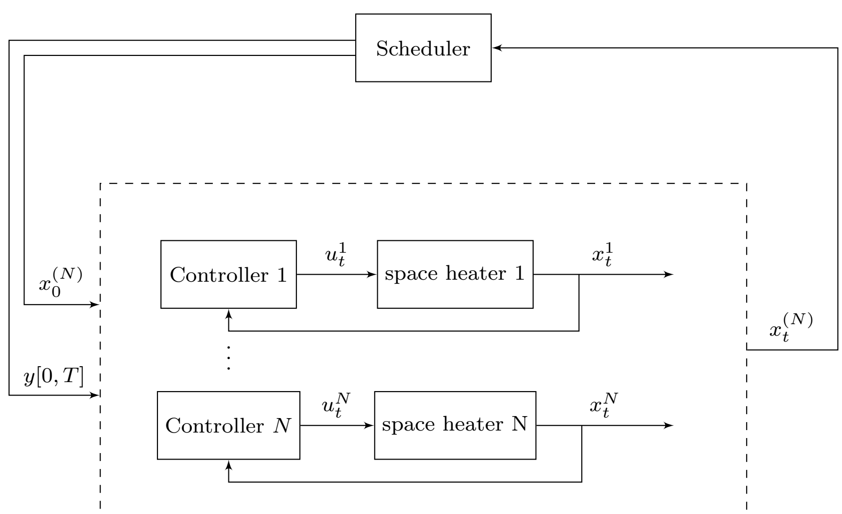

Figure 2: Control architecture in practice

4 Homogeneous case fixed point analysis

Here we present the fixed point analysis for the collective target tracking MF equation system for the particular case $a^i = a$, $b^i = b, 1 \leq i \leq N$, for brevity of notation. We assume that $y \leq \bar{x}_0$, where the target temperature of the central authority is less than or equal to the initial mean temperature of the population. The analysis is very similar for the case $\bar{x}_0 \leq y$; which therefore will be omitted. Note that for $y \leq \bar{x}_0$, the population direction z is set to be less than y; i.e., $z = l \leq y \leq \bar{x}_0 \leq h$, so that the agents *collectively decrease* their temperatures by moving towards that target (see Figure 3).

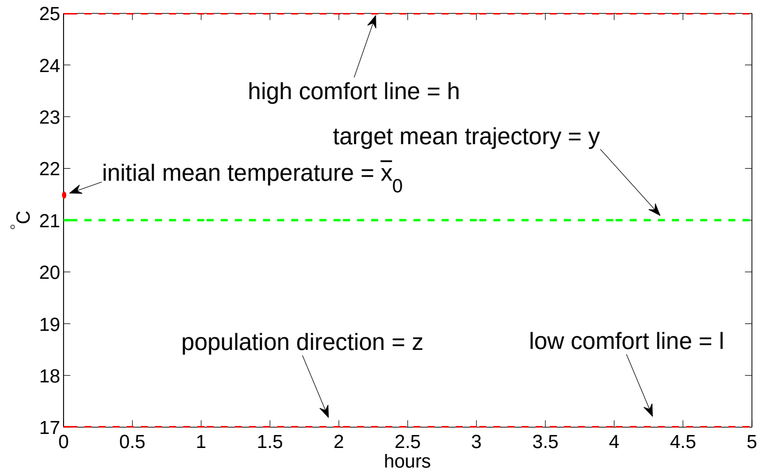

Figure 3: An energy release system ($z \leq y \leq \bar{x}_0$)

The corresponding equation system is given below.

Proposition 4.1 *Collective Target Tracking (CTT) MF Equation System on* $t \in [0, \infty)$*:*

$$\begin{aligned} q_t^y &= \left| \lambda \int_0^t (\bar{x}_\tau - y) d\tau \right|, \\ -\frac{d\pi_t}{dt} &= (-2a - \delta)\pi_t - b^2 r^{-1} \pi_t^2 + q_t^y + q^{x_0}, \\ -\frac{ds_t}{dt} &= (-a - \delta - b^2 \pi_t r^{-1}) s_t + a\bar{x}_0 \pi_t - q_t^y z - q^{x_0} \bar{x}_0, \\ \frac{d\bar{x}_t}{dt} &= (-a - b^2 \pi_t r^{-1})\bar{x}_t - b^2 r^{-1} s_t + a\bar{x}_0, \end{aligned} \tag{17}$$

The proof is given in Appendix B.

4.1 CTT MF equation system

In this section we analyze the existence of a fixed point for CTT MF Equation System (17). We first introduce the operators Δ and $\mathcal{T}$, where their composition $\mathcal{M} \triangleq \mathcal{T} \circ \Delta$, $\mathcal{M}(\bar{x}) : \mathbf{C}_b[0, \infty) \to \mathbf{C}[0, \infty)$, characterizes equation system (17). Then, we define the function set $\mathcal{G} \in \mathbf{C}_b[0, \infty)$ endowed with $\|\cdot\|_\infty$, which includes all the continuous functions of interest in our context. i.e., bounded within the comfort zones and all sharing the same initial point. In order to employ Schauder's fixed point theorem [21] for $\mathcal{M}$ on $\mathcal{G}$, we need to show that (i) $\mathcal{G}$ is a non-empty and closed convex subset of $\mathbf{C}_b[0, \infty)$, (ii) the operator $\mathcal{M} : \mathcal{G} \to \mathcal{G}$ is a compact operator (i.e., the operator is continuous and bounded sets in $\mathcal{G}$ are mapped into sequentially compact sets). We first show that $\mathcal{G}$ is closed in $\mathbf{C}_b[0, \infty)$, and is a non-empty convex set. Then, we show that $\mathcal{M}$'s image set on $\mathcal{G}$ is within $\mathcal{G}$. Next we show that $Im(\mathcal{M})$ is bounded and forms an equicontinuous family of functions, followed by the continuity of $\mathcal{M}$. The latter properties imply that $Im(\mathcal{M})$ is indeed a sequentially compact subset of $\mathcal{G}$ based on the Arzela-Ascoli theorem [21], thus establishing the existence, by Schauder's Theorem, of a fixed point for the operator $\mathcal{M}$ on $\mathcal{G}$.

First we define the operator $\Delta(\bar{x}; \lambda) : \mathbf{C}_b[0, \infty) \to \mathbf{C}[0, \infty)$:

$$\begin{aligned} q_t^y &= \left| \lambda \int_0^t (\bar{x}_\tau - y) d\tau \right| \\ &\triangleq \Delta(\bar{x}; \lambda)(t), \end{aligned} \tag{18}$$

where $\lambda > 0$. Next we define $\mathcal{T} : \mathbf{C}[0,\infty) \to \mathbf{C}[0,\infty)$ for the equation system below with input q_t^y and output $\bar{x}_t$.

$$-\frac{d\pi_t}{dt} = (-2a-\delta)\pi_t - b^2 r^{-1}\pi_t^2 + q_t^y + q^{x_0}, \tag{19}$$

$$-\frac{ds_t}{dt} = (-a-\delta-b^2\pi_t r^{-1})s_t + a\bar{x}_0\pi_t - q_t^y z - q^{x0}\bar{x}_0, \tag{20}$$

$$\frac{d\bar{x}_t}{dt} = (-a-b^2\pi_t r^{-1})\bar{x}_t - b^2 r^{-1} s_t + a\bar{x}_0, \tag{21}$$

which is equivalent to

$$\bar{x}_t \triangleq (\mathcal{T}q)(t). \tag{22}$$

Hence, one can write the MF equation system for CTT as

$$\begin{aligned} \bar{x}_t &= (\mathcal{T}\circ\Delta)(\bar{x})(t) \\ &\triangleq (\mathcal{M}\bar{x})(t), \quad t\in[0,\infty). \end{aligned} \tag{23}$$

Definition 4.1 *Define the set $\mathcal{G}$ endowed with the norm $\|\cdot\|_\infty$ as that including all functions such that $f \in \mathbf{C}_b[0,\infty)$, $f(0) = \bar{x}_0$ and $z \le f(t) \le \bar{x}_0$ for $t \in [0,\infty)$. Note that $\mathcal{G}$ is nonempty.*

Note that the proofs of the following lemmas and propositions are given in Appendix B. Appendix A incorporates other necessary preliminary results and their proofs.

Lemma 4.2 *$\mathcal{G}$ is convex and closed in $\mathbf{C}_b[0,\infty)$ under $\|\cdot\|_\infty$.*

Proposition 4.3 *For $\mathcal{M} : \mathcal{G} \to \mathbf{C}[0,\infty)$ defined in (23) and $\mathcal{G}$ specified in Definition 4.1:*

(i) $Im(\mathcal{M})$ is uniformly bounded,

(ii) $Im(\mathcal{M}) \subset \mathcal{G}$ and forms a family of equicontinuous functions.

Proposition 4.4 *For operator $\mathcal{M} : \mathcal{G} \to \mathbf{C}[0,\infty)$ defined in (23), where $Im(\mathcal{M})$ is shown to be in $\mathcal{G}$ in Proposition 4.3, for any x', $x'' \in \mathcal{G}$ we have*

$$\|\mathcal{M}x' - \mathcal{M}x''\|_\infty \le f(\lambda)\|x' - x''\|_\infty, \tag{24}$$

where λ is given in (18), $f(0) = 0$ and $f(\cdot)$ is a strictly monotonically increasing function of λ.

4.2 Fixed point theorem

Following the lemmas we present our fixed point existence theorem.

Theorem 4.5 *There exists a fixed point for the map $\mathcal{M} : \mathcal{G} \to \mathcal{G}$.*

Proof. It has been shown in Lemma 4.2 that the set $\mathcal{G}$ is non-empty, convex and closed in $\mathbf{C}_b[0,\infty)$. Then it has been shown in Proposition 4.3 that $\mathcal{M}$ is a mapping from $\mathcal{G}$ onto itself, $Im(\mathcal{M})$ is bounded, and $Im(\mathcal{M})$ forms a family of equicontinuous functions. In Proposition 4.4 it has been shown that $\mathcal{M}$ is a continuous operator.

Now take a sequence $\{\bar{x}_k\}_{k\in\mathbb{N}} \in \mathcal{G}$ converging to $\bar{x}^\dagger \in (\mathcal{G}, \|\cdot\|_\infty)$ (note that $\mathcal{G}$ is closed in $\mathbf{C}_b[0,\infty)$ due to Lemma 4.2). Proposition 4.3 implies the uniform boundedness and equicontinuity of $\{\mathcal{M}(\bar{x}_k)\}_{k\in\mathbb{N}}$. By Arzela-Ascoli Theorem there exists a convergent subsequence of $\{\mathcal{M}(\bar{x}_k)\}_{k\in\mathbb{N}}$. Therefore, $Im(\mathcal{M})$ is a sequentially compact subset of $\mathcal{G}$.

Hence, since $\mathcal{G}$ is a closed convex subset of $\mathbf{C}_b[0,\infty)$, $\mathcal{M} : \mathcal{G} \to \mathcal{G}$, and $\mathcal{M}$ is a compact operator, Schauder's Fixed Point Theorem [22] dictates the existence of a fixed point for the map $\mathcal{M} : \mathcal{G} \to \mathcal{G}$. □

From the structure of the problem, it appears fairly obvious that the only possible candidates at steady-state for such an equilibrium are a desirable one ($\bar{x}_\infty = y$), and an undesirable one ($\bar{x}_\infty = z$). The existence of a fixed point for (17) in essence implies the existence of a Nash Equilibrium for an infinite population game. In the equilibrium, the prescribed control actions are the best responses for infinitesimal agents, and there is no unilateral profitable deviation. The existence of an algorithm to reach one such equilibrium by tuning integral control gain parameter λ in (12) will be presented below.

5 Numerical algorithm

5.1 Restricted operator

In order to concentrate on the monotonicity of the trajectories for a system when $z \le y \le \bar{x}_0$, we employ a so-called restricted operator whereby anytime the state trajectory $\bar{x}_t$, $t \ge 0$, hits y, it is frozen there.

Note that $y \le \bar{x}_0$. Define $T_h \in (\mathbb{R}_+ \cup \infty)$ as the first time that $\bar{x}_t \le y$, $t \ge 0$.

Then, define $\mathcal{M}^r \triangleq \mathcal{T}^r \circ \Delta$ where $\mathcal{T}^r : \mathbf{C}[0,\infty) \to \mathbf{C}_b[0,\infty)$:

$$\mathcal{T}^r \triangleq \begin{cases} \mathcal{T}(q) & \text{for } [0, T_h), \\ y & \text{for } [T_h, \infty). \end{cases} \tag{25}$$

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

Proposition 5.1 *If $\mathcal{M}^r(\hat{x}) = \hat{x}$ for some $\hat{x} \in \mathcal{G}^r$ such that $\hat{x} \in \mathbf{C}_b^1[0,\infty)$, then $\hat{x}$ is a fixed point of the operator $\mathcal{M}$ on $\mathcal{G}$; i.e., $\mathcal{M}(\hat{x}) = \hat{x}$.*

Proof. Note that $M^r(f) \equiv M(f)$ for $y \le f(t) \le x_0$, $t \in [0,\infty)$. Since $\hat{x} \in \mathbf{C}_b^1[0,\infty)$, either (i) $T_h = \infty$, i.e., $\hat{x}$ asymptotically converges to y, or (ii) $\hat{x}$ never crosses y. In both cases, the operators act identically. Therefore, $M^r(\hat{x}) = \hat{x}$, $\hat{x} \in \mathcal{G}^r, \hat{x} \in \mathbf{C}_b^1[0,\infty)$ implies $M(\hat{x}) = \hat{x}$. □

Theorem 5.2 *There exists a $\lambda^* > 0$ that guarantees the existence of a unique fixed point for the map $\mathcal{M} : \mathcal{G} \to \mathcal{G}$.*

Proof. Note that $\mathbf{C}_b[0,\infty)$ equipped with $\|\cdot\|_\infty$ is a complete metric space. Since $\mathcal{G}$ is closed in $\mathbf{C}_b[0,\infty)$ due to Lemma 4.2, and a closed subset of a complete metric space is complete, $\mathcal{G}$ equipped with $\|\cdot\|_\infty$ is complete. For $\lambda \in [0, \lambda^*)$ where

$$\lambda^* \triangleq f^{-1}(1), \tag{26}$$

and $f(\cdot)$ is given in (24), $\mathcal{M}$ is a contraction mapping $\mathcal{M} : \mathcal{G} \to \mathcal{G}$. Employing Banach fixed point theorem provides the uniqueness of the fixed point. □

5.2 A numerical algorithm for the restricted operator

In this section we present an algorithm that always finds a fixed point to the operator $\mathcal{M} : \mathcal{G} \to \mathcal{G}$. The algorithm acts on the restricted space $\mathcal{G}^r$, and decreases λ until a fixed point is achieved such that the fixed point $x^* \in \mathcal{G}^r$ also has the property that $x^* \in \mathbf{C}_b^1[0,\infty)$. Once that is achieved, Proposition 5.1 provides that x^* is also a fixed point to the original operator $\mathcal{M}$ given in (23). Note that Theorem 5.2 guarantees the existence of a lower bound for λ, so it is guaranteed that the algorithm is bound to end in a finite number of iterations. There is the (unfortunate) inevitable trade-off that lowering λ in essence lowers the convergence speed of the mean trajectory to the target trajectory.

The numerical algorithm not only tries to find any fixed point, but also tries to find a desirable fixed point; in other words tries to achieve $\bar{x}_\infty = y$. At steady state, the mean field equations are written as

$$(-2a-\delta)\pi_\infty - b^2 r^{-1}\pi_\infty^2 + q_\infty^y + q^{x_0} = 0,$$
$$(-a-\delta-b^2\pi_\infty r^{-1})s_\infty + a\bar{x}_0\pi_\infty - q_\infty^y z - q^{x_0}\bar{x}_0 = 0,$$
$$(-a-b^2\pi_\infty r^{-1})\bar{x}_\infty - b^2 r^{-1} s_\infty + a\bar{x}_0 = 0,$$

where π_∞, s_∞, $\bar{x}_\infty$ denote the steady state values for Riccati, offset and mean field state equations respectively. Solving the equation system for q_∞^y when $\bar{x}_\infty = y$ gives

$$q_\infty^{y\,*} = \frac{[a(a+\delta)r + q^{x_0}b^2]}{b^2}\left(\frac{\bar{x}_0 - y}{y-z}\right), \tag{27}$$

where $q_\infty^{y\,*}$ is the corresponding cost coefficient for the desirable fixed point of the system.

Define $T_{q_\infty} \in (\mathbb{R}_+ \cup \infty)$ as the first time that $q_t^y \geq q_\infty^{y\,*}$, $t \geq 0$.

Then, define $\Delta^r : \mathbf{C}_b[0,\infty) \to \mathbf{C}_b[0,\infty)$:

$$\Delta^r \triangleq \begin{cases} \Delta(x_t;\lambda) & \text{for } [0, T_{q_\infty}), \\ q_\infty^{y\,*} & \text{for } [T_{q_\infty}, \infty). \end{cases}$$

The so-called *Restricted Operator Algorithm* is presented below.

Definition 5.1 *Restricted Operator Algorithm For the iterative algorithm first pick sufficiently small $\epsilon_1 > 0$ and $\epsilon_2 > 0$, where the former will be compared to the infinity norm between two successive iterations as a threshold rule for convergence, and the latter will be used to decrease λ when the iterations for the incumbent λ enter a zone of non-desirable convergence or divergence.*

- $k = 0$
- **while** $\|\bar{x} - \bar{x}^{old}\|_\infty > \epsilon_1$, **do**
  - $\bar{x}^{old(2)} = \bar{x}^{old}$
  - $\bar{x}^{old} = \bar{x}$
  - **if** $\mathrm{mod}(k,2) == 1$ **then**
    - $q^y = \Delta(\bar{x};\lambda)$
    - **if** $q_\infty^y > q_\infty^{y\,*}$ **then**
      $\lambda = \lambda \times \frac{1}{1+\epsilon_2}$
      $k = 0$
      $\bar{x}_t = \bar{x}_0$, $t \in [0,\infty)$
      $q^y = \Delta(\bar{x};\lambda)$
    - **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** $\bar{x}$.

Remark 1 The algorithm basically iterates $\mathcal{M}^r$ on $\mathcal{G}^r$. We initiate the algorithm with $x_t^0 = \bar{x}_0$, $t \geq 0$, and at the end of each iteration the algorithm calculates the sequence (x^i), $i \in \mathbb{Z}$. Note that x^{2k}, $k \in \mathbb{Z}$, denotes the even subsequence whereas x^{2k+1}, $k \in \mathbb{Z}$, denotes the odd subsequence.

Lemma 5.3 *$x^{2k} > x^{2k+1}$ for all $k \in \mathbb{Z}$.*

Proof. Note that we initialize x^0 as $x^0 \triangleq \mathcal{T}(q^{y0})$ where $q_t^{y0} = 0, \forall t \geq 0$. Hence, $x_t^0 = x_0$ for all $t \geq 0$. Then define $q^{y1} \triangleq \Delta(x^0)$, and note that $\|q^{y1}\|_\infty > 0$ since $x^0 > y$. Since $q_t^{y1} \geq q_t^{y0}$ for all $t \geq 0$, and $\|q^{y1}\|_\infty > \|q^{y0}\|_\infty$, $x^1 \triangleq \mathcal{T}(q^{y1})$ is uniformly less than x^0; i.e., $x_t^1 \leq x_t^0$ for all $t \geq 0$, and $\inf_{t\geq 0}(x_t^1) < \inf_{t \geq 0}(x_t^0)$. Hence, $x^1 < x^0$.

Now assume there exists some $\hat{k} \in \mathbb{Z}$ such that $x^{2\hat{k}} \leq x^{2\hat{k}+1}$. By definition, $x^{2\hat{k}} = (\mathcal{T} \circ \Delta)(x^{2\hat{k}-1})$ and $x^{2\hat{k}+1} = (\mathcal{T} \circ \Delta)(x^{2\hat{k}})$. Then, $x^{2\hat{k}} \leq x^{2\hat{k}+1}$ implies $\Delta(x^{2\hat{k}-1}) \geq \Delta(x^{2\hat{k}})$. This implies $x^{2\hat{k}-1} \geq x^{2\hat{k}}$ which also implies $\Delta(x^{2\hat{k}-2}) \geq \Delta(x^{2\hat{k}-1})$, which then again implies $x^{2k-2} \leq x^{2k-1}$. Continuing this iteration gives $x^0 \leq x^1$ which is a contradiction. Therefore the lemma follows. □

Lemma 5.4 *The limits exist for $\lim_{k\to\infty} x^{2k} = x^{(2k)^*}$ and $\lim_{k\to\infty} x^{2k+1} = x^{(2k+1)^*}$ such that*

$$x^{(2k)^*} = x^{(2k+1)^*}.$$

Proof. x^{2k} and x^{2k+1} are both bounded from above by $\bar{x}_0$ and bounded below by y. Since both are monotonic sequences due to Lemma 5.3, the limits exist. For all finite $k \in \mathbb{Z}$, Lemma 5.3 asserts $x^{2k} > x^{2k+1}$; i.e., the sequence x^{2k} is uniformly higher than x^{2k+1} for all $t \geq 0$. Hence, at the limit one has $x^{(2k)^*} \geq x^{(2k+1)^*}$.

Assume $x^{(2k)^*} > x^{(2k+1)^*}$. Since $x^{(2k)^*}$ is the limit, then $x^{(2k)^*} = \mathcal{M}(\mathcal{M}(x^{(2k)^*}))$. On the other hand $\Delta^r(\mathcal{M}(\mathcal{M}(x^{(2k)^*}))) < \Delta(x^{(2k)^*})$ which implies $\mathcal{M}(\mathcal{M}(\mathcal{M}(x^{(2k)^*}))) > \mathcal{M}(x^{(2k)^*})$. However, this implies $x^{(2k+1)^*} > x^{(2k+1)^*}$, which is a contradiction. Hence,

$$x^{(2k)^*} = x^{(2k+1)^*}.$$

□

Theorem 5.5 *Let us define $\hat{q}^y$ to be equal to $\Delta(x^*)$, where $x^* = x^{(2k)^*} = x^{(2k+1)^*}$. Then, $\hat{q}^y_\infty = q^{y*}_\infty$ given in (27).*

Proof. Note that the algorithm employs Δ^r operator on the odd sequence when $q^y_\infty > q^{y*}_\infty$. Therefore, by definition, $\Delta(x^{(2k+1)^*}) \leq q^{y*}_\infty$. Now, if we had $\Delta(x^{(2k+1)^*}) < q^{y*}_\infty$, then $x^{(2k)^*}$ would not cross y; therefore, $\Delta(x^{(2k)^*})$ would tend to ∞, which violates $x^* = x^{(2k)^*} = x^{(2k+1)^*}$. Hence, $\hat{q}^y_\infty = q^{y*}_\infty$ is established. □

We have shown in Theorem 5.5 that the limit of the iterations of the numerical algorithm provides a trajectory whose integral gives a cost function $\hat{q}^y$ that is equal to q^{y*}_∞ asymptotically. The resulting response to such a cost function is a smooth function that asymptotically converges to y as $t \to \infty$. Therefore, the algorithm provides a desirable smooth trajectory.

6 ϵ-Nash theorem

Here we present the main theorem of the paper. More specifically it develops an MF stochastic control law that achieves a Nash equilibrium at the population limit when applied by all agents in the system. Moreover, the control law induces an ϵ-Nash equilibrium property for a finite population.

Theorem 6.1 *Collective Target Tracking MF Stochastic Control Theorem*

*For systems (9) with the cost function (11), let **A1**–**A5** hold.*

(i) For $\lambda \in [0, \lambda^*)$ (26), the CTT MF Equations have a unique solution which induces a family of decentralized feedback control policies $\mathcal{U}^N_{col} \triangleq \{(u^i)^\circ; 1 \leq i \leq N\}$, $1 \leq N < \infty$, generated by (15) such that

(ii) a desirable fixed point is achieved; i.e., the mean trajectory converges to the target trajectory;

(iii) all agent system trajectories x^i, $1 \le i \le N$, *are exponentially bounded in the sense that*

$$\mathbb{E}\int_0^\infty e^{-\delta t}\|x_t^i\|^2 dt < \infty;$$

(iv) $\{\mathcal{U}_{col}^N; 1 \le N < \infty\}$ *yields an* ϵ*-Nash equilibrium in the sense that, for all* $\epsilon > 0$*, there exists* $N(\epsilon)$ *such that for all* $N \ge N(\epsilon)$

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

Proof.

(i) The proof is given in Theorem 5.2.

(ii) The proof is given in Theorem 5.5.

(iii) Note that the closed loop system is given by

$$dx_t^i = (A^i - B^i R^{-1} {B^i}^\top \Pi_t^i) x_t^i dt - B^i R^{-1} {B^i}^\top s_t^i dt + c^i dt + D dw_t^i, \quad t \ge 0.$$

First of all (a) **A2** ensures that $A^i - B^i R^{-1} {B^i}^\top \Pi_t^i - (\delta/2)I$ is Hurwitz, and (b) **A1** ensures that x_0^i is independent of $w^i(.)$ and $\mathbb{E}\|x_0^i\|^2 < \infty$. Moreover, $s^i \in \mathbf{C}_{\delta/2}[0,\infty)$, and hence $\mathbb{E}\int_0^\infty \exp(-\delta t)\|u_t^i\|^2 dt < \infty$. Therefore, Lemma A.4 ensures that

$$\mathbb{E}\int_0^\infty e^{-\delta t}\|x_t^i\|^2 dt < \infty, \quad 1 \le i \le N.$$

(iv) The proof is similar to the proof of Theorem 5.6 in [9], and is therefore omitted.

□

7 Simulations

For our numerical experiments we simulate a population of 200 space heaters. We take a uniform population of heaters; adopt a one layer ETP model given in (7), where the capacitance (C_a) and conductance (U_a) parameters are chosen to be 10 kWh/°C and 0.2 kW/°C respectively, the ambient temperature is set to -10°C, and the volatility parameter is set to 0.25°C/$\sqrt{h}$. The initial temperatures of the heaters are drawn from a Gaussian distribution with a mean of 21°C and a variance of 1. The cost function parameters δ, q^{x_0} and r are uniformly chosen to be 0.001, 200 and 1 respectively.

For the first simulation the central authority provides the target temperature trajectory (20°C) to each controller, and local controllers apply a classical LQG tracking algorithm. Figure 4 shows that as a result each agent in the population tracks the target degree of 20°C. Notice that in this implementation *all* agents track 20°C in order for the *population mean temperature* to track 20°C, where all agents are heavily disturbed for the global goal.

In the second scenario the central authority sets the target temperature to 20°C (y parameter), and all agents are assigned to track 17°C (z parameter). The $\bar{x}$ based Picard iterations of the collective target tracking MF equation system without restriction are provided in Figure 5 with $\bar{x}_t^0 = \bar{x}_0$, $t \in [0,\infty)$. One immediately notices the monotonic behaviour in the early stages of the trajectories before any of the curves encounters the y target line; the so-called restricted operator algorithm is introduced to take advantage of that monotonicity. In Figure 6 we plot the iterations of the restricted algorithm. Note that the even curves are frozen at the point y when the mean trajectory reaches y. The algorithm converges to the same mean field trajectory as in the previous Figure 5. We provide the corresponding cost function trajectory given by $q^{y*} \triangleq \Delta(\bar{x}^*)$ in Figure 7, where $\bar{x}^*$ is the fixed point to the mean field equation system. The resulting controlled trajectories are presented and the mean temperature trajectory is shown in Figure 8. It can be seen that while the mean temperature still settles at 20°C, individuals in the population are disturbed much less than in the LQG tracking implementation.

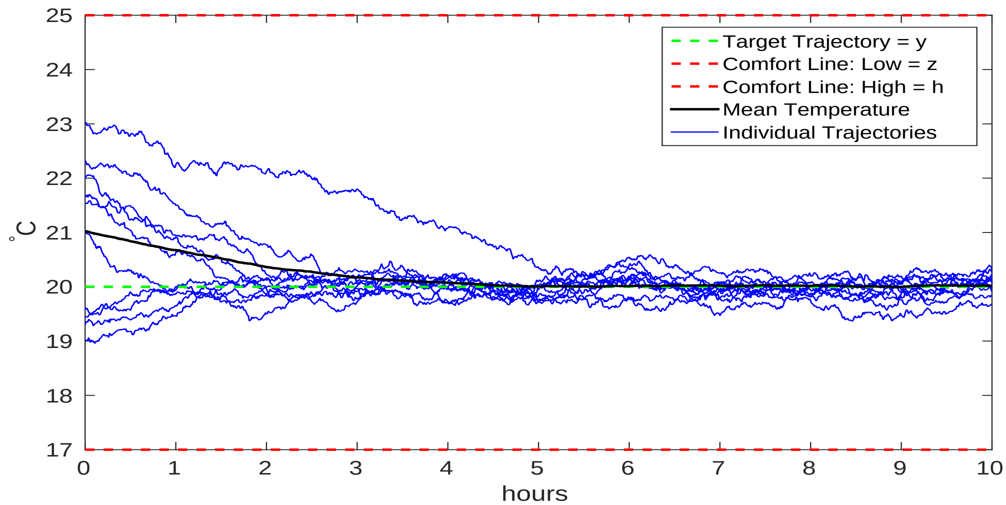

Figure 4: Agents applying classical LQG tracking

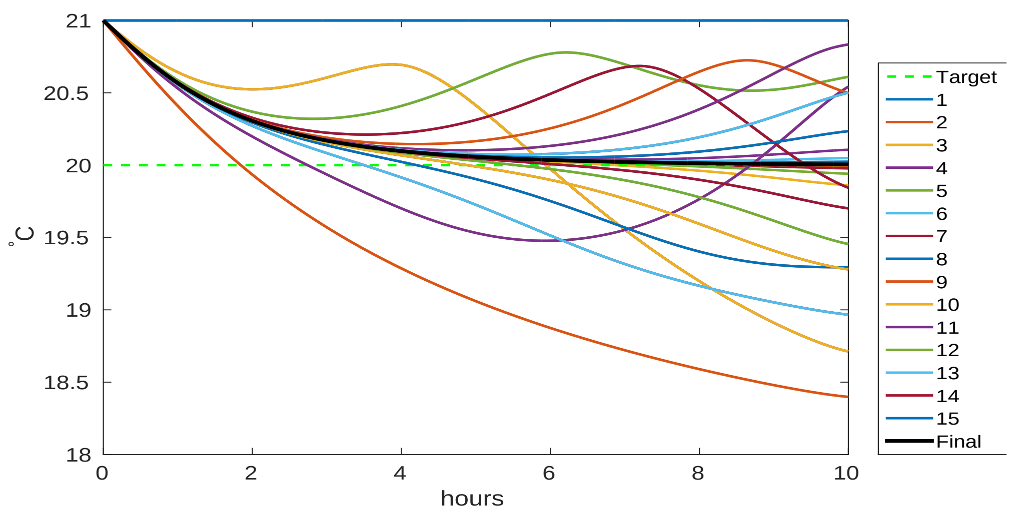

Figure 5: Collective target tracking MF iterations

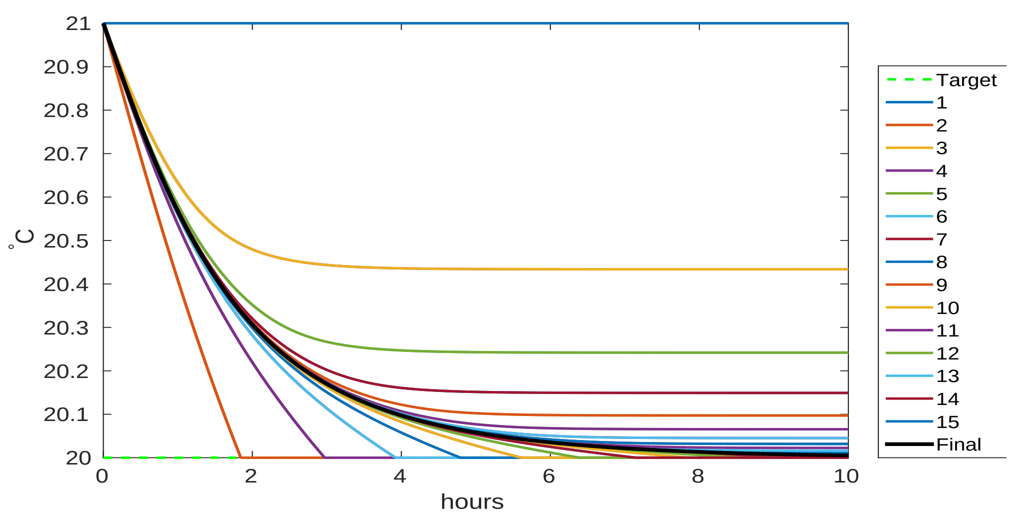

Figure 6: Collective target tracking MF iterations: Restricted algorithm

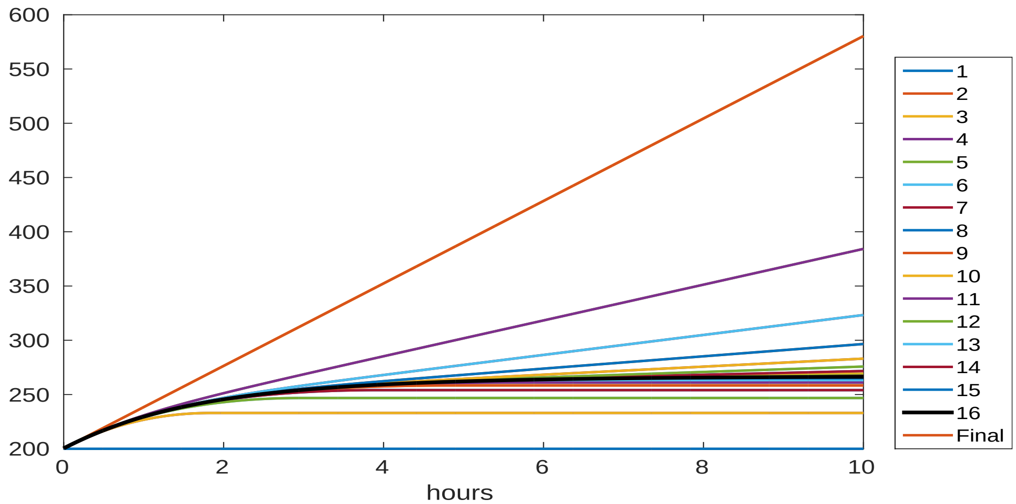

Figure 7: Collective target tracking: Iterations for cost function

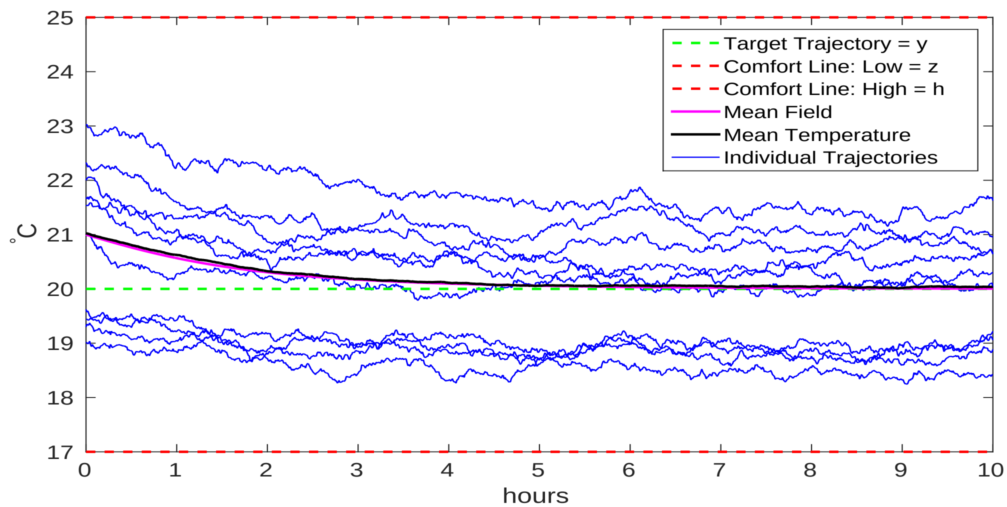

Figure 8: Agents applying collective target tracking MF control: All agents following the low comfort level

In practical implementation, in order to increase the convergence speed of the mean field trajectory to the steady state, we implement an *accelerated engineering solution*: the agents calculate their individualized steady states (which contribute to the global mean field steady state), and employ *no control* until their temperature reaches their individual steady state (in the case of a desired increase in mean temperature, the individuals would instead apply maximum control to reach their own calculated steady-state). This calculation can be carried out as follows. One writes (27) for individual agent $\mathcal{A}_i, 1 \leq i \leq N$, as

$$q_\infty^{y\,*} = \frac{[a(a+\delta)r + q^{x_0}b^2]}{b^2}\left(\frac{x_0^i - (x_\infty^i)^*}{(x_\infty^i)^* - z}\right). \tag{28}$$

Solving (28) for $(x_\infty^i)^*$ yields

$$(x_\infty^i)^* = x_0^i - \frac{b^2(q_\infty)^*(x_0^i - z)}{ar(a+\delta) + (q_\infty)^* b^2 + q^{x_0}b^2}.$$

For the same set of parameters, the central authority sets the target temperature to 20°C. The resulting controlled trajectories are presented and the mean temperature trajectory is shown in Figure 9. It can be

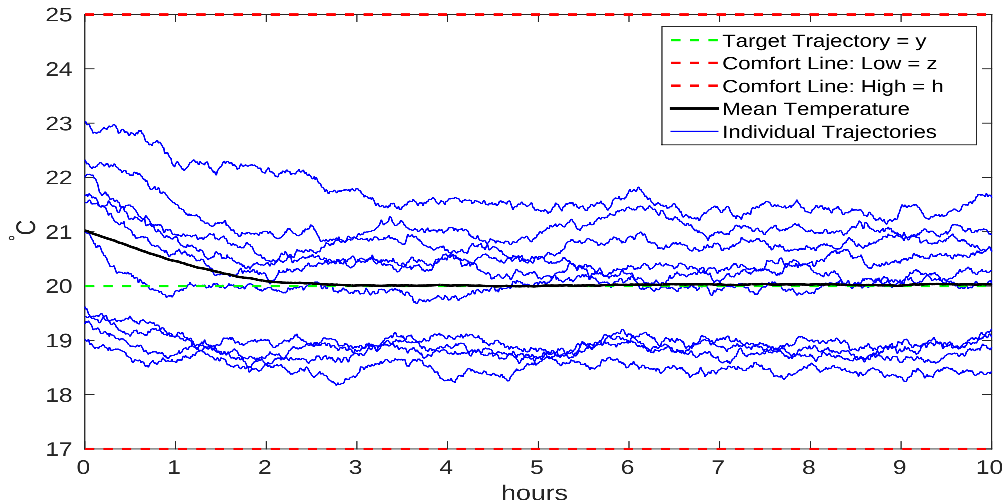

Figure 9: Agents applying collective target tracking MF control: Accelerated engineering solution

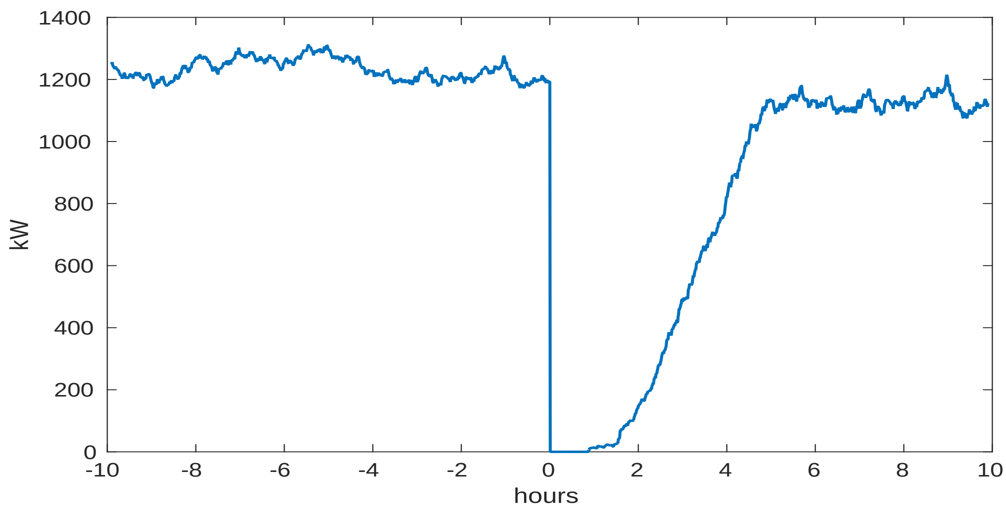

Figure 10: Agents applying collective target tracking MF control: All agents following the low comfort level signal

seen that the mean temperature settles at 20°C in about two and a half hours, a rate noticeably faster than the convergence rate of the mean field dictated by collective target tracking MF control.

For the same set of parameters of the previous simulation, this time the target is set to 19°C. The corresponding aggregate power consumption plot is provided in Figure 10. Not only does the control algorithm provide immense relief at the early stages of the horizon, but it also provides a smooth transition to the steady state power consumption profile without a delayed payback peak typical of direct control schemes of thermostats. Notice on the other hand that while thermostats were eliminated from the original formulation, the final controls correspond to simply changing the set points of individual thermostats from their initial values to the desired individualized steady-state values and could be implemented in this manner.

For the next experiment we separate the population in two groups where the first group consists of the heaters above the 21°C initial population mean temperature, and the second group consists of the ones below that temperature. The central authority sets the target temperature to 22°C, and both groups are assigned to track 25°C. In order to achieve a level of fairness among the agents, the first group is assigned a higher control penalty coefficient r thus making it more reluctant to change its initial temperature. Collective target tracking MF control is applied to these groups together, however using the *max control until the individual* *steady state option* which has been described above. The simulation result is provided in Figure 11. It can be seen that the collective target tracking MF control scheme leads the mean temperature of the global heaters population to 22°C while soliciting the agents with lower initial temperatures more intensely than those with higher initial temperatures.This illustrates how one could shape the collective tracking response for greater fairness. The mean trajectories of the sub-populations with this scheme compared to the case when sub-populations share the same control penalty coefficients is given in Figure 12.

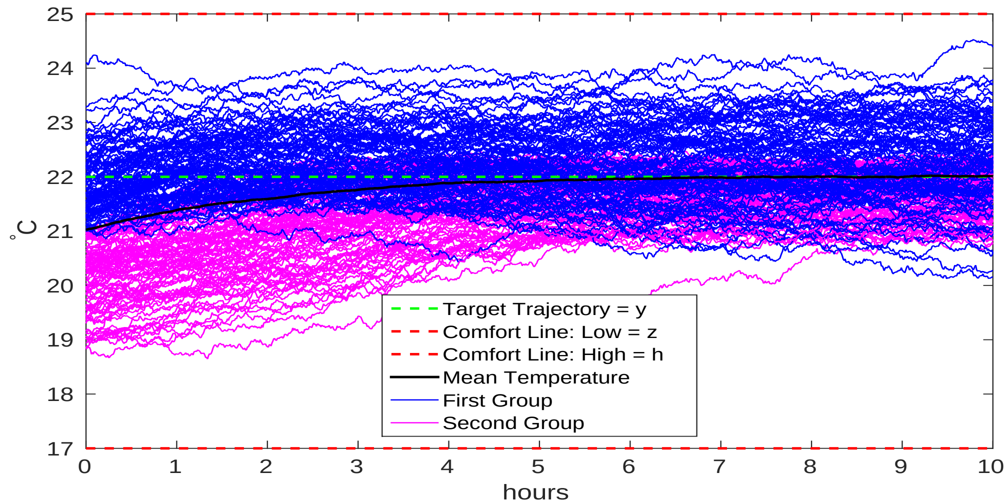

Figure 11: Agents applying collective target tracking MF control: Different r for subpopulations. The blue heated spaces are initially warmer and are thus asked for a lower contribution to the overall energy reduction effort

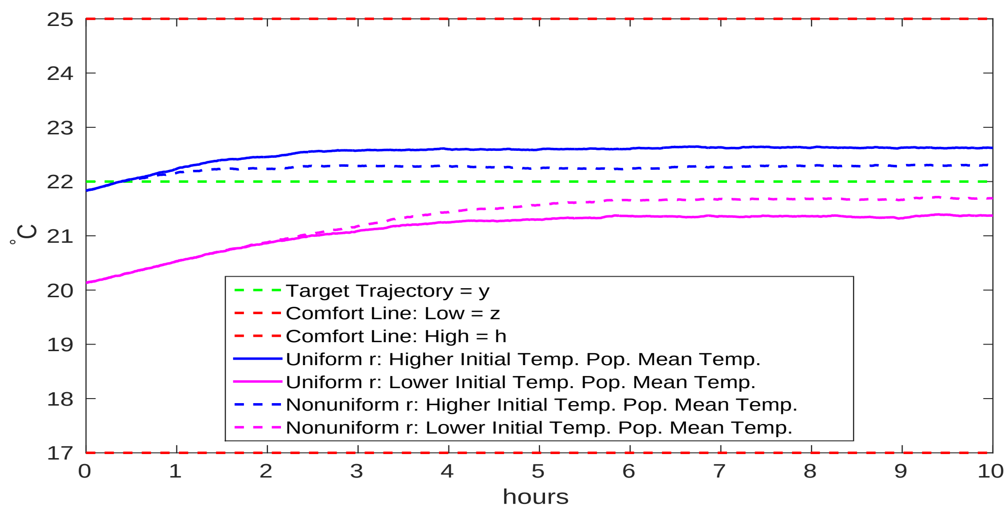

Figure 12: The mean trajectories of the subpopulations in Figure 11

8 Concluding remarks

In this paper the presence of large numbers of electric devices associated with energy storage is employed to develop a decentralized mean field control based approach to the problem of these devices following a desired mean trajectory. Provided with the mean temperature target trajectory as well as the initial mean temperature in the controlled group, the devices generate their own control locally, and thus enforce their safety and comfort constraints locally as well. The proposed solution deviates from the classical formulation which would have each element track the desired mean temperature thus introducing unnecessary control actions. The solution made possible by mean field theory enforces collective mean temperature tracking while leaving individual devices freer to remain, if possible within their comfort zone. In the proposed method, the 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 currently prevailing mean field control formulations where the mean field effect is concentrated on the tracking signal. The equations that provide the best response action are presented together with the MF equation system; a fixed point analysis is given and the existence of a solution is shown. A numerical algorithm for the calculation of a desirable smooth fixed point is provided, and finally an ϵ-Nash Theorem is presented. Numerical simulation results are provided illustrating the flexibility and potential of this form of decentralized collective control.

In future work, the analysis of a cooperative framework and nondiffusion load dynamics which involve jump Markov models (such as in electric water heaters) will be studied. Also, the extension to time varying target tracking problems and the impact of constraints on the synthesis of control laws are subjects of future study. Finally, note that higher dimensional elemental models such as presented in [23] can be easily accommodated in the current theory.

Appendix A Preliminary results

We declare the following lemmas.

Lemma A.1 *For $\mathcal{M} : \mathcal{G} \to \mathbf{C}[0,\infty)$ in (23), the cost function q_t^y, $t \in [0,\infty)$, in (12) and the associated Riccati equation $\pi_t, t \in [0,\infty)$, in (13) have the following properties:*

(i) $q_t^y \leq \lambda \max[y - z, \bar{x}_0 - y]t$, $t \geq 0$,

(ii) π_t *is bounded above by* π^u, *and below by* π^b, *where* $\pi^b > 0$ *is a constant, and* π^u *is monotonically increasing in time and strictly concave,*

(iii) $\pi(q')(t) - \pi(q'')(t) \leq \lambda(\kappa_1 t + \kappa_1^2)\|\bar{x}' - \bar{x}''\|_\infty$, *where* q' *and* q'' *are given by (18) respectively for* $\bar{x}'$ *and* $\bar{x}''$ *in* $\mathcal{G}$.

Proof.

(i) Since $\bar{x} \in \mathcal{G}$, and q^y is given in (18), (i) immediately follows.

(ii) Since q^y is upper bounded by $(q^y)^u = \lambda \max[y - z, \bar{x}_0 - y]t$, then the corresponding $\pi(q^y)$ is upper bounded by the corresponding $\pi^u((q^y)^u)$. Also, $\pi(q^y)$ is bounded below by $\pi^l((q^y)^l)$ where $(q^y)^l(t) = 0$, $t \in [0,\infty)$.

To show the monotonicity properties of $(q^y)^u$ we use a similar technique to the one used in [24, Lemma 10.2]. Now, for any $t \in [0,\infty)$, we have

$$-\frac{d\pi_t^u}{dt} = -(2a+\delta)\pi_t^u - b^2 r^{-1}(\pi_t^u)^2 + q^{x_0} + \lambda \max[y - z, \bar{x}_0 - y]t, \quad t \geq 0. \tag{29}$$

Taking the derivative gives

$$\begin{aligned} -\frac{d^2\pi_t^u}{dt^2} &= -(2a+\delta)\frac{d\pi_t^u}{dt} - b^2 r^{-1} 2\pi_t^u \frac{d\pi_t^u}{dt} + \lambda \max[y - z, \bar{x}_0 - y], \\ &= \left[-(2a+\delta) - 2b^2 r^{-1}\pi_t^u\right] \frac{d\pi_t^u}{dt} + \lambda \max[y - z, \bar{x}_0 - y], \quad t \geq 0. \end{aligned} \tag{30}$$

Now, the solution is given by

$$\frac{d\pi_t^u}{dt} = -\int_\infty^t \Gamma(t,\tau)\lambda \max[y - z, \bar{x}_0 - y]d\tau,$$

where

$$\frac{\partial \Gamma(t,\tau)}{\partial t} = 2a + \delta + 2b^2 r^{-1}\pi_t^u, \tag{31}$$

which implies $\frac{d\pi_t^u}{dt} > 0$ for all $t \in [0,\infty)$.
Similarly take the derivative of (30) to obtain

$$-\frac{d^3\pi_t^u}{dt^3} = -(2a+\delta)\frac{d^2\pi_t^u}{dt^2} - 2b^2 r^{-1}\pi_t^u \frac{d^2\pi_t^u}{dt^2} - 2b^2 r^{-1}\left(\frac{d\pi_t^u}{dt}\right)^2, \tag{32}$$

$$= \left[-(2a+\delta) - 2b^2 r^{-1}\pi_t^u\right]\frac{d^2\pi_t^u}{dt^2} - 2b^2 r^{-1}\left(\frac{d\pi_t^u}{dt}\right)^2, \quad t \geq 0. \tag{33}$$

The solution is given by

$$\frac{d^2\pi_t^u}{dt^2} = -\int_\infty^t \Gamma(t,\tau)\left[-2b^2 r^{-1}\left(\frac{d\pi_t^u}{d\tau}\right)^2\right]d\tau,$$

where $\Gamma(t,\tau)$ is given in (31), which implies $\frac{d^2\pi_t^u}{dt^2} < 0$ for all $t \in [0,\infty)$.
Hence, π^u is monotonically increasing in time and strictly concave.

(iii) First define $\tilde{q}_t \triangleq q_t' - q_t''$ and $\tilde{\pi}_t \triangleq \pi(q')(t) - \pi(q'')(t)$. Then, (29) gives

$$-\frac{d\tilde{\pi}_t}{dt} = \left[-(2a+\delta) - \frac{b^2}{r}2\pi(q'')(t)\right]\tilde{\pi}_t - b^2 r^{-1}\tilde{\pi}_t^2 + \tilde{q}_t, \quad t \geq 0. \tag{34}$$

where $\tilde{\pi}_t$ is the solution to the Riccati equation (34). Hence, for $\tilde{q}_t \geq 0$, $t \in [0,\infty)$, $\tilde{\pi}_t \geq 0, t \in [0,\infty)$, follows.
Since $\pi(q'')(t) \geq 0$, $t \in [0,\infty)$, and $\tilde{\pi}_t^2 \geq 0$, $t \in [0,\infty)$, we get

$$-\frac{d\tilde{\pi}_t}{dt} + (2a+\delta)\tilde{\pi}_t \leq \tilde{q}_t, \quad t \geq 0.$$

Since $\tilde{q}_t \leq \lambda t\|\bar{x}' - \bar{x}''\|_\infty$ we can write

$$-\frac{d\tilde{\pi}_t}{dt} + (2a+\delta)\tilde{\pi}_t \leq \lambda t\|\bar{x}' - \bar{x}''\|_\infty, \quad t \geq 0. \tag{35}$$

Taking derivative of both sides gives

$$-\frac{d^2\tilde{\pi}_t}{dt^2} + (2a+\delta)\frac{d\tilde{\pi}_t}{dt} \leq \lambda\|\bar{x}' - \bar{x}''\|_\infty, \quad t \geq 0.$$

It is shown in (ii) that $\tilde{\pi}$ is concave. Therefore,

$$\frac{d\tilde{\pi}_t}{dt} \leq \frac{\lambda}{2a+\delta}\|\bar{x}' - \bar{x}''\|_\infty, \quad t \geq 0. \tag{36}$$

Then, (35) implies

$$(2a+\delta)\tilde{\pi}_t \leq \frac{d\tilde{\pi}_t}{dt} + \lambda t\|\bar{x}' - \bar{x}''\|_\infty, \quad t \geq 0.$$

Employing (36) gives

$$\tilde{\pi}_t \leq \frac{1}{2a+\delta}\left(\lambda t + \frac{\lambda}{2a+\delta}\right)\|\bar{x}' - \bar{x}''\|_\infty,$$

$$\triangleq \lambda(\kappa_1 t + \kappa_1^2)\|\bar{x}' - \bar{x}''\|_\infty.$$

□

Lemma A.2 *For a system with the transition matrix*

$$\frac{\partial \Phi(t,s)}{\partial t} = (-\alpha - \beta t^{\gamma})\Phi(t,s), \tag{37}$$

for $\alpha \geq 0$, $\beta \geq 0$, $0 < \gamma \leq 1$, the following assertions hold:

1) $\sup_{t\geq 0} \int_0^t \Phi(t,\tau)d\tau \leq \frac{1}{\alpha}$,

2) $\sup_{t\geq 0} t \int_0^t \Phi(t,\tau)d\tau < \kappa_2$,

3) $\Phi(t,\tau) - \Phi(t+\Delta t,\tau) \leq \kappa_3 \Delta t$ *as* $\Delta t \to 0$,

4) $\int_0^t (\Phi(t,\tau) - \Phi(t+\Delta t,\tau))\, d\tau \leq \kappa_4 \Delta t$ *as* $\Delta t \to 0$.

Proof. When needed, consider in the following $\alpha = 0$ and $\gamma = 1$ without loss of generality. Then,

1) $\sup_{t\geq 0} \int_0^t \Phi(t,\tau)d\tau \leq \frac{1}{\alpha}$.
Since $\Phi(t,\tau) \leq \exp[-\alpha(t-\tau)]$, we get the assertion.

2) $\sup_{t\geq 0} t \int_0^t \Phi(t,\tau)d\tau < \kappa_2$.
First write

$$\begin{aligned} t\int_0^t \Phi(t,\tau)d\tau &= t\int_0^t \exp\left(\int_\tau^t -\beta s ds\right) d\tau \\ &= t\int_0^t \exp\left(\frac{-\beta}{2}(t^2-\tau^2)\right) d\tau. \end{aligned}$$

Now, [25, 7.8.7] implies

$$\begin{aligned} t\int_0^t \Phi(t,\tau)d\tau &< \frac{2}{3\beta}\left(2 + \frac{\beta}{2}\exp\left(\frac{-\beta}{2}t^2\right)t^2 - 2\exp\left(\frac{-\beta}{2}t^2\right)\right) \\ &< \frac{2}{3\beta}\left(2 + \frac{\beta}{2}\exp\left(\frac{-\beta}{2}t^2\right)t^2\right). \end{aligned}$$

Take the derivative of $\exp\left(\frac{-\beta}{2}t^2\right)t^2$ to obtain $\exp\left(\frac{-\beta}{2}t^2\right)\left(1-\frac{\beta}{2}t^2\right)$, which is positive for $t \in [0,(2/\beta)^{0.5})$ and negative for $t > (2/\beta)^{0.5}$; hence achieves maximum at $t = (2/\beta)^{0.5}$.
Therefore,

$$\begin{aligned} \sup_{t\geq 0} t\int_0^t \Phi(t,\tau)d\tau &< \frac{2\exp(-1)+4}{3\beta} \\ &\triangleq \kappa_2. \end{aligned}$$

3) $\Phi(t,\tau) - \Phi(t+\Delta t,\tau) \leq \kappa_3 \Delta t$ as $\Delta t \to 0$.
First write

$$\begin{aligned} \Phi(t,\tau) - \Phi(t+\Delta t,\tau) &= \Phi(t,\tau)(1-\Phi(t+\Delta t,t) \\ &= \exp\left[\frac{-\beta}{2}(t^2-\tau^2)\right]\left(1-\exp\left[\frac{-\beta}{2}(2t+\Delta t)\Delta t\right]\right). \end{aligned}$$

Then, again without loss of generality take $\tau = 0$, $\Delta t << 1$ and $\Delta t << t$. Then,

$$\begin{aligned}\Phi(t,\tau)-\Phi(t+\Delta t,\tau) &\leq \exp\left(\frac{-\beta}{2}t^2\right)\frac{\beta}{2}(2t+1)\Delta t\\ &= \exp\left(\frac{-\beta}{2}t^2\right)\beta t\Delta t+\exp\left(\frac{-\beta}{2}t^2\right)\frac{\beta}{2}\Delta t\\ &\leq \kappa_3\Delta t.\end{aligned}$$

4) $\int_0^t (\Phi(t,\tau)-\Phi(t+\Delta t,\tau))d\tau \leq \kappa_4\Delta t$ as $\Delta t\to 0$.
We write

$$\int_0^t (\Phi(t,\tau)-\Phi(t+\Delta t,\tau))d\tau = (1-\Phi(t+\Delta t,t))\int_0^t \Phi(t,\tau)d\tau.$$

Take $\alpha=0$ and $\gamma=1$ without loss of generality. Then,

$$\begin{aligned}\int_0^t (\Phi(t,\tau)-\Phi(t+\Delta t,\tau))d\tau &= \left(1-\exp\left(\int_t^{t+\Delta t}(-\beta\tau)d\tau\right)\right)\int_0^t \Phi(t,\tau)d\tau\\ &= \left(1-\exp\left[-\frac{\beta}{2}(\Delta t(2t+\Delta t))\right]\right)\int_0^t \Phi(t,\tau)d\tau.\end{aligned}$$

Take $\Delta t << t$ and $\Delta t << 1$. Then,

$$\int_0^t (\Phi(t,\tau)-\Phi(t+\Delta t,\tau))d\tau \leq \frac{\beta}{2}\Delta t(2t+1)\int_0^t \Phi(t,\tau)d\tau.$$

With the assumption $\Delta t << t$ and $\Delta t << 1$, one gets

$$\begin{aligned}\int_0^t (\Phi(t,\tau)-\Phi(t+\Delta t,\tau))d\tau &\leq \frac{\beta}{2}\Delta t(2t+1)\int_0^t \Phi(t,\tau)d\tau\\ &= \Delta t\beta t\int_0^t \Phi(t,\tau)d\tau+\Delta t\frac{\beta}{2}\int_0^t \Phi(t,\tau)d\tau\\ &\leq \Delta t\beta\sup_{t\geq 0}\left\{t\int_0^t \Phi(t,\tau)d\tau\right\}+\Delta t\frac{\beta}{2}\sup_{t\geq 0}\left\{\int_0^t \Phi(t,\tau)d\tau\right\}\\ &= \beta I_1\Delta t+\frac{\beta}{2}I_2\Delta t.\end{aligned}$$

We have shown in 2) that $I_1\leq\kappa_2$ and in 1) that $I_2\leq 1/\alpha$. Therefore, the assertion holds. □

Lemma A.3 *For a system with the transition matrix*

$$\frac{\partial\Psi(t,s)}{\partial t}\triangleq(\alpha+\beta t^{\gamma})\Psi(t,s),$$

where $\alpha\geq 0$, $\beta\geq 0$, $\gamma>0$, we get

$$\int_{t_1}^{t_2}\Psi(t_2,\tau)(\alpha+\beta\tau^{\gamma})d\tau=\exp[\alpha(t_2-t_1)]\exp\left[\frac{\beta}{\gamma+1}(t_2^{\gamma+1}-t_1^{\gamma+1})\right]-1,$$

which implies

$$\int_{\infty}^{t_2}\Psi(t_2,\tau)(\alpha+\beta\tau^{\gamma})d\tau=-1.$$

The proof is trivial, and is therefore omitted.

Lemma A.4 *For a system with the dynamics*

$$dx_t=(Ax_t+Bu_t+c_t)dt+Ddw_t,\quad t\geq 0,$$

$a>0$ where (i) $A-(\delta/2)I$ is Hurwitz, (ii) x_0 is independent of $w(\cdot)$, (iii) $\mathbb{E}\|x_0\|^2<\infty$, (iv) $c\in\mathbf{C}_{\delta/2}[0,\infty)$, (v) $\mathbb{E}\int_0^{\infty}\exp(-\delta t)\|u_t\|^2dt\leq c_1$, the following holds

$$\mathbb{E}\int_0^{\infty}\exp(-\delta t)\|x_t\|^2dt\leq c_2.$$

Proof. Define $x'_t = \exp[-(\delta/2)t]x_t$ and $u'_t = \exp[-(\delta/2)t]u_t$. The SDE for x' is given by

$$dx'_t = [A - (\delta/2)I]x'_t dt + Bu'_t dt + \exp[-(\delta/2)t]c_t dt + \exp[-(\delta/2)t]Ddw_t.$$

Taking the integral gives

$$\mathbb{E}\int_0^\infty \|x'_t\|^2 dt \le C + \mathbb{E}\int_0^\infty \left\|\int_0^t \exp\left[(A-(\delta/2)I)(t-\tau)\right]Bu'_\tau d\tau\right\|^2 dt.$$

Since $A - (\delta/2)I$ is Hurwitz, there exists a $\rho > 0$ such that

$$\mathbb{E}\int_0^\infty \|x'_t\|^2 dt \le C + \|B\|^2\mathbb{E}\int_0^\infty \left\|\int_0^t \exp[-\rho(t-\tau)]u'_\tau d\tau\right\|^2 dt.$$

Now employ Cauchy-Schwarz Inequality to obtain

$$\begin{aligned}
&\le C + \|B\|^2\mathbb{E}\int_0^\infty \left(\int_0^t \exp[-\rho(t-\tau)]d\tau\right)\left(\int_0^t \exp[-\rho(t-\tau)]\|u'_\tau\|^2 d\tau\right)dt \\
&= C + (\|B\|^2/\rho)\mathbb{E}\int_0^\infty \left(\int_0^t \exp[-\rho(t-\tau)]\|u'_\tau\|^2 d\tau\right)dt.
\end{aligned}$$

Using Tonelli's theorem, a change-of-order of integration gives

$$\begin{aligned}
&= C + (\|B\|^2/\rho)\mathbb{E}\int_0^\infty \exp(\rho\tau)\|u'_\tau\|^2\left(\int_\tau^\infty \exp(-\rho t)dt\right)d\tau \\
&\le C + (\|B\|^2/\rho^2)\mathbb{E}\int_0^\infty \|u'_\tau\|^2 d\tau,
\end{aligned}$$

which implies

$$\begin{aligned}
\mathbb{E}\int_0^\infty \exp(-\delta t)\|x_t\|^2 dt &\le C + (\|B\|^2/\rho^2)\mathbb{E}\int_0^\infty \exp(-\delta\tau)\|u_\tau\|^2 d\tau \\
&= C + \frac{\|B\|^2}{\rho^2}c_1 \\
&= c_2.
\end{aligned}$$

□

Appendix B Proofs

Proof of Proposition 4.1. Employing **A5**, for systems (8) where the individual controllers are generated according to (15), the fixed point equation system is given by

$$\begin{aligned}
-\frac{d\pi_t}{dt} &= (-2a-\delta)\pi_t - b^2 r^{-1}\pi_t^2 + q_t^y + q^{x_0}, \\
-\frac{ds_t^\theta}{dt} &= (-a-\delta-b^2\pi_t r^{-1})s_t^\theta + a\bar{x}_0^\theta\pi_t - q_t^y z - q^{x_0}x_0^\theta, \qquad (38)\\
\frac{d\bar{x}_t^\theta}{dt} &= (-a-b^2\pi_t r^{-1})\bar{x}_t^\theta - b^2 r^{-1}s_t^\theta + a\bar{x}_0^\theta, \\
\bar{x}_t &= \int_\Theta \bar{x}_t^\theta dF^\theta, \qquad (39)\\
q_t^y &= \left|\lambda\int_0^t (\bar{x}_\tau - y)d\tau\right|.
\end{aligned}$$

The state transition and offset transition coefficients under the feedback control (15) evolve respectively according to

$$\frac{\partial \Phi(t,s)}{\partial t} \triangleq (-a - b^2\pi_t r^{-1})\Phi(t,s),$$
$$\frac{\partial \Psi(t,s)}{\partial t} \triangleq (a + b^2\pi_t r^{-1} + \delta)\Psi(t,s).$$

The unique solutions to the state and offset equations are given respectively by

$$\bar{x}_t^\theta = \Phi(t,0)\bar{x}_0^\theta + \int_0^t \Phi(t,\tau)\big(-b^2 r^{-1} s_\tau^\theta + a\bar{x}_0^\theta\big)d\tau, \tag{40}$$

$$s_t^\theta = \int_\infty^t \Psi(t,\tau)\big(-a\bar{x}_0^\theta \pi_\tau + q_\tau^y z + q^{x_0} x_0^\theta\big)d\tau. \tag{41}$$

Integrating (38) with respect to F^θ, $s_t = \int_\Theta s_t^\theta dF^\theta$, gives

$$\begin{aligned} s_t &= \int_\Theta \int_\infty^t \Psi(t,\tau)\big(-a\bar{x}_0^\theta \pi_\tau + q_\tau^y z + q^{x_0} x_0^\theta\big)d\tau dF^\theta \\ &= \int_\infty^t \Psi(t,\tau)\big(-a\bar{x}_0 \pi_\tau + q_\tau^y z + q^{x_0} \bar{x}_0\big)d\tau. \end{aligned}$$

Taking the time derivative of s_t gives

$$-\frac{ds_t}{dt} = (-a - \delta - b^2\pi_t r^{-1})s_t + a\bar{x}_0\pi_t - q_t^y z - q^{x_0}\bar{x}_0.$$

Integrating (39) with respect to F^θ gives

$$\begin{aligned} \bar{x}_t &= \int_\Theta \Phi(t,0)\bar{x}_0^\theta dF^\theta + \int_\Theta \int_0^t \Phi(t,\tau)\big(-b^2 r^{-1} s_\tau^\theta + a\bar{x}_0^\theta\big)d\tau dF^\theta \\ &= \Phi(t,0)\bar{x}_0 + \int_0^t \Phi(t,\tau)\big(-b^2 r^{-1} s_\tau + a\bar{x}_0\big)d\tau. \end{aligned}$$

Taking the time derivative of $\bar{x}_t$ gives

$$\frac{d\bar{x}_t}{dt} = (-a - b^2\pi_t r^{-1})\bar{x}_t - b^2 r^{-1} s_t + a\bar{x}_0.$$

Hence, (17) is achieved. □

Proof of Lemma 4.2. Let $(f_n) \in \mathcal{G}$ be a Cauchy sequence in $\mathbf{C}_b[0,\infty)$. Since $\mathcal{G} \in \mathbf{C}_b[0,\infty)$, and $\mathbf{C}_b[0,\infty)$ is a Banach space, $f \triangleq \lim_{n\to\infty}(f_n)$ exists in $\mathbf{C}_b[0,\infty)$. Assume that there exists an $\epsilon > 0$ such that $\sup_{t\geq 0} f(t) \geq \bar{x}_0 + \epsilon$ (or $\inf_{t\geq 0} f(t) \leq z - \epsilon$ without loss of generality). This implies that there exists some f_N, $N \geq 0$, such that $\|f_N\|_\infty \geq \bar{x}_0 + \epsilon/2$. Clearly, this is a contradiction to the fact that $(f_n) \in \mathcal{G}$ for $n \geq 0$. Therefore, the limit f has the bound $z \leq f \leq \bar{x}_0$ which implies that $\mathcal{G}$ is closed in $\mathbf{C}_b[0,\infty)$.

For convexity of $\mathcal{G}$, let $f_1, f_2 \in \mathcal{G}$. Consider $\hat{f} = f_1 + (1-\gamma)f_2$ for $0 \leq \gamma \leq 1$. Note that $\hat{f}(0) = \bar{x}_0$, $z \leq \hat{f} \leq \bar{x}_0$ and $\hat{f}$ is continuous for $0 \leq \gamma \leq 1$, which implies convexity. □

Proof of Proposition 4.3.

Step 1: $Im(\mathcal{M})$ is bounded between z and $\bar{x}_0$.

Note that $q_t^y = 0$, $t \in [0,\infty)$, gives $\mathcal{T}(q) = \bar{x}_0$ since the optimal control is $u_t^\circ = 0$ for all $t \in [0,\infty)$. Also, note that as $q_t \to \infty$ $\bar{x}_t \to z$; and for fixed r this behaviour is monotone in q_t. Moreover since $\bar{x}_0 \leq y \leq z$, and $0 \leq q_t^y \leq \lambda \max[y - z, \bar{x}_0 - y]t$ due to Lemma A.1, for all $t \in [0,\infty)$, $\mathcal{T}(q)$ is bounded between $\bar{x}_0$ and z.

Step 2: $Im(\mathcal{M}) \subset \mathcal{G}$ and forms a family of equicontinuous functions.

In order to establish (2), we need first to rely on Lemmas A.1, A.2 and A.3 established earlier in Appendix A.

The unique solutions to the state and offset equations are given respectively by (40) and (41).

Injecting $q^y_\tau = |\lambda \int_0^\tau (\bar{x}_s - y)ds|$ gives

$$s_t = \int_\infty^t \Psi(t,\tau)\Big(- a\bar{x}_0\pi_\tau + \Big|\lambda \int_0^\tau (\bar{x}_s - y)ds\Big| z + q^{x_0}\bar{x}_0\Big) d\tau.$$

Then, the mean field state equation is written as

$$\begin{aligned}\bar{x}_t = \Phi(t,0)\bar{x}_0 + \int_0^t \Phi(t,\tau)\Big[- b^2 r^{-1} \int_\infty^\tau \Psi(\tau,s) \\ \Big(- a\bar{x}_0\pi_s + \Big|\lambda \int_0^s (\bar{x}_\zeta - y)d\zeta\Big| z + q^{x_0}\bar{x}_0\Big) ds + a\bar{x}_0\Big] d\tau \\ = (\mathcal{M}\bar{x})(t).\end{aligned} \tag{42}$$

For $t_2 = t_1 + \Delta t$, $0 \le t_1 \le t_2 < \infty$,

$$\begin{aligned}(\mathcal{M}x)(t_2) - (\mathcal{M}x)(t_1) &= \Phi(t_2,0)\bar{x}_0 + \int_0^{t_2} \Phi(t_2,\tau)(\cdot)d\tau \\ &\quad - \Phi(t_1,0)\bar{x}_0 + \int_0^{t_1} \Phi(t_1,\tau)(\cdot)d\tau \\ &= (\Phi(t_2,0) - \Phi(t_1,0))\bar{x}_0 \\ &\quad + \int_0^{t_2} \Phi(t_2,\tau)\Big[- b^2 r^{-1} \int_\infty^\tau \Psi(\tau,s)\Big(- a\bar{x}_0\pi_s + q^{x_0}\bar{x}_0\Big) ds\Big] d\tau \\ &\quad - \int_0^{t_1} \Phi(t_1,\tau)\Big[- b^2 r^{-1} \int_\infty^\tau \Psi(\tau,s)\Big(- a\bar{x}_0\pi_s + q^{x_0}\bar{x}_0\Big) ds\Big] d\tau \\ &\quad + \int_0^{t_2} \Phi(t_2,\tau)\Big[- b^2 r^{-1} \int_\infty^\tau \Psi(\tau,s)\Big|\cdot\Big| z ds\Big] d\tau \\ &\quad - \int_0^{t_1} \Phi(t_1,\tau)\Big[- b^2 r^{-1} \int_\infty^\tau \Psi(\tau,s)\Big|\cdot\Big| z ds\Big] d\tau \\ &\quad + \int_0^{t_2} \Phi(t_2,\tau)a\bar{x}_0 d\tau \\ &\quad - \int_0^{t_1} \Phi(t_1,\tau)a\bar{x}_0 d\tau \\ &\triangleq I_1 + I_2 - I_3 + I_4 - I_5 + I_6 - I_7,\end{aligned}$$

where the terms I_i, $1 \le i \le 7$, are defined respectively.

(a) $|I_1|$ is given as

$$\begin{aligned}|I_1| &= |\Phi(t_2,0) - \Phi(t_1,0)|\bar{x}_0 \\ &\le h|\Phi(t_2,0) - \Phi(t_1,0)|,\end{aligned} \tag{43}$$

since $\bar{x}_0 \le h$, the high comfort level. Then, we employ Lemma A.1 together with Lemma A.2 and get

$$|I_1| \le h\lambda\kappa_1\kappa_3|t_2 - t_1|.$$

Note that in Subsections (b), (c) and (d) below, the definitions of variable Δ_1 and Δ_2 are strictly local to the subsections.

(b) $I_2 - I_3$ is given as

$$\begin{aligned} I_2 - I_3 &= \int_0^{t_2} \Phi(t_2,\tau)\Big[b^2 r^{-1}\int_\infty^\tau \Psi(\tau,s)\Big(a\bar{x}_0\pi_s - q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\ &\quad - \int_0^{t_1} \Phi(t_1,\tau)\Big[b^2 r^{-1}\int_\infty^\tau \Psi(\tau,s)\Big(a\bar{x}_0\pi_s - q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\ &= \int_0^{t_1} (\Phi(t_2,\tau)-\Phi(t_1,\tau))\Big[b^2 r^{-1}\int_\infty^\tau \Psi(\tau,s)\Big(a\bar{x}_0\pi_s - q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\ &\quad + \int_{t_1}^{t_2} \Phi(t_2,\tau)\Big[b^2 r^{-1}\int_\infty^\tau \Psi(\tau,s)\Big(a\bar{x}_0\pi_s - q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\ &\triangleq \Delta_1 + \Delta_2. \end{aligned}$$

We start with Δ_1 and write

$$|\Delta_1| \le b^2 r^{-1}\bar{x}_0 \int_0^{t_1} |\Phi(t_2,\tau)-\Phi(t_1,\tau)| \Big|\int_\infty^\tau \Psi(\tau,s)(a\pi_s + q^{x_0})ds\Big| d\tau.$$

Note that $\bar{x}_0 \le h$; we employ Lemmas A.1 and A.3, and obtain

$$|\Delta_1| \le b^2 r^{-1} h(a+q^{x_0}) \int_0^{t_1} |\Phi(t_2,\tau)-\Phi(t_1,\tau)|d\tau.$$

We employ Lemma A.2 and get

$$|\Delta_1| \le b^2 r^{-1} h(a+q^{x_0})\lambda\kappa_1\kappa_4|t_2-t_1|.$$

Likewise, for Δ_2 we get

$$\begin{aligned} |\Delta_2| &\le b^2 r^{-1} h(a+q^{x_0}) \int_{t_1}^{t_2} \Phi(t_2,\tau)d\tau \\ &\le b^2 r^{-1} h(a+q^{x_0})a^{-1}\left(1-\exp[-a(t_2-t_1)]\right). \end{aligned}$$

Assuming $|t_2 - t_1| << 1$, we achieve the bound

$$|\Delta_2| \le b^2 r^{-1} h(a+q^{x_0})|t_2-t_1|.$$

(c) $I_4 - I_5$ is given as

$$\begin{aligned} I_4 - I_5 &= \int_0^{t_1} (\Phi(t_2,\tau)-\Phi(t_1,\tau))\Big[-b^2 r^{-1}\int_\tau^\infty \Psi(s,\tau)q_s^y z ds\Big]d\tau \\ &\quad + \int_{t_1}^{t_2} \Phi(t_2,\tau)\Big[b^2 r^{-1}\int_\tau^\infty \Psi(s,\tau)q_s^y z ds\Big]d\tau \\ &\triangleq \Delta_1 + \Delta_2. \end{aligned}$$

We start with Δ_1. Employing Lemma A.1 gives

$$|\Delta_1| \le \lambda b^2 r^{-1} z \max[y-z, \bar{x}_0 - y] \int_0^{t_1} |\Phi(t_2,\tau)-\Phi(t_1,\tau)| \left(\int_\tau^\infty \Psi(s,\tau)s ds\right) d\tau.$$

Employing Lemma A.1 together with Lemmas A.2 and A.3 gives

$$|\Delta_1| \le \lambda^2 b^2 r^{-1} z \max[y-z, \bar{x}_0 - y]\kappa_1\kappa_4|t_2-t_1|.$$

Then, employing Lemmas A.1 and A.3 gives

$$|\Delta_2| \leq \lambda b^2 r^{-1} z \max[y-z, \bar{x}_0 - y] \int_{t_1}^{t_2} \Phi(t_2,\tau) \left(\int_{\tau}^{\infty} \Psi(s,\tau) s ds \right) d\tau$$
$$\leq \lambda b^2 r^{-1} z \max[y-z, \bar{x}_0 - y] \int_{t_1}^{t_2} \Phi(t_2,\tau) d\tau.$$

The bound $\Phi(t_2,\tau) \leq \exp[-a(t-\tau)]$ gives

$$|\Delta_2| \leq \lambda b^2 r^{-1} z \max[y-z, \bar{x}_0 - y] a^{-1}(1 - \exp[-a(t_2 - t_1)]).$$

For $|t_2 - t_1| << 1$ we get

$$|\Delta_2| \leq \lambda b^2 r^{-1} z \max[y-z, \bar{x}_0 - y]|t_2 - t_1|.$$

(d) $I_6 - I_7$ is given as

$$I_6 - I_7 = \int_0^{t_1} (\Phi(t_2,\tau) - \Phi(t_1,\tau)) a\bar{x}_0 d\tau$$
$$+ \int_{t_1}^{t_2} \Phi(t_2,\tau) a\bar{x}_0 d\tau$$
$$\triangleq \Delta_1 + \Delta_2.$$

We start with Δ_1, employ $\bar{x}_0 \leq h$, Lemmas A.1 and A.2 to obtain

$$|\Delta_1| \leq ah\lambda\kappa_1\kappa_4|t_2 - t_1|.$$

Likewise for Δ_2, employing $\bar{x}_0 \leq h$ gives

$$|\Delta_2| \leq ah \int_{t_1}^{t_2} \exp[-a(t_2-\tau)] d\tau$$
$$\leq aha^{-1}\left(1 - \exp[-a(t_2 - t_1)]\right).$$

For $|t_2 - t_1| << 1$, we get

$$|\Delta_2| \leq ah|t_2 - t_1|.$$

Finally,

$$|(\mathcal{M}x)(t_2) - (\mathcal{M}x)(t_1)| \leq \Big(\lambda\kappa_1\kappa_3 h + (\lambda\kappa_1\kappa_4 + 1)(b^2 r^{-1} h)(a + q^{x_0})$$
$$+ (\lambda^2\kappa_1\kappa_4 + \lambda) b^2 r^{-1} z \max[y-z, \bar{x}_0 - y]$$
$$+ (\lambda\kappa_1\kappa_4 + 1) ah \Big) |t_2 - t_1|$$
$$\triangleq M|t_2 - t_1|. \qquad (44)$$

Note that $\mathcal{M}x(0) = \bar{x}_0$. We have shown in Step 1 that $z \leq \mathcal{M}(x) \leq \bar{x}_0$ for all $x \in \mathcal{G}$. Then, we have shown in Step 2 that $\mathcal{M}(x)$ is uniformly Lipschitz continuous for all $x \in \mathcal{G}$. This means (i) $Im(\mathcal{M}) \subset \mathcal{G}$, and (ii) $Im(\mathcal{M})$ forms a family of equicontinuous functions. □

Proof of Proposition 4.4. The operator $\mathcal{M}$ has been shown in (42) to be

$$\bar{x}_t = \Phi^{\bar{x}}(t,0)\bar{x}_0 + \int_0^t \Phi^{\bar{x}}(t,\tau) \Big[-b^2 r^{-1} \int_{\infty}^{\tau} \Psi^{\bar{x}}(\tau,s)$$
$$\Big(-a\bar{x}_0 \pi_s^{\bar{x}} + \Big|\lambda \int_0^s (\bar{x}_\zeta - y) d\zeta\Big| z + q^{x_0}\bar{x}_0 \Big) ds + a\bar{x}_0 \Big] d\tau$$
$$= (\mathcal{M}\bar{x})(t).$$

where

$$\frac{d\Phi^{\bar{x}}(t,s)}{dt} \triangleq (-a - b^2\pi_t^{\bar{x}}r^{-1})\Phi^{\bar{x}}(t,s),$$
$$\frac{d\Psi^{\bar{x}}(t,s)}{dt} \triangleq (a + b^2\pi_t^{\bar{x}}r^{-1} + \delta)\Psi^{\bar{x}}(t,s).$$

For x', $x'' \in \mathcal{G}$,

$$\begin{aligned}
\mathcal{M}x' - \mathcal{M}x'' &= \Phi^{\bar{x}'}(t,0)\bar{x}_0 + \int_0^t \Phi^{\bar{x}'}(t,\tau)(\cdot)^{\bar{x}'}d\tau \\
&\quad - \Phi^{\bar{x}''}(t,0)\bar{x}_0 + \int_0^{t_1} \Phi^{\bar{x}''}(t,\tau)(\cdot)^{\bar{x}''}d\tau \\
&= (\Phi^{\bar{x}'}(t,0) - \Phi^{\bar{x}''}(t,0))\bar{x}_0 \\
&+ \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big(q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}''}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big(q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\
&+ \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big(- a\bar{x}_0\pi_s^{\bar{x}'}\Big)ds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}''}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big(- a\bar{x}_0\pi_s^{\bar{x}''}\Big)ds\Big]d\tau \\
&+ \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big|\cdot\Big|^{\bar{x}'} zds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}''}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big|\cdot\Big|^{\bar{x}''} zds\Big]d\tau \\
&+ \int_0^t \Phi^{\bar{x}'}(t,\tau)a\bar{x}_0 d\tau \\
&- \int_0^t \Phi^{\bar{x}''}(t,\tau)a\bar{x}_0 d\tau \\
&\triangleq I_1 + I_2 - I_3 + I_4 - I_5 + I_6 - I_7 + I_8 - I_9,
\end{aligned}$$

where the terms I_i, $1 \le i \le 9$, are defined respectively.

(a) $|I_1|$ is given as

$$\begin{aligned}
|I_1| &= |\Phi^{\bar{x}'}(t,0) - \Phi^{\bar{x}''}(t,0)|\bar{x}_0 \\
&\le h|\Phi^{\bar{x}'}(t,0) - \Phi^{\bar{x}''}(t,0)|,
\end{aligned} \tag{45}$$

since $\bar{x}_0 \le h$, the high comfort level. Then, we employ Lemma A.1 together with Lemma A.2 and get

$$\|I_1\|_\infty \le h\lambda\kappa_1\kappa_3\|\bar{x}' - \bar{x}''\|_\infty.$$

Note that in Subsections (b), (c) and (d) below, the definitions of variable Δ_1 and Δ_2 are strictly local to the subsections.

(b) $I_2 - I_3$ is given as

$$\begin{aligned}
I_2 - I_3 &= \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big(q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\
&\quad - \int_0^t \Phi^{\bar{x}''}(t,\tau)\Big[- b^2r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big(q^{x_0}\bar{x}_0\Big)ds\Big]d\tau \\
&= \int_0^t \Phi^{\bar{x}'}(t,\tau)(\cdot)'d\tau - \int_0^t \Phi^{\bar{x}''}(t,\tau)(\cdot)'d\tau + \int_0^t \Phi^{\bar{x}''}(t,\tau)(\cdot)'d\tau - \int_0^t \Phi^{\bar{x}''}(t,\tau)(\cdot)''d\tau \\
&\triangleq \Delta_1 - \Delta_2 + \Delta_3 - \Delta_4.
\end{aligned}$$

With similar analysis to (b) in the Proof of Proposition 4.2 we get

$$\|\Delta_1 - \Delta_2\|_\infty \le b^2 r^{-1} h q^{x_0} \lambda \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty.$$

Likewise, for $\Delta_3 - \Delta_4$ we get

$$\|\Delta_3 - \Delta_4\|_\infty \le b^2 r^{-1} h q^{x_0} \lambda \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty,$$

so

$$\|I_2 - I_3\|_\infty \le 2b^2 r^{-1} h q^{x_0} \lambda \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty.$$

(c) Likewise, we get

$$\|I_4 - I_5\|_\infty \le 3b^2 r^{-1} h a \lambda \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty,$$

(d) $I_6 - I_7$ is given as

$$\begin{aligned}
I_6 - I_7 = &\int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big|\lambda\int_0^s(\bar{x}'_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}''}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big|\lambda\int_0^s(\bar{x}''_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
= &\int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big|\lambda\int_0^s(\bar{x}'_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big|\lambda\int_0^s(\bar{x}''_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
&+ \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}'}(\tau,s)\Big|\lambda\int_0^s(\bar{x}''_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big|\lambda\int_0^s(\bar{x}''_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
&+ \int_0^t \Phi^{\bar{x}'}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big|\lambda\int_0^s(\bar{x}''_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
&- \int_0^t \Phi^{\bar{x}''}(t,\tau)\Big[-b^2 r^{-1}\int_\infty^\tau \Psi^{\bar{x}''}(\tau,s)\Big|\lambda\int_0^s(\bar{x}''_\zeta - y)d\zeta\Big| z ds\Big]d\tau \\
\triangleq &\ \Delta_1 - \Delta_2 + \Delta_3 - \Delta_4 + \Delta_5 - \Delta_6.
\end{aligned}$$

For $\Delta_1 - \Delta_2$ we employ

$$\Big|\lambda\int_0^s(\bar{x}'_\zeta - y)d\zeta\Big| - \Big|\lambda\int_0^s(\bar{x}'_\zeta - y)d\zeta\Big| \le \lambda s\|\bar{x}' - \bar{x}''\|_\infty,$$

and obtain

$$\|\Delta_1 - \Delta_2\|_\infty \le b^2 r^{-1} z a^{-1} \lambda \|\bar{x}' - \bar{x}''\|_\infty.$$

Likewise

$$\|\Delta_3 - \Delta_4\|_\infty \le b^2 r^{-1} z \lambda^2 \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty,$$

and

$$\|\Delta_5 - \Delta_6\|_\infty \le b^2 r^{-1} z \lambda^2 \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty,$$

which implies

$$\|I_6 - I_7\|_\infty \le b^2 r^{-1} z (a^{-1}\lambda + 2\lambda^2 \kappa_1 \kappa_4) \|\bar{x}' - \bar{x}''\|_\infty.$$

(e) For $I_8 - I_9$ we get

$$\|I_8 - I_9\|_\infty \le a h \lambda \kappa_1 \kappa_4 \|\bar{x}' - \bar{x}''\|_\infty.$$

Finally,

$$\begin{aligned}\|\mathcal{M}x' - \mathcal{M}x''\|_\infty &\leq \Big[\lambda\big(h\kappa_1\kappa_3 + 2b^2r^{-1}hq^{x_0}\kappa_1\kappa_4 + 3b^2r^{-1}ha\kappa_1\kappa_4 + b^2r^{-1}za^{-1} + ah\kappa_1\kappa_4\big) \\ &\quad + \lambda^2\big(2b^2r^{-1}z\kappa_3\kappa_4\big)\Big]\|x' - x''\|_\infty \\ &\triangleq f(\lambda)\|x' - x''\|_\infty,\end{aligned}$$

where κ_1 is given in Lemma A.1, and κ_3 and κ_4 are given in Lemma A.2.

□

References

[1] R.P. Malhamé and C.-Y. Chong, Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system, IEEE Transactions on Automatic Control, 30(9), 854–860, Sep. 1985.

[2] J.C. Laurent and R.P. Malhamé, A physically-based computer model of aggregate electric water heating loads, IEEE Transactions on Power Systems, 9(3), 1209–1217, 1994.

[3] D.S. Callaway, Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy, Energy Conversion and Management, 50, 1389–1400, 2009.

[4] J. Mathieu, S. Koch, and D. Callaway, State estimation and control of electric loads to manage real-time energy imbalance, in Power and Energy Society General Meeting (PES), 2013 IEEE, July 2013, 1–1.

[5] L.C. Totu, Large scale demand response of thermostatic loads, Ph.D. dissertation, Aalborg University, 2015.

[6] J.M. Lasry and P.-L. Lions, Jeux à champ moyen. i - le cas stationnaire, C. R. Acad. Sci. Paris, Ser. I, 343, 619–625, 2006.

[7] M. Huang, P.E. Caines, and R.P. Malhamé, Individual and mass behaviour in large population stochastic wireless power control problems: Centralized and Nash equilibrium solutions, in 42nd IEEE Conference on Decision and Control, Maui, Hawaii, Dec. 2003, 98–103.

[8] M. Huang, R.P. Malhamé, and P.E. Caines, Large population stochastic dynamic games: Closed loop McKean-Vlasov systems and the Nash certainty equivalence principle, Special issue in honour of the 65th birthday of Tyrone Duncan, Communications in Information and Systems, 6, 221–252, Nov. 2006.

[9] M. Huang, P.E. Caines, and R.P. Malhamé, Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ – Nash equilibria, IEEE Transactions on Automatic Control, 52(9), 1560–1571, Sep. 2007.

[10] P.E. Caines, Mean field games, in Encyclopedia of Systems and Control, T. Samad and J. Ballieul, Eds. London: Springer-Verlag, 2014, 1–6.

[11] Y. Xu, L. Xie, and C. Singh, Optimal scheduling and operation of load aggregator with electric energy storage in power markets, in North American Power Symposium (NAPS), 2010, Sept 2010, 1–7.

[12] S. Meyn, P. Barooah, A. Bušić, and J. Ehren, Ancillary service to the grid from deferrable loads: The case for intelligent pool pumps in Florida, in Proceedings of the 52nd IEEE Conf. on Decision and Control, Dec 2013, 6946–6953.

[13] Y. Chen, A. Bušić, and S. Meyn, Individual risk in mean-field control models for decentralized control, with application to automated demand response, in Proceedings of the 53rd IEEE Conf. on Decision and Control, Dec 2014, 6425–6432.

[14] Y. Chen, A. Bušić, and S. Meyn, State estimation and mean field control with application to demand dispatch, 2015. [Online]. Available: http://arxiv.org/abs/1504.00088.

[15] J.L. Mathieu, D.S. Callaway, and S. Kiliccote, Examining uncertainty in demand response baseline models and variability in automated response to dynamic pricing, in 49th IEEE Conference on Decision and Control, 2010, 4332–4339.

[16] F. Tahersima, J. Stoustrup, S. Afkhami, and H. Rasmussen, Contribution of domestic heating systems to smart grid control, in 50th IEEE Conference on Decision and Control, 2011, 3677–3681.

[17] Z. Ma, D.S. Callaway, and I. Hiskens, Decentralized charging control for large populations of plug-in electric vehicles, in 49th IEEE Conference on Decision and Control, 2010, 206–212.

[18] B. Anderson and J.B. Moore, Optimal Control: Linear Quadratic Methods. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[19] R. Sonderegger, Dynamic models of house heating based on equivalent thermal parameters, Ph.D. dissertation, Princeton University, 1978.

[20] M. Nourian, P.E. Caines, R.P. Malhamé, and M. Huang, Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation, IEEE Trans. Autom. Cont., 57(11), 2801–2816, 2012.

[21] J.B. Conway, A Course in Functional Analysis - 2nd ed. Springer, 1990.

[22] A. Friedman, Partial Differential Equations. Dover Publications, 2008.

[23] W. Zhang, J. Lian, C.-Y. Chang, and K. Kalsi, Aggregated modeling and control of air conditioning loads for demand response, IEEE Transactions on Power Systems, 28(4), 4655–4664, Nov 2013.

[24] R.R. Bitmead and M. Gevers, The Riccati Equation. Springer-Verlag Berlin, Heidelberg, 1991, ch. Riccati Difference and Differential Equations: Convergence, Monotonicity and Stability, 263–291.

[25] F.W.J. Olver, D.W. Lozier, R.F. Boisvert, and C.W. Clark, NIST Handbook of Mathematical Functions. Cambridge University Press, 2010.