52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

Mean Field Based Control of Power System Dispersed Energy Storage Devices for Peak Load Relief

Arman C. Kizilkale Roland P. Malhamé

Abstract— **The demand response problem is investigated where it is required that the mean temperature of a massive number of electric devices associated with energy storage (e.g. electric water heaters, electric space heaters, etc.) follow a target temperature trajectory computed as part of a system wide optimization problem in smart grids. The classical control approach is to compute the required signal centrally for each device and to send that signal. However, the corresponding computational and communication requirements can be forbidding. Instead, in this paper the presence of large numbers of such devices justifies the use of a decentralized mean field control based approach to the problem. A novel agent cost structure including an integral term is utilized for this purpose. The corresponding system of mean field equations is developed, a fixed point analysis is given for a particular case, an ϵ-Nash theorem is presented, and numerical simulation results are provided.**

I. INTRODUCTION

The advances in smart grid systems enable users to track the electricity price signals periodically. At first look, 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, this intuition might be misleading. Indeed, individual responses without a proper dynamic optimization perspective can possibly lead to oscillations in the aggregate load, and in fact possibly lead to amplified albeit delayed peaks.

One implementation alternative aimed at achieving peak load reduction 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. On the theoretical front, while some challenges exist, the optimization problem is not insurmountable, but the implementation part faces several difficulties. To begin with, the computation power required to globally generate the control action for the demand response of thousands of households is immense. Moreover, it is not desirable to interfere with individual load frequently. And lastly, such an approach requires a significant amount of signal exchange between the controller and the households, whose demand remains essentially stochastic and in the event of a communication loss, unexpected outcomes are possible both on the aggregate level and on the individual user level.

In the global energy generation equation several reasons favor renewable energy (wind, solar, etc.) as a crucial option for the power grid; the most notable are to be cleanliness in terms of carbon emissions, cost independence from the highly volatile and mostly increasing oil prices, and the potential cost of carbon emission taxes. Nevertheless, renewable sources tend to be structurally problematic; they lack continuous availability due to factors out of the power system direct control. Thus, even though a wind farm is highly reliable and cheap to sustain, it usually has a low capacity factor: it works at a low percentage of the nameplate capacity. This fact leads to more spinning reserve production which is typically costly and polluting. An alternative approach explored in this paper is to employ 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 paper, a demand response *mechanism* is implemented that will employ the free storage in the grid (water heaters, space heaters, etc.) 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 homogeneous 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.

(1) Each controller has to be situated locally.

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

(3) User disturbance should be kept at minimum.

We elaborate further on (3). Since the centralized authority is solely interested in the aggregate consumption, individual trajectories do not need to necessarily follow the targets set by the authority. On the contrary, in fact in the case of a population of controlled electric space heaters, or air conditioners, or electric water heaters for example, it is desirable to shape the mean temperature of the population with least disturbance; ideally without customers even noticing the

The authors gratefully acknowledge the support of Natural Resources Canada.

Arman C. Kizilkale and Roland P. Malhamé are with the Department of Electrical Engineering at École Polytechnique de Montréal and Groupe d'Études et de Recherche en Analyse des Décisions (GERAD), Canada. {arman.kizilkale;roland.malhame}@polymtl.ca

978-1-4673-5716-6/13/$31.00 ©2013 IEEE

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.

Statistical mechanics inspired load models of large aggregates of energy storage associated loads, particularly space heating and cooling loads, to be controlled within peak shaving and valley filling load management programs by direct control were presented in [1], whereas the modeling methodology was applied to electric water heaters in [2].

Using dispersed storage for accommodating renewable intermittency is a growing area of research. Dispersed energy storage for frequency regulation in the presence of wind energy is investigated in [3]. This work uses the aggregate load modeling framework in [1] and extends it for improved transient analysis. Dynamic pricing for controlling the load of aggregates of large commercial buildings is analyzed in [4], and domestic heating systems are employed as heat buffers in [5]. A decentralized charging control strategy for large populations of plug-in electric vehicles (PEVs) using the mean field methodology is presented in [6].

The rest of the paper is organized as follows. In Section II we briefly introduce linear quadratic Gaussian (LQG) stochastic optimal tracking control together with a brief LQG mean field theory. In Section III we introduce the model that will be used throughout the paper, propose our collective target tracking mean field formulation, present a fixed point analysis for a particular case, introduce a numerical algorithm that achieves a desirable fixed point, and present our ϵ-Nash Theorem indicating that an approximate Nash Equilibrium is attained. Lastly, in Section IV, we provide simulation results together with comparisons to a prevailing target tracking control formulation.

We denote the set of nonnegative real numbers by $\mathbb{R}_+$. The norm $\|\cdot\|$ denotes the 2-norm of vectors and matrices, and $\|x\|_Q^2 \triangleq x^\top Qx$. $\mathbf{C}_b = \{x : x \in \mathbf{C}, \sup_{t\geq 0}\|x_t\| < \infty\}$ denotes the family of all bounded continuous functions, and for any $x \in \mathbf{C}_b$, $\|\cdot\|_\infty$ denotes the supremum norm: $\|x\|_\infty \triangleq \sup_{t\geq 0}\|x_t\|$. $\mathrm{Tr}(X)$ denotes the trace, and $X^\top$ denotes the transpose of a matrix X.

II. BACKGROUND ON LINEAR QUADRATIC GAUSSIAN MEAN FIELD CONTROL

In order to meet the three requirements above the principles of the mean field control methodology [7] are employed. This framework is based on a decentralized scheme whereby each agent calculates its individual best response to the anticipated action profile of the population. Under technical constraints together with the assumption of each agent's individual rationality, the approach looks for a Nash Equilibrium as the number of users goes to infinity when agents implement their stabilizing best response actions. The anticipation of the action profile is carried out offline and locally with statistical information obtained based on a parsimonious population measurement scheme and available at the start of the control horizon. The control scheme is fully decentralized; i.e., no communication takes place among the controllers throughout the horizon.

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

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

$1 \leq i \leq N$, 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 $\bar{x}^N = (1/N)\sum_{i=1}^N x^i$.

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

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

where the cost-coupling is assumed to be in the form of an averaging function $m_t^N \triangleq m(\bar{x}_t^N + \eta)$, $\eta \in \mathbb{R}^n$. 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 $\mathcal{A}_{-i} = \{\mathcal{A}_j, j \neq i, 1 \leq j \leq N\}$.

For the basic MF control problem, the following assumptions are adopted.

A1: The disturbance processes w^i, $1 \leq i \leq N$, 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: Θ is a compact set such that for each $\theta = [A^i, B^i, c^i] \in \Theta$, $[A^i, B^i]$ is controllable and $[Q_t^{1/2}, A^i]$, $t \in [0, \infty)$, is observable. ■

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

Each agent $\mathcal{A}_i, 1 \leq i \leq N$, obtains the positive definite solution to the algebraic Riccati equation

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

for $t \in [0, \infty)$. Moreover, for a given posited *mass tracking signal* $x^* \in \mathbf{C}_b[0, \infty)$ the mass offset function s^i is generated by the differential equation

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

for $t \in [0, \infty)$. Then, the optimal tracking control law [8] is given by

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

Note that x^* 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 in Definition 2.1 below.

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

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

Note in the above the potential heterogeneity of the agent dynamic parameters is indicated by a parameter θ considered as a random vector on a compact set (see [7] for further details).

Theorem 2.1: MF Stochastic Control Theorem [7]

Let **A1**-**A3** hold; under further technical assumptions (see [7]), the MF Stochastic Control Law (5) generates a set of controls $\mathcal{U}_{MF}^N \triangleq \{(u^i)^0; 1 \le i \le N\}$, $1 \le N < \infty$, with

$$(u_t^i)^0 = -R^{-1}(B^i)^\top (\Pi_t^i x_t^i + s_t^i), \quad t \ge 0, \tag{7}$$

such that

(i) the MF equations (6) have a unique solution.
Note in the above that potential heterogeneity of the agent dynamic parameters is indicated by a parameter θ considered as a random vector on a compact set (see [7] for further details).

(ii) All agent system trajectories x^i, $1 \le i \le N$, are L^2 stable;

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

$$
\begin{aligned}
J_i^N\left((u^i)^0, (u^{-i})^0\right) - \epsilon &\le \inf_{u^i \in \mathcal{U}_g^N} J_i^N\left(u^i, (u^{-i})^0\right) \\
&\le J_i^N\left((u^i)^0, (u^{-i})^0\right).
\end{aligned}
$$

In essence Theorem 2.1 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.

III. ELECTRIC SPACE HEATER MODELS

The currently prevailing mean field LQG control formulation provides the requirements (1) and (2) in the Introduction Section, but fails short of objective (3). 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 [9]) together with 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 as compared to prevailing mean field control formulations where the mean field effect is on the tracking signal. The resulting concept will be called *mean field based collective target tracking*.

A. Classical Linear Quadratic Gaussian (LQG) Tracking

In this paper we assume that each heating device is modeled by a diffusion process

$$dx_t^i = [a(x_t^i - x_0^i) + bu_t^i]dt + \sigma dw_t^i, \quad t \ge 0, \tag{8}$$

$1 \le i \le N$. Note that the model is a simpler version of the model given in [1], where the thermostat control in [1] is exchanged with a linear control. Also, recalling implementation principle (3), we consider that users "naturally" would like their devices to stay at their initial temperatures (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.

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

$$J_i^N(u^i, u^{-i}) = \mathbb{E}\int_0^\infty e^{-\delta t}\left[(x_t^i - y)^2 q + b(u_t^i)^2 r\right] dt. \tag{9}$$

The problem with this approach is that each agent minimizes its own cost function and tracks the same signal. 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.

B. Collective Target Tracking Mean Field Model

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

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

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, $t \in [0, \infty)$, which captures the mean field information and calculated as the following *integrated error signal*:

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

where y is the main control center dictated mean target signal.

The justification for the above cost function is that by pointing individual agents towards what is considered as the minimum comfort temperature z, it dictates a global decrease in their individual temperatures. This pressure for decrease 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 which, in the steady-state, should maintain the mean population temperature at y. When this happens, individual agents reach themselves their steady state (in general different from y and closer to their comfort zone than classical LQG tracking would dictate).

In order to derive the limiting infinite population MF equation system, and analogously to the more classical MF LQG case in Section II, we start this time assuming a constant (although unknown) cost penalty trajectory q_t. Given q_t, individual agents $\mathcal{A}_i$, $1 \le i \le N$, solve a classical target tracking LQG problem with time varying cost coefficient with Riccati gain π_i and offset term s_i evolving as follows:

$$
\begin{aligned}
-\frac{d\pi_t^i}{dt} &= (2a-\delta)\pi_t^i - b^2 r^{-1}(\pi_t^i)^2 + q_t, \quad t \ge 0,\\
-\frac{ds_t^i}{dt} &= (a-\delta-b^2\pi^i r^{-1})s_t^i - a x_0^i \pi^i - q_t z, \quad t \ge 0,
\end{aligned}
$$

Then, the optimal tracking control law [8] is given by

$$
u_t^i = -br^{-1}(\pi_t^i x_t^i + s_t^i), \quad t \ge 0. \tag{10}
$$

The calculation of the unknown q_t, $t \ge 0$, is obtained by requiring that $q_t, t \ge 0$, must be such that the individual agents carrying their corresponding optimal responses must collectively replicate $q_t, t \ge 0$, itself. This fixed point requirement leads to the specification in Definition 3.1 below of the Collective Target Tracking MF Equation System.

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

$$
\begin{aligned}
-\frac{d\pi_t^\theta}{dt} &= (2a-\delta)\pi_t^\theta - b^2 r^{-1}(\pi_t^\theta)^2 + q_t,\\
-\frac{ds_t^\theta}{dt} &= (a-\delta-b^2\pi^\theta r^{-1})s_t^\theta - a x_0^\theta \pi^\theta - q_t z,\\
\frac{d\bar{x}_t^\theta}{dt} &= (a-b^2\pi^\theta r^{-1})\bar{x}_t^\theta - b^2 r^{-1} s_t^\theta - a x_0^\theta,\\
\bar{x}_t &= \int_\Theta \bar{x}_t^\theta dF^\theta,\\
q_t &= \left|\int_0^t (\bar{x}_t - y)dt\right|.
\end{aligned} \tag{11}
$$

Note that the MF Equations for this model is significantly different from (6). Indeed system (6) is 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 (11) is fundamentally nonlinear (because of the form of q_t, $t \ge 0$,) and special arguments have to be developed for analysis.

C. Fixed Point Analysis

Here we present the fixed point analysis for the Collective Target Tracking MF Equation System for the particular case $z < y < x_0$. Due to the lack of space, the lemmas and theorems will be provided without proofs.

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

$$
\begin{aligned}
q_t &= \left|\int_0^t (\bar{x}_t - y)dt\right|\\
&\triangleq \Delta(\bar{x}_t).
\end{aligned}
$$

Next we define $\mathcal{T} : \mathbf{C}_b[0,\infty) \to \mathbf{C}_b[0,\infty)$ for the equation system below:

$$
\begin{aligned}
-\frac{d\pi_t^\theta}{dt} &= (2a-\delta)\pi_t^\theta - b^2 r^{-1}(\pi_t^\theta)^2 + q_t,\\
-\frac{ds_t^\theta}{dt} &= (a-\delta-b^2\pi^\theta r^{-1})s_t^\theta - a x_0^\theta \pi^\theta - q_t z,\\
\frac{d\bar{x}_t^\theta}{dt} &= (a-b^2\pi^\theta r^{-1})\bar{x}_t^\theta - b^2 r^{-1} s_t^\theta - a x_0^\theta,\\
\bar{x}_t &= \int_\Theta \bar{x}_t^\theta dF^\theta,
\end{aligned}
$$

which is equivalent to $x_t \triangleq \mathcal{T}(q_t)$.

Hence, one can write the MF Equation System for Collective Target Tracking as

$$
\begin{aligned}
x_t &= (\mathcal{T} \circ \Delta)(x_t)\\
&= \mathcal{M}(x_t), \quad t \in [0,\infty).
\end{aligned}
$$

Define the set $\mathcal{G}$ endowed with $\|\cdot\|_\infty$, where $f \in \mathbf{C}_b[0,\infty), f(0) = x_0$ and $z \le f(t) \le x_0$ for $t \in [0,\infty)$. Note that $\mathcal{G}$ is clearly nonempty.

Lemma 3.1: Under $\|\cdot\|_\infty$, $\mathcal{G}$ is closed and convex in $\mathbf{C}_b[0,\infty)$.

Lemma 3.2: We have $\mathcal{M}(x) \in \mathcal{G}$ for all $x \in \mathcal{G}$.

Theorem 3.3: For $z < y < x_0$ there exists a fixed point for the map $\mathcal{M} : \mathcal{G} \to \mathcal{G}$.

The existence of a fixed point for (11) 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 this equilibrium will be presented below.

D. Chopped Operator

In order to concentrate on the monotonicity of the trajectories for a system when $z < y < x_0$, we employ a so-called chopped operator whereby anytime the mean state hits the target y, it is frozen at y.

First define the set $\mathcal{G}^c \subset \mathcal{G}$ endowed with $\|\cdot\|_\infty$, where $f \in \mathbf{C}_b[0,\infty), f(0) = x_0$ and $y \le f(t) \le x_0$ for $t \in [0,\infty)$. Also, define $\mathcal{T}^c : \mathbf{C}_b[0,\infty) \to \mathbf{C}_b[0,\infty)$, where

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

Next define $\mathcal{M}^c \triangleq \mathcal{T}^c \circ \Delta : \mathbf{C}_b[0,\infty) \to \mathbf{C}_b[0,\infty)$.

The two theorems below provide the existence of a fixed point for the operator $\mathcal{M}^c$ on $\mathcal{G}^c$ and its relation to the fixed point analysis of $\mathcal{M}$ on $\mathcal{G}$.

Theorem 3.4: For fixed $z < y < x_0$, there exists a fixed point under the operator $\mathcal{M}^c$ on $\mathcal{G}^c$.

Theorem 3.5: If for $\mathcal{M}^c(\hat{x}) = \hat{x}$ for $\hat{x} \in \mathcal{G}^c$, $\hat{x}$ is smooth, then $\hat{x}$ is the fixed point to the operator $\mathcal{M}$ on $\mathcal{G}$. Also conversely, if $\mathcal{M}(\hat{x}) = \hat{x}$ for $\hat{x} \in \mathcal{G}$, $\hat{x}$ is smooth and converges to y, then $\hat{x}$ is the unique fixed point to the operator $\mathcal{M}^c$ on $\mathcal{G}^c$.

E. A Numerical Algorithm for the Chopped Operator

Here we present an algorithm that converges to the fixed point presented in Theorem 3.5.

First define q^*_∞: the steady state cost, $\Delta(\bar{x})$ of $\bar{x}$ for $\bar{x} \to y$ as $t \to \infty$, which is calculated below:

$$q^*_\infty = \frac{a(a-\delta)r}{b^2}\left(\frac{x_0 - x_\infty}{x_\infty - z}\right).$$

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

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

Definition 3.2: Chopped Operator Algorithm

1) set $k = 0$
2) set $\bar{x}_t = x_0$ for $t \geq 0$
3) set $x^{old} = \infty$
4) for $\|\bar{x} - x^{old}\|_\infty > \epsilon$
   - set $x^{old} = \bar{x}$
   - calculate $q = \Delta(\bar{x})$
   - if mod(k,2) == 1 & $q_\infty > q^*_\infty$, EXIT
   - if $q_\infty > q^*_\infty$, set $q_t = q^*_\infty$ for $t \in [T_{q^*_\infty}, \infty)$
   - calculate $\bar{x} = T^c(q)$
   - set $k = k + 1$
5) EXIT.

Remark 1: The convergence of the algorithm to a smooth fixed point on $\mathcal{G}^c$ is a necessary and sufficient condition for the existence of a desirable smooth fixed point for $\mathcal{M}$ on $\mathcal{G}$.

F. ϵ-Nash Theorem

Here we present the main theorem of the paper. We have seen that for an infinite population the system achieves a fixed point theorem for the particular case $z < y < x_0$. The next theorem provides an MF stochastic control law that achieves a Nash equilibrium at the population limit when applied by all agents in the system. Moreover, an ϵ-Nash equilibrium property holds for any finite population.

Theorem 3.6: Collective Target Tracking MF Stochastic Control Theorem

For $z < y < x_0$ if a desirable smooth fixed point exists for $\mathcal{M}$ on $\mathcal{G}$, the Collective Target Tracking MF Stochastic Control Law (10) generates a set of controls $\mathcal{U}^N_{col} \triangleq \{(u^i)^0; 1 \leq i \leq N\}$, $1 \leq N < \infty$, with

$$(u^i_t)^0 = -br^{-1}(\pi^i_t x^i_t + s^i_t), \quad t \geq 0, \tag{12}$$

such that

(i) $\mathcal{M}^c$ has a unique fixed point on $\mathcal{G}^c$.
(ii) all agent system trajectories x^i, $1 \leq i \leq N$, are L^2 stable;
(iii) $\{\mathcal{U}^N_{col}; 1 \leq N < \infty\}$ yields an ϵ-Nash equilibrium for all $\epsilon > 0$.

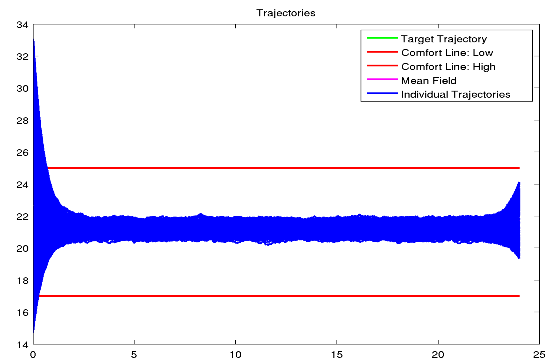

Fig. 1. Agents Applying Classical LQG Tracking

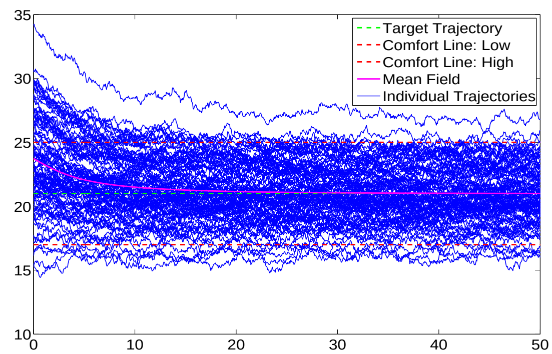

Fig. 2. Agents Applying Collective Target Tracking MF Control: All Agents Following the Low Comfort Level Signal

IV. SIMULATIONS

Here in the first simulation we simulate a population of 400 heaters. The mean initial temperature is 24 degrees. The central authority sets the target temperature to 21 degrees over a 24 hours horizon. The central authority provides the target temperature trajectory to each controller, and local controllers solve an LQG tracking problem as provided in Section III-A. Figure 1 shows that each agent in the population tracks the target degree of 21 degrees. Notice that in this implementation all agents track 21 degrees in order for the population mean temperature to track 21 degrees, where all agents are heavily disturbed for the global goal.

In another scenario the central authority sets the target temperature to 21 degrees, and all agents are assigned to track 17 degrees. The trajectories are presented and the calculated mean field signal is shown in Figure 2. It can be seen that while the mean temperature still settles at 21 degrees, the population is disturbed much less than the LQG tracking implementation.

In Figure 3 for the same simulation we plot the iterations of the Collective Target Tracking Algorithm without chopping. 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 chopped operator algorithm is introduced to take advantage of that

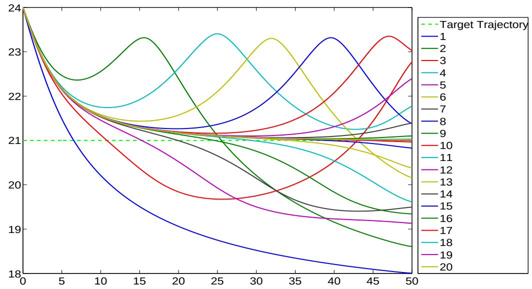

Fig. 3. Collective Target Tracking MF Iterations

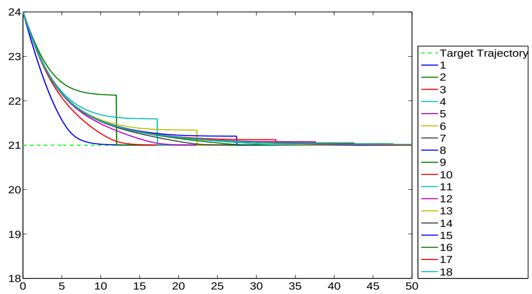

Fig. 4. Collective Target Tracking MF Iterations: Chopped Algorithm

monotonicity.

In Figure 4 we plot the iterations of the chopped algorithm. Note that the even curves are chopped at the point when $q_t = q^*_\infty$ and the odd curves are chopped when they cross the target y. The algorithm converges to the same mean field trajectory as in the previous Figure 3.

For the next experiment, we study a population of initial mean temperature at 25 degrees. We separate the population in two groups where the first group consists of the agents above 25 degrees initial temperature and the second group consists of the ones below 25 degrees. Both groups are assigned to track 17 degrees. In order to achieve a level of fairness, the first group is assigned a smaller control penalty coefficient r. Collective Target Tracking MF control is applied to these groups, and the simulation result is provided in Figure 5. It can be seen that the MF based integral error control scheme takes the mean temperature of the whole population to 21 degrees disturbing the agents less than in Figure 1.

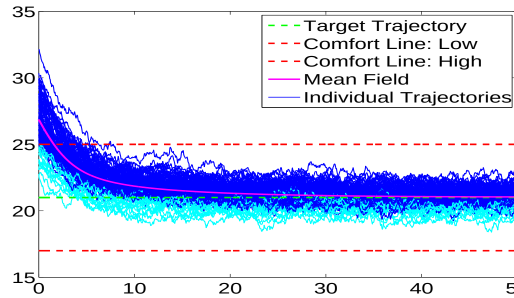

Fig. 5. Agents Applying Collective Target Tracking MF Control: Different r for subpopulations

V. 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 integral error, as compared to currently prevailing mean field control formulations where the mean field effect is concentrated on the tracking signal. The corresponding system of optimal control equations is developed, a fixed point analysis is provided for a particular case, an ϵ-Nash Theorem is presented and numerical simulation results are provided.

In future work, the analysis of nondiffusion load dynamics which involve jump Markov models 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.

REFERENCES

[1] R. P. Malhamé and C.-Y. Chong, "Eletric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Transactions on Automatic Control*, vol. 30, no. 9, pp. 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*, vol. 9, no. 3, pp. 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*, vol. 50, pp. 1389–1400, 2009.

[4] 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, pp. 4332–4339.

[5] 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, pp. 3677–3681.

[6] 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, pp. 206–212.

[7] 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*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.

[8] A. Bensoussan, *Stochastic Control of Partially Observable Systems*. U. K.: Cambridge Univ. Press, 1992.

[9] 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.*, vol. 57, no. 11, pp. 2801–2816, 2012.