2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

Mean Field Game Based Control of Dispersed Energy Storage Devices with Constrained Inputs

Feng Li, Roland P. Malhamé and Jerome Le Ny

Abstract— **In this paper, we study the mean field control of a large population of electric space heaters with linear dynamics and saturation constraints on the inputs. The mean field model is described by the fixed point of a system of coupled partial differential equations. A numerical algorithm is proposed to find this fixed point, which takes advantage of the special form of the individual control problem. We derive a decentralized control mechanism based on the mean field equilibrium solution under which the mean temperature of the population follows a set target temperature, while controls for each device are generated locally and attempt to keep individual temperature deviations small. We illustrate the results using numerical simulations, and compare the solutions obtained to the case where control inputs are unconstrained.**

I. INTRODUCTION

With the increasing levels of renewable power generation (wind turbines, solar panels) connected to power grids, it is always important to balance between the load demand and the generation. However, due to the intermittent characteristics of renewable generation (which amounts to high variability), auxiliary energy (such as bulk energy storage) is often required to maintain such balance [1]. In this context, the dispersed household devices (such as electric water heaters, electric space heaters, heat pumps, refrigerators, etc.), constitute a readily available source of energy to help mitigate the variability of renewable power generation [2]. These devices are naturally present in the power system and come in large quantities.

In order to utilize and control these dispersed devices, it is desirable to develop a decentralized control mechanism, which ideally should meet the following requirements: 1) control actions should be computed locally by each device yet preserve global optimality; 2) the level of data exchange between the central authority and users should be kept to a minimum; and 3) the disturbance to users load profile relative to the uncontrolled situation should be kept small.

A class of decentralized control mechanisms for load management of power systems were previously designed in [3]–[5] based on the mean field game (MFG) equilibrium solution of linear quadratic Gaussian (LQG) models with integral control in the cost coefficients. For example, electric space heaters with diffusion dynamics are controlled to collectively reach a target temperature under a non-cooperative framework in [3] and under a cooperative framework in [4]. In [5] water heaters with Markovian jump-driven hot water demand models are used as agents under the mean field control. The equilibrium solution concept used to derive these decentralized control mechanisms is based on the ϵ-Nash Equilibrium theorem derived in [6]. However, the cited literature does not consider constraints on control inputs when deriving optimal control laws. The power demands of a large number of home appliances are controlled in a decentralized fashion based on frequency regulation in [7], and at equilibrium each agent is following a bang-bang-like switching control under control and state constraints. The set of feasible aggregated power trajectories are characterized in [8] such that a population of thermostatically controlled loads (TCLs) can follow them, where each TCL has a hybrid dynamics and the mode switching rate is constrained. Mean field problems under convex state and control constraints are considered in [9], and several decentralized iteration approaches and conditions to converge to a fixed point are proposed.

In this paper, we extend the study of decentralized control of a large number of space heaters as energy storage devices initiated in [3], but we consider the case where the local control inputs are constrained. Indeed, in most engineering applications control inputs may be constrained due to physical or design limitations of the system. In the presence of control input constraints, the formulation under the LQG MFG setup fails to fully characterize the optimal control solutions, and the mean field effect must be described by partial differential equations (PDE's) rather than ordinary differential equations (ODE's; see [10]).

The paper makes three contributions. First, when applying a nonlinear optimal control law under saturation into the dynamics, we produce a time varying random population distribution, the mean of which is described by a PDE that we characterize. We also present the fixed point mean field game equations characterizing the limiting infinite population Nash equilibrium, and describe a numerical algorithm to solve and find the equilibrium solution as a fixed point. Second, we argue that if all quantities converge to a steady-state equilibrium, the latter must be the same in the constrained and unconstrained cases. Thus it is only the transient dynamics which are affected by saturation effects. Third, we develop sufficient conditions under which any device's best response involves either a single switching from saturated to unsaturated control, or a control which never saturates, and establish under these conditions a numerical scheme to compute the aggregate response of the devices.

The rest of the paper is organized as follows. Section 2 presents the MFG model to control the mean temperature

This work was funded by a grant from Natural Resources Canada. The authors are with the department of Electrical Engineering, Polytechnique Montreal and with GERAD, Montreal, Quebec, Canada H3T1J4. {feng.li, roland.malhame, jerome.le-ny}@polymtl.ca.

of a large population of electric heaters under constrained control inputs. In Section 3 we analyze the control policy, present our sufficient conditions for at most single switching best response policies of individuals given a posited mean field, develop the fixed point equations characterizing the associated equilibrium mean field, and propose a numerical algorithm for their solution. In Section 4 we illustrate the equilibrium solution under constrained control inputs by simulation.

II. Mean Field Game Model with Constrained Control Inputs

We consider a large population of N space heating devices (agents of the game) in the power grid. It is assumed that a central authority (either the power system authority or an "aggregator" managing the group of devices) would like the mean state of all agents to follow a target trajectory. We assume that the dynamics of each agent $i \in [N] := \{1, \dots, N\}$ are described by the following process

$$\frac{dx^i}{dt} = -a(x^i - x_a) + bu^i_{tot}, \;\; t \geq 0,$$

where $a > 0$, $b > 0$, u^i_{tot} is the control action, and x_a is the outside ambient temperature. Let $x^i_0 \in \mathbb{R}$ be agent i's random initial condition, distributed according to some known probability density function (pdf) ρ_0. We assume $x_a < x^i_0, \;\; \forall i \in [N]$. Each device heater wants naturally to remain at its initial temperature, which requires a (constant) control

$$u^i_{free} := ab^{-1}(x^i_0 - x_a).$$

Such u^i_{free} is considered a *free* control and is not penalized in the cost function defined below. Hence, we are only interested in the control effort u^i that is required to deviate from x^i_0 so as to drive the population mean temperature towards a target temperature set by the authority and denoted y. Hence by writing $u^i_{tot} = u^i + u^i_{free}$, we can reformulate and simplify the process model to

$$\frac{dx^i}{dt} = -a(x^i - x^i_0) + bu^i \triangleq f(x^i, u^i), \;\; t \geq 0. \quad (1)$$

Under the control constraint, it is the total control $u^i_{tot} = u^i + u^i_{free}$ that is constrained to a common set of values $u^i_{tot} \in \mathcal{U}_{tot} \triangleq \{u | u_{\min} \leq u \leq u_{\max}\}$ for all N agents. The constraints $u_{\min}$ and $u_{\max}$ need to be chosen carefully in order to make sense physically. Specifically, we must have $0 \leq u_{\min} \leq \min\{u^i_{free}, \forall i \in [N]\}$, and $u_{\max} \geq \max\{u^i_{free}, \forall i \in [N]\}$. These are imposed because at $t = 0$ all u^i_{free} should fall within the admissible control set $[u_{\min}, u_{\max}]$. Also, the minimum control u^i_{tot} that agent i can exert is to not heat, hence $u_{\min} \geq 0$. Hence we have the admissible control $u^i \in \mathcal{U}^i \triangleq \{u | u_{\min} - u^i_{free} \leq u \leq u_{\max} - u^i_{free}\}$, where $u_{\min}$ and $u_{\max}$ meet the conditions mentioned above. Note in particular that the constraint set for u^i depends on the agent i.

As in [3], the goal is to move the *average* temperature of the agents' population to the target y, while keeping each individual temperature relatively close to its initial value. In particular, we want to avoid controls u^i that move each individual temperature x^i to y. This motivates the definition of the following cost function for each agent i [3]

$$J_i(u^i, u^{-i}) = \int_0^\infty e^{-\delta t}[(x^i - z)^2 q_t + (x^i - x^i_0)^2 q_0 + (u^i)^2 r] dt, \quad (2)$$

where δ, q_0, r are positive constants and the temperature z serves as a direction signal to all agents, such that all agents should move toward z but not beyond. The penalization coefficient q_t, defined below, is calculated according to the integrated difference between the mean field temperature of the entire population and the constant target y

$$q_t = \left| \lambda \int_0^t (\bar{x} - y) dt \right| + k_c, \quad (3)$$

where $\lambda, k_c > 0$, and $\bar{x} = \frac{1}{N} \sum_{i=1}^N x^i$ is the mean temperature of the population. The cost function J_i is defined so that any agent i feels the pressure q_t built up from the difference between the mean field temperature $\bar{x}$ of the population and the target temperature y. In the case where $y < \bar{x}_0$, we are asking all population to reduce the temperature, and we must have $\bar{x} \geq y > z > x_a$. If q_t is to achieve a steady-state, then $\bar{x}$ must approach y asymptotically. At this point, all agents reach an individualized, initial condition dependent steady-state and maintain the population mean temperature at y. k_c is a positive constant term, which provides some initial pressure to all agents at the start of the control horizon. As q_t is calculated based on mean temperature $\bar{x}$, and the latter becomes deterministic as the number of agents goes to infinity, q_t can be viewed in the limit as a given function of time t, and we then write $J_i(u^i, q_t)$. For a posited q_t trajectory, each agent chooses its best response $u^i_* = \arg\min_{u^i \in \mathcal{U}^i} J_i(u^i, q_t)$ to minimize J_i while respecting the control constraint. By defining the value function $V_i(x^i, t)$ as the optimal cost to go starting from x^i at time t, we write the Hamilton-Jacobi-Bellman (HJB) equation for agent i to find optimal u^i_* under constraint

$$-\frac{\partial V_i}{\partial t} = \inf_{u^i \in \mathcal{U}^i} \left(L(x^i, u^i, t) + \frac{\partial V_i}{\partial x^i} f(x^i, u^i) \right), \quad (4)$$

where

$$L(x^i, u^i, t) = \frac{1}{2} e^{-\delta t}[(x^i - z)^2 q_t + (x^i - x^i_0)^2 q_0 + (u^i)^2 r],$$
$$f(x^i, u^i) = -a(x^i - x^i_0) + bu^i.$$

A. MFG Approximation

As in the usual MFG set up, we consider a limiting infinite population situation where the mass effect $\bar{x}$ characterizing q_t, the weight in the cost function of individual agents, is posited as given. The solution of the HJB Equation describes the optimal control u^i_* as a best response to the *mass* behavior for each agent i. However, in contrast to the standard linear quadratic MFG setup, the dynamics (1), the cost function in (2) and the constraint set $\mathcal{U}^i$ depend throughout the control horizon on the initial condition of the particular agent. To address this difficulty, we still view the population of agents as a continuum, but approximate the mass dynamics by discretizing the agents' initial temperature pdf ρ_0. First, we partition the set of initial temperatures into disjoint intervals $\Theta_k, k = 1\dots,K$. We define $m^{\Theta_k}(x,t)$ to be the temperature pdf conditional on the initial temperature falling in the set Θ_k. For Θ_k, $k = 1,\dots,K$, the initial conditional pdf $m^{\Theta_k}(x,0)$ is supported and uniform over the intervals Θ_k. Moreover, we attach to a set Θ_k a single control policy, taken to be the one associated with the particle initially at the mean temperature $\bar{\theta}_k$ within Θ_k. In practice we compute $\bar{\theta}_k$ as the empirical mean temperature of the agents falling within Θ_k.

We can then compute $m^{\Theta_k}(x,t)$ using the advection equation

$$\frac{\partial m^{\Theta_k}(x,t)}{\partial t} + \frac{\partial}{\partial x}\{m^{\Theta_k} v^{\bar{\theta}_k}\} = 0, \tag{5}$$

where $v^{\bar{\theta}_k} := f_k(x, u_*^{\bar{\theta}_k}) = -a(x^* - \bar{\theta}_k) + b u_*^{\bar{\theta}_k}$. The optimal control $u_*^{\bar{\theta}_k}$ is obtained from the HJB equation

$$-\frac{\partial V_k}{\partial t} = \inf_{u\in\mathcal{U}^{\bar{\theta}_k}} \left(L_k(x,u,t) + \frac{\partial V_k}{\partial x} f_k(x,u) \right). \tag{6}$$

Finally, in order to calculate the mean $\bar{x}$ of the mass, we have

$$\bar{x} = \int_{-\infty}^{+\infty} x m(x,t) dx, \tag{7}$$

where the mean-field $m(x,t)$ is computed as

$$m(x,t) = \frac{1}{K}\sum_{k=1}^{K} m^{\Theta_k}(x,t) \left(\int_{\Theta_k} \rho_0(\alpha) d\alpha \right).$$

The MFG equation system to solve consists of (3), (5), (6) and (7), with $k = 1,\dots,K$. A fixed point must be found for this system, in the sense that starting from a posited family of $m^{\Theta_k}(x,t)$ flows, one recovers the same $m^{\Theta_k}(x,t)$ when solving the advection equations under the associated optimal control laws for each k.

III. Structure of the Optimal Policies

Note that in (4), $L(x^i, u^i, u^{-i}, t)$ is a time variant function. For our further analysis, we perform a change of variable to make it time invariant. Denote $\tilde{x}^i = e^{-\frac{\delta t}{2}} x^i, \tilde{x}_0^i = e^{-\frac{\delta t}{2}} x_0^i, \tilde{z} = e^{-\frac{\delta t}{2}} z$, and $\tilde{u}^i = e^{-\frac{\delta t}{2}} u^i$. Then we can rewrite (4) using the new variables $(\tilde{x}^i, \tilde{u}^i)$, which gives

$$-\frac{\partial \tilde{V}_i}{\partial t} = \inf_{\tilde{u}^i\in\tilde{\mathcal{U}^i}} \left(\tilde{L}(\tilde{x}^i,\tilde{u}^i) + \frac{\partial \tilde{V}_i}{\partial \tilde{x}^i} f(\tilde{x}^i,\tilde{u}^i) \right), \tag{8}$$

where

$$\begin{aligned} \tilde{L}(\tilde{x}^i,\tilde{u}^i) &= \frac{1}{2}[(\tilde{x}^i - \tilde{z})^2 q_t + (\tilde{x}^i - \tilde{x}_0^i)^2 q_0 + (\tilde{u}^i)^2 r], \\ f(\tilde{x}^i,\tilde{u}^i) &= -a(\tilde{x}^i - \tilde{x}_0^i) + b\tilde{u}^i \\ \tilde{\mathcal{U}^i} &= \{\tilde{u}^i | \tilde{u}_{\min} - \tilde{u}^i_{free} \le \tilde{u}^i \le \tilde{u}_{\max} - \tilde{u}^i_{free}\} \end{aligned}$$

In general, assuming q_t is a given function (itself depending on a posited $m(x,t)$), an optimal solution $(\tilde{u}_*^i, \tilde{x}_*^i)$ to (8) can be obtained numerically using dynamic programming. Let $(\tilde{u}_*^i, \tilde{x}_*^i)$ be an optimal solution to (8), we can define an *costate variable* $\tilde{p}^i$ and the Hamiltonian function $\tilde{H}(\tilde{p}^i, \tilde{x}_*^i, \tilde{u}_*^i)$ such that

$$\begin{aligned} \tilde{p}^i &\triangleq \frac{\partial}{\partial \tilde{x}^i} \tilde{V}_i(\tilde{x}^i, t), \\ \tilde{H}(\tilde{p}^i, \tilde{x}_*^i, \tilde{u}_*^i) &= \tilde{L}(\tilde{x}_*^i, \tilde{u}_*^i) + \tilde{p}^i f(\tilde{x}_*^i, \tilde{u}_*^i). \end{aligned} \tag{9}$$

Then for all admissible controls $\tilde{u}^i \in \tilde{\mathcal{U}^i}$, we must have

$$\tilde{H}(\tilde{p}^i, \tilde{x}_*^i, \tilde{u}_*^i) \le \tilde{H}(\tilde{p}^i, \tilde{x}_*^i, \tilde{u}^i). \tag{10}$$

Hence an optimal solution to (8) also satisfies (9) and (10), which are conditions for Pontryagin's Maximum Principle (PMP). As PMP accommodates constraints more easily, we will look for control solution in an analytic format based on PMP conditions. From (9) and (10), we get

$$(\tilde{u}_*^i)^2 r + b\tilde{p}^i \tilde{u}_*^i \le (\tilde{u}^i)^2 r + b\tilde{p}^i \tilde{u}^i. \tag{11}$$

The inequality in (11) can be expressed in terms of (u^i, x^i, p^i) using the earlier change of variables. By denoting $\tilde{\pi}^i = \pi^i$ and $\tilde{s}^i = e^{-\frac{\delta t}{2}} s^i$, the following control law for $u^i \in \mathcal{U}^i$ can be obtained

$$u_*^i = \begin{cases} u_-^i, & h(p^i) < u_-^i \\ h(p^i), & h(p^i) \in [u_-^i, u_+^i] \\ u_+^i, & h(p^i) > u_+^i \end{cases}, \tag{12}$$

where for an agent i, $h(p^i) = -br^{-1}p^i$, $u_-^i = u_{\min} - u^i_{free}$ is the lower bound of the constraint $\mathcal{U}^i$, and $u_+^i = u_{\max} - u^i_{free}$ is the upper bound of the constraint.

An exact optimal control law can be derived by following some PMP-based approaches (see [11], [12] for example), but the procedure to compute such optimal solution is complex.

A. Single Switching Control Policies

To continue our analysis we implement a control policy where controls can switch from saturation to unsaturation at most once. For a linear quadratic (LQ) tracking problem, the costate $p^i(t)$ is of the form

$$p^i(t) = \pi^i(t) x^i(t) + s^i(t),$$

where π^i and s^i are two functions to determine.

From PMP, $p^i(t)$ should satisfy the costate equation

$$\dot{p}^i = (a+\delta)p^i - q_t(x^i - z) - q_0(x^i - x_0^i). \tag{13}$$

While based on (12), on a given q_t path, control could go in theory from unsaturation to saturation, and then back to unsaturation several times, we shall impose conditions on q_t that would make $p^i(t)$ monotonic and thus would allow saturation, if any, only at the start of the control horizon until some time t_*^i. Past t_*^i the optimal control becomes unsaturated for the rest of the control horizon. Considering without loss of generality the temperature decrease case, we thus require $\frac{dp^i}{dt} \le 0$. From (13), this imposes a condition on q_t whereby,

$$q_t \ge \frac{(a+\delta)p^i}{x^i - z} + q_0 \frac{x_0^i - x^i}{x^i - z}. \tag{14}$$

In what follows, we produce a lower bound q_t^- on q_t and an upper bound for the right hand side of (14). In order to do so, we shall assume that δ is small enough that the steady-state mean temperature trajectory $\bar{x}$ asymptotically reaches the target y. At such steady-state, q_t reaches a constant value q_t^∞. It turns out that q_t^∞ is the unique constant weight such that the mean of all individual initial condition dependent temperature steady-states is equal to y.

Proposition 3.1: The unique q_t^∞ can be computed by

$$q_t^\infty = c_1 \frac{\bar{x}_0 - y}{y - z}, \tag{15}$$

where $c_1 = \frac{a(a+\delta)r + b^2 q_0}{b^2}$.

To see this, first note that the steady-state at y is inconsistent with control saturation, because all trajectories are monotone decreasing (in temperature decrease case), and persistent saturated controls would send temperatures to the ambient temperature x_a where $x_a < z$. So at equilibrium, one can rely on the steady-state equations from the unconstrained case.

In the unconstrained case, π^i and s^i must satisfy the following Riccati equation [13]

$$\begin{aligned} \dot{\pi}^i &= (2a+\delta)\pi^i - q_t - q_0 + b^2 r^{-1} \pi^{i2}, \\ \dot{s}^i &= (a + \delta + b^2 r^{-1}\pi^i) s^i - a\pi^i x_0^i + q_t z + q_0 x_0^i, \end{aligned} \tag{16}$$

By substituting the optimal control into (1), we get,

$$\dot{x}^i = (-a - b^2 r^{-1} \pi^i) x^i - b^2 r^{-1} s^i + a x_0^i. \tag{17}$$

At steady-state, we have $\dot{\pi}^i = 0$, $\dot{s}^i = 0$, and $\dot{x}^i = 0$, and solving (16) and (17) gives,

$$q_t^\infty = \frac{a(a+\delta)r + b^2 q_0}{b^2} \frac{x_0^i - x_\infty^i}{x_\infty^i - z}, \tag{18}$$

where x_∞^i denotes the steady-state temperature for agent i.

Given that at steady-state, the mean temperature $\bar{x}_\infty = y$, we can compute q_t^∞ in terms of $\bar{x}_0$ and y by taking expected values of both sides in (18), and get the result in (15).

By applying a constant q_t^∞ on any initial condition x_0^i, we can produce a state trajectory $x_-^i(t)$ and a costate trajectory $p_-^i(t)$, such that

$$\frac{(a+\delta)p_-^i}{x_-^i - z} + q_0 \frac{x_0^i - x_-^i}{x_-^i - z} \geq \frac{(a+\delta)p^i}{x^i - z} + q_0 \frac{x_0^i - x^i}{x^i - z}. \tag{19}$$

By applying the same constant q_t^∞ to $\bar{x}_0$, we produce a trajectory $\bar{x}_0^-(t)$ below $\bar{x}_0(t)$ such that,

$$q_t^- + k_c = \lambda \left| \int_0^t (\bar{x}_0^-(\tau) - y) d\tau \right| + k_c < q_t. \tag{20}$$

Note that this is in view of the monotonicity of the best response to the level of the weighting term of $(x^i - z)^2$ in the cost function. Therefore by combining (14), (19), and (20) we can write a sufficient condition under which $p^i(t)$ is monotonically decreasing for the temperature decrease case.

$$\begin{aligned} q_t + k_c > q_t^- + k_c &\geq \frac{(a+\delta)p_-^i}{x_-^i - z} + q_0 \frac{x_0^i - x_-^i}{x_-^i - z} \\ &\geq \frac{(a+\delta)p^i}{x^i - z} + q_0 \frac{x_0^i - x^i}{x^i - z}. \end{aligned} \tag{21}$$

For the rest of the paper, we assume that the above condition is verified for all initial conditions and all times, which would produce single switching behaviors for the best responses.

For a given q_t that meets the condition in (21), and for any initial condition x_0^i, the control will be saturated by u_-^i at $t = 0$ until some switching point t_*^i. From t_*^i to the end of the control horizon, control is unsaturated and follows the optimal policy $-br^{-1}p^i$. The t_*^i can be determined whenever $-br^{-1}p^i(t_*^i) = u_-^i$ is satisfied. For some agent i, we could have $\left|-br^{-1}p^i(0)\right| \leq u_-^i$ at $t = 0$, in which case $t_*^i = 0$ and controls are unsaturated all the time.

IV. Computation of the Equilibrium Solution

In this section, we analyze and find the equilibrium solution $(m(x,t), \{u_*\})$ that solves the MFG system.

A. Discretization and Numerical Solution of Advection Equation

It is recognized that the advection equation in (5) is a balance equation and conservative. Therefore, if one wishes to conserve probability, a discretization based on finite volume methods is recommended [14]. By discretizing the $x-t$ plane with time-step $k = \Delta t$ and x-step $h = \Delta x$, we define a grid with points (x_j, t_n) where $x_j = jh, j = 0, 1, 2, \ldots$ and $t_n = nk, n = 0, 1, 2, \ldots$. The solution $m^{\Theta_k}(x,t)$ should satisfy the integral form of the conservation laws. Denote $g(m^{\Theta_k}) := m^{\Theta_k} v^{\bar{\theta_k}}$. Then we have

$$\int_{x_{j-1/2}}^{x_{j+1/2}} m^{\Theta_k}(x, t_{n+1}) dx = \int_{x_{j-1/2}}^{x_{j+1/2}} m^{\Theta_k}(x, t_n) dx - \left[\int_{t_n}^{t_{n+1}} g(m^{\Theta_k}(x_{j+1/2}), t) dt - \int_{t_n}^{t_{n+1}} g(m^{\Theta_k}(x_{j-1/2}), t) dt \right]. \tag{22}$$

We define M_j^n as an approximation to the cell average of $m^{\Theta_k}(x, t_n)$ for $x \in [x_{j-1/2}, x_{j+1/2})$ at t_n, and the numerical flux $F(M_j^n, M_{j+1}^n)$ which is the average "probability current" through $x_{j+1/2}$ over the time interval $[t_n, t_{n+1}]$.

$$\begin{aligned} M_j^n &\simeq \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} m^{\Theta_k}(x, t_n) dx, \\ F(M_j^n, M_{j+1}^n) &\simeq \frac{1}{k} \int_{t_n}^{t_{n+1}} g(m^{\Theta_k}(x_{j+1/2}), t)) dt. \end{aligned}$$

Accordingly the integral form of the conservation laws in (22) can be expressed as

$$M_j^{n+1} = M_j^n - \frac{k}{h}[F(M_j^n, M_{j+1}^n) - F(M_{j-1}^n, M_j^n)]. \tag{23}$$

Using a finite difference method such as Lax-Friedrichs or Lax-Wendroff [14], we can numerically solve for M_j^n which is approximation to the solution $m^{\Theta_k}(x,t)$ on the defined grid. For example, if we use the Lax-Friedrichs scheme, which has the following form

$$M_j^{n+1} = \frac{1}{2}\left(M_{j+1}^n + M_{j-1}^n\right) - \frac{k}{2h}\left(g(M_{j+1}^n) - g(M_{j-1}^n)\right).$$

We can therefore write the expression $F(M_j^n, M_{j+1}^n)$ as

$$F(M_j^n, M_{j+1}^n) = \frac{h}{2k}(M_j^n - M_{j+1}^n) + \frac{1}{2}\left(g(M_j^n) + g(M_{j+1}^n)\right). \tag{24}$$

Using initial density $M_0^0 = m^{\Theta_k}(x_0, 0)$, we can get the propagation of an approximated solution M_j^n on the defined grid points (x_j, t_n) from (23) and (24).

B. Equilibrium Solution by Iteration Approach

As the equilibrium solution is a fixed point of the MFG system, we can compute it using an iterative approach. Given initial mass density $m(x_0, 0)$, we first make an arbitrary guess of a family of $m^{\Theta_k}(x,t)^0$ as the fixed point. Then we compute a q_t^0 from $m^{\Theta_k}(x,t)^0$. We inject the computed q_t^0 into the MFG system and get a new candidate fixed point $m^{\Theta_k}(x,t)^1$ and a new q_t^1 by following the optimal control law and solving the advection equations. We repeat the procedure until $m^{\Theta_k}(x,t)$ converges to a fixed point, from which we get the equilibrium solution $(m^*(x,t), \{u_*\})$. During each iteration, each q_t candidate obtained should satisfy the condition in (21), in order to guarantee the single switching behaviors of the control law.

V. Simulations

In the numerical study, we make the following assumptions. The distribution of initial temperature is uniform between 18 and 28 degrees, hence the initial mean temperature is 23 degrees. We wish to achieve a target mean temperature y of 22 degrees over a 4 hours horizon, and set parameter z in (2) to 17 degrees so every device will tend to decrease their temperature. Other parameters used in the simulation are as follows: $a = 0.03, b = 0.2, r = 1, q_0 = 200, \lambda = 40, k_c = 5.207$. The constraint we impose is that the total control $(u^i + u^i_{free})$ must be always greater than zero, as it is inadmissible to have a negative control to cool down the space. Hence for u^i, the constraint is $u^i \geq -u^i_{free}$.

We initialize the iteration process by using q_t^-. Figure 1 illustrates the evolution of q_t at each iteration which is constructed from the solutions to the MFG system until convergence. In all iterations, the condition in (21) is respected.

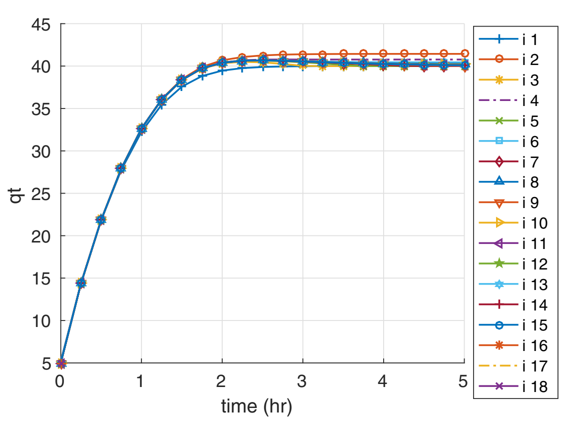

Fig 1. Iteration approach to find equilibrium in terms of q_t

Figure 2 shows the propagation of mass density $m^*(x,t)$ computed from $m^{\Theta_k}(x,t)$ at the converged equilibrium. We start from an initial mass with uniform distribution, and at steady state the mass shifts towards lower temperature and we arrive at another uniform distribution. Note that we are using the Lax-Friedrichs scheme, which is a first order finite volume method and the edges of the uniform distribution at steady- state are smeared off. By using a finer $x - t$ grid plane, we can reduce the smearing effects such that steady-state mass shapes more similar to a uniform distribution, but the CFL (Courant–Friedrichs–Lewy; see [14]) condition must always be satisfied in order to ensure stability of the numerical method. Also, note that the partition Θ_k of initial temperatures should be sufficiently refined in order to avoid significant ripples in the transient mass probability density.

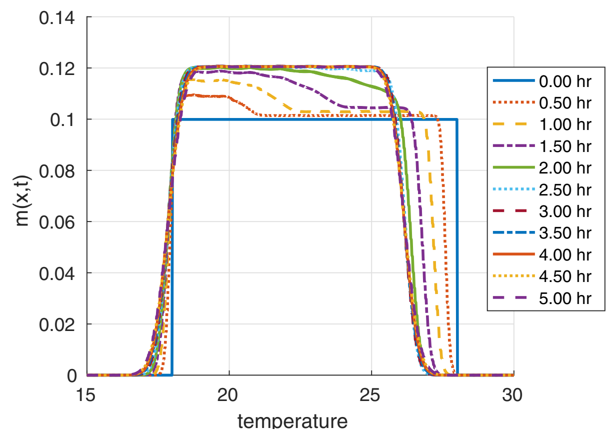

Fig 2. Propagation of mass density at equilibrium solution

In Figure 3 we take a few particle samples whose initial temperatures are randomly taken from the initial uniform distribution. We show the single switching behaviors of control trajectories $\{u_*^i\}$, and the corresponding state trajectories $\{x_*^i\}$ at equilibrium. From the setup of the quadratic cost function in (2), the particles with higher initial temperatures feel more pressure to quickly decrease the temperature; however, due to the presence of the control constraint, their rates of temperature decrease are constrained at $t = 0$ (trajectories in red color). On the other hand, particles starting from lower initial temperatures feel less stress to decrease temperature, and their controls are not affected by the control constraint (trajectories in blue color). It is seen that the mean temperature of the population remains at the desired target temperature of 22 degrees at steady state.

The derived model can also be used in the unconstrained case by simply removing the constraint and let $u^i \in \mathbb{R}, \ \forall i$. It is verified that under the same iteration approach we get an equilibrium solution identical to that in [3].

In the next experiment, we wish to explore the difference in state trajectories under constrained versus unconstrained cases. In particular we are interested in the mean temperature and particle samples starting from the extreme initial temperatures. Figure 4 shows the comparison results. For particles starting from 28 degrees, as the imposed control constraints limit their rates of decrease since the start of the control horizon, the temperatures do not fall as fast as they do in the unconstrained case. As a consequence, q_t increases due to a larger error recorded between $\bar{x}$ and y, as compared to the unconstrained case at the same point in time. With a larger q_t, particles starting from 18 degrees are pressured to contribute more to reduce the global mean temperature. However, they do not feel enough stress to completely make up the gap caused by the constrained particles. Once controls become unsaturated, the constrained particles continue to decrease more than they would do in the unconstrained case, while the particles with low initial temperatures start to contribute less correspondingly. Eventually at steady-state, *all particles including the mean temperature have the same values as those in the unconstrained case*. This result is expected as all controls fall within the saturation limits at steady-state.

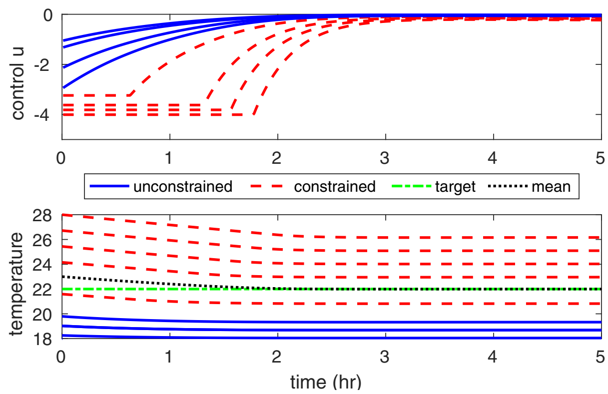

Fig 3. Trajectories of controls **(top)** and states **(bottom)** for particle samples from the mass

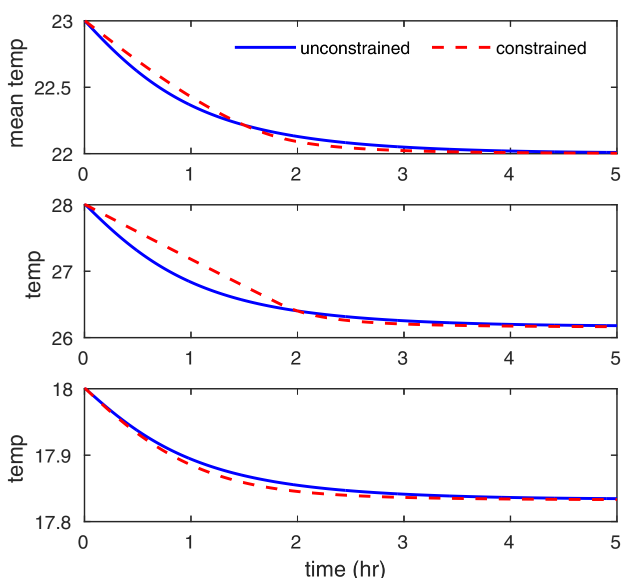

Fig 4. State trajectories comparisons of control constrained vs. unconstrained cases: mean temperature trajectories **(top)**; particle with high initial temperature **(middle)**; particle with low initial temperature **(bottom)**

VI. Concluding Remarks

The non-linearity of the optimal control law under control constraints makes the computation of the MFG fixed point more challenging than in the linear quadratic case, requiring the solution of coupled PDE's, which moreover depends on the initial conditions of the agents. We describe a feasible optimal control policy with single switching behaviors and an algorithm to compute an approximation of the fixed point. We verify that the steady-states in the constrained case are identical to those in the unconstrained case. For future work, the existence and uniqueness of the fixed point of the constrained MFG equation system remains to be proved, while the sufficient condition in (21) must be further analyzed. Also, we wish to extend the analysis to the stochastic case where a noise process is included in (1). Finally, extensions of the analysis to the multidimensional case, including non diffusion models (electric water heating loads) would be of interest.

References

[1] B. Tarroja, F. Mueller, J. D. Eichman, and S. Samuelsen, "Metrics for evaluating the impacts of intermittent renewable generation on utility load-balancing," *Energy*, vol. 42, no. 1, pp. 546–562, 2012.

[2] 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, no. 5, pp. 1389–1400, 2009.

[3] A. C. Kizilkale and R. P. Malhamé, "Mean field based control of power system dispersed energy storage devices for peak load relief," in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*, pp. 4971–4976, Dec 2013.

[4] A. C. Kizilkale and R. P. Malhamé, "A class of collective target tracking problems in energy systems: cooperative versus non-cooperative mean field control solutions," in *2014 IEEE 53rd Annual Conference on Decision and Control (CDC)*, pp. 3493–3498, Dec 2014.

[5] A. C. Kizilkale and R. P. Malhamé, "Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads," in *IFAC World Congress*, 2014.

[6] M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized epsilon-Nash Equilibria," *IEEE Transactions on Automatic Control*, vol. 52, pp. 1560–1571, Sept 2007.

[7] F. Bagagiolo and D. Bauso, "Mean-Field Games and Dynamic Demand Management in Power Grids," *Dynamic Games and Applications*, vol. 4, no. 2, pp. 155–176, 2013.

[8] D. Paccagnan, M. Kamgarpour, and J. Lygeros, "On the range of feasible power trajectories for a population of thermostatically controlled loads," in *2015 IEEE 54th Annual Conference on Decision and Control (CDC)*, Dec. 2015.

[9] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash Equilibria in constrained mean field control," *CoRR*, vol. abs/1410.4421, 2014.

[10] 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," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[11] W. Heemels, S. Van Eijndhoven, and A. Stoorvogel, "Linear quadratic regulator problem with positive controls," *International Journal of Control*, vol. 70, no. 4, pp. 551–578, 1998.

[12] R. Goebel and M. Subbotin, "Continuous time constrained linear quadratic regulator - convex duality approach," in *American Control Conference, 2005. Proceedings of the 2005*, pp. 1401–1406 vol. 2, June 2005.

[13] A. E. Bryson, *Applied optimal control: optimization, estimation and control*. CRC Press, 1975.

[14] R. J. LeVeque, *Numerical methods for conservation laws*, vol. 132. Springer, 1992.

[15] D. E. Kirk, *Optimal control theory: an introduction*. Prentice-Hall networks series, Prentice-Hall, 1970.

[16] P. Cardaliaguet, "Notes on mean field games," tech. rep., 2010.