Dyn Games Appl (2014) 4:155–176
DOI 10.1007/s13235-013-0097-4

Mean-Field Games and Dynamic Demand Management in Power Grids

Fabio Bagagiolo · Dario Bauso

Published online: 12 November 2013

Abstract This paper applies mean-field game theory to dynamic demand management. For a large population of electrical heating or cooling appliances (called agents), we provide a mean-field game that guarantees desynchronization of the agents thus improving the power network resilience. Second, for the game at hand, we exhibit a *mean-field equilibrium*, where each agent adopts a bang-bang switching control with threshold placed at a nominal temperature. At equilibrium, through an opportune design of the terminal penalty, the switching control regulates the mean temperature (computed over the population) and the mains frequency around the nominal value. To overcome Zeno phenomena we also adjust the bang-bang control by introducing a thermostat. Third, we show that the equilibrium is stable in the sense that all agents' states, initially at different values, converge to the equilibrium value or remain confined within a given interval for an opportune initial distribution.

Keywords Mean field games · Dynamic demand management · Viscosity solutions · Distributional solutions

1 Introduction

This paper applies mean-field game theory to dynamic demand management in the same spirit as [14, 37]. The latter is a recent and promising research area aiming at improving

F. Bagagiolo · D. Bauso (✉)
Dipartimento di Matematica, Università di Trento, Via Sommarive 14, 38050 Povo-Trento, Italy
e-mail: dario.bauso@unipa.it

F. Bagagiolo
e-mail: bagagiol@science.unitn.it

D. Bauso
DICGIM, Università di Palermo, Viale delle Scienze, 90128 Palermo, Italy

Present address:
D. Bauso
Department of Engineering Science, University of Oxford, Oxford, UK

resilience in power networks [3, 15, 16]. In a nutshell, the functioning of a power network is characterized by the system frequency, also called the *mains frequency*, which represents an indicator of the balance level between energy demand and supply. This frequency usually needs to be stabilized around a nominal value (50 Hz in Europe). If electrical demand exceeds generation then frequency will decline, and vice versa.

In this context, dynamic demand management aims at assigning part of the regulation burden to the consumers by using "frequency responsive" appliances. In other words, each appliance regulates automatically and in a decentralized fashion its power demand based on the mains frequency. A similar concept characterizes a recent literature on "load control" in power systems [12, 14, 26, 28, 37]. In particular, [12] surveys issues related to the redistribution of the load away from peak hours and the design of decentralized strategies to produce a predefined load trajectory (see also [14]). A main challenge of decentralized control is that local controllers may give rise to conflicts while pursuing their goals. Indeed, local decisions may result in an over- or undersupply of the required response. Thus the aforementioned conflicting objectives have led scientists to adopt noncooperative games as paradigmatic models. In [26] the authors present a large population game where the agents are plug-in electric vehicles and the Nash-equilibrium strategies (see [9]) correspond to distributed charging policies that redistribute the load away from peaks (called valley-filling strategies). In this paper we adopt the same perspective in that we show that network frequency stabilization can be achieved by giving incentives to the agents to adjust their strategies in order to converge to a mean-field equilibrium. To do this, in the spirit of *prescriptive game theory* [6], a central planner or game designer has to design the individual objective function so to penalize those agents that are in the ON state in peak hours, as well as those who are in the OFF state in off-peak hours.

Valley-filling and coordination strategies have been shown particularly efficient in thermostatically controlled loads such as refrigerators, air conditioners and electric water heaters [28]. Thermostatically controlled loads are also the focus of the present paper. Indeed, in most cases the capability for these electric heating or cooling appliances (henceforth simply called *appliances*) of storing thermal energy is greater than the capability of a battery of storing chemical energy.

The results obtained in this paper are in accordance with the recent results in [3], according to which, stochastic control laws are in general more appropriate than deterministic ones when it comes to desynchronize the appliances functioning.

Highlights of Main Results This paper presents three main contributions. First, it provides a mean-field game which captures the interactions among a large number of appliances. Each single agent is characterized by a temperature and can be in one of the two states ON or OFF. The dynamics of a single agent describes the time evolution of its temperature, and takes on the form of a linear ordinary differential equation. In addition to this, each agent is given a cost function that accounts for (i) energy consumption, (ii) deviation of mains frequency from the nominal one, and (iii) deviation of the agent's temperature from reference value. With respect to item (ii), we introduce in the cost function a mean-field term that incentivizes the agent to switch OFF if the mains frequency is below the nominal value and to switch ON if the mains frequency is above the nominal value.

A main feature of the game at hand is that it gives rise to the formation of atomic parts in the distribution of temperatures. As such, the game can describe both the case with a continuum of agents and the case with a finite number of agents. While in the former case, the resulting mean-field game should be justified by a limit procedure of Nash equilibria of games with a finite number of players, in the latter case such a limit procedure is no longer needed. On one hand, this simplifies the tractability as, except for some special cases, such a procedure is still far from being understood and satisfactory, see [13] for a comprehensive account. On the other hand, the formation of atomic parts in the distribution requires a suitably modified weak solution concept, which is formally defined in Sect. 3.

In addition, the provided game differs from a standard mean-field game in [13] in at least two other aspects: first both controls and states are bounded, and second we have an additional cross-term on distribution and controls in the objective function, which is not monotonic on distributions.

As a second contribution, for the mean-field game at hand, we compute a *mean-field equilibrium* and show that at the equilibrium each agent adopts a bang-bang-like switching control with threshold placed at the nominal temperature.

Through an opportune design of the final penalty of value function, we show that the equilibrium regulates the mean temperature (computed over the population) and the means frequency as well around the nominal value. By doing this, we address two main issues: one is related to the macroscopic behavior of the system and the other one involves the microscopic behavior. The first issue accounts for the synchronization of the appliances which is recognized as the root cause of the mains frequency oscillation. At the equilibrium each agent switches to ON with probability 1/2 and this gives a stochastic flavor to the implemented control law in agreement with the results in [3]. The second issue regards the so-called Zeno phenomenon which is common to many switching control problems. To overcome such a phenomenon we adjust the bang-bang control by introducing a static nonlinearity in the form of a hysteresis. To do this we expand the state space by adding an additional state variable that accounts for the number of switches up to the current time.

A third contribution analyzes the system behavior around the equilibrium. Under certain assumptions, we show that the equilibrium is stable in the sense that all agents states, initially at different values, converge to the equilibrium value or remain confined within a given interval.

Related literature on mean-field games The mean-field theory of dynamical games with large but finite populations of asymptotically negligible agents (as the population size goes to infinity) originated in the work of M.Y. Huang, P.E. Caines and R. Malhamé [18–20] and independently in that of J. M. Lasry and P.L. Lions [23–25], where the now standard terminology of Mean Field Games (MFG) was introduced. In addition to this, the closely related notion of Oblivious Equilibria for large population dynamic games was introduced by G. Weintraub, C. Benkard, and B. Van Roy [35] in the framework of Markov Decision Processes.

Mean-field games arise in several application domains such as economics, physics, biology, and network engineering (see [1, 17, 20, 22, 37]).

With regard to network engineering, mean-field games have been applied to medium access control in wireless networks, resource allocation problem, flow control, congestion management, demand and price formation in the power grid market, energy management in cloud computing, consensus and synchronization problems, state estimation, etc. More details can be found in [31]. Mean-field game formulations apply also to multi-inventory systems with quadratic cost and additional set-up costs, as discussed in [29]. Decision problems with mean-field coupling terms have also been formalized and studied in [11]. An example from production engineering has been first introduced by [17] (see also [31]). A robust formulation of the production problem is also available in [10].

From a mathematical point of view, the mean-field approach leads to a study of a system of two partial differential equations (PDEs), The first PDE is the Hamilton–Jacobi–Bellman equation which is usually solved backwards in time with penalty on final state and distribution (suppose a finite horizon formulation). The second PDE is the Fokker–Planck equation which describes the density of the players and is solved forwards in time with boundary conditions on the initial population distribution (see, e.g., the Fokker–Planck–Kolmogorov equation in [2, 25, 32, 36] and in the lecture notes [13] and M.S. thesis [27]).

Explicit solution in terms of mean-field equilibria are not common unless the problem has a linear–quadratic structure, see [7]. In this sense, a variety of solution schemes have been recently proposed based on discretization and or numerical approximations. In [2], for instance, a fully discrete finite difference approximation scheme of the coupled system has been proposed and studied.

Mean-field games have connections to another stream of literature: evolutionary games ([21, 30, 33]). Here, the so-called *anonymous games* and *aggregative games* build upon the notion of mass interaction and can be seen as a stationary mean field. A dynamic discrete time version of the mean-field game has been studied by [21] where a fundamental mean-field system consisting of value function and mean-field evolution was proposed. This corresponds to a backward-forward system. The equation satisfied by the value is essentially a Bellman equation and the equation satisfied by the mean-field term is a Kolmogorov equation. In [21] the authors provide sufficiency conditions for existence of solutions to such systems.

More recently, robustness has been brought into the picture. Robust mean-field games aim to achieve robust performance or stability in the presence of unknown disturbances when there is a large number of players; see [32]. Still in [32] relations with risk-sensitive games and risk-neutral games have been analyzed in [32].

The rest of the paper is organized as follows. In Sect. 2 we illustrate the problem and introduce the model. In Sect. 3 we address the concept of weak solution for the resulting mean-field system. In Sect. 4 we exhibit a mean-field equilibrium for the problem at hand. In Sect. 5 we introduce a thermostat in the switching control law. In Sect. 6 we analyze the system behavior around the equilibrium. In Sect. 7 we provide numerical examples. Finally, in Sect. 8 we draw some conclusions.

2 Model and Problem Set-up

Consider a population of homogeneous electric appliances (players), each one characterized by a temperature $X(s)$ at time $t \le s \le T$, where $[t, T]$ is the time horizon window. The control variable is a measurable function of time $\pi_{\mathrm{ON}}(\cdot)$ defined as $s \mapsto \{0, 1\}$ and such that $\pi_{\mathrm{ON}}(s) = 1$ means that, at time s, the appliance is set to *ON* and $\pi_{\mathrm{ON}}(s) = 0$ means that the appliance is set to *OFF*.

When the appliance is ON the temperature decreases exponentially up to a fixed lower temperature whereas in OFF position the temperature increases exponentially up to a higher temperature. Then, the temperature of each appliance evolves according to the following differential equations:

$$X'(s) = \begin{cases} -\alpha(X(s) - X_{\mathrm{ON}}), & \text{if } \pi_{\mathrm{ON}}(s) = 1 \\ -\alpha(X(s) - X_{\mathrm{OFF}}), & \text{if } \pi_{\mathrm{ON}}(s) = 0 \end{cases}, \quad t < s < T \tag{1}$$

with initial state $X(t) = x$ and where $\alpha > 0$ is a given scalar (the rate) and X_{ON}, X_{OFF} are the steady-state temperatures of the appliances when in state ON or OFF, respectively. Here, considering a same rate for the two states has the only meaning of simplifying future computations. We will see that a different rate, though more realistic, adds no value as it affects in no ways the solution approach.

2.1 Control Design

The control $\pi_{ON}(t)$ has to satisfy the following requirements, which will appear as additive terms in the cost function to minimize:

- minimization of power: $W_{ON}\pi_{ON}(s) + W_{OFF}(1-\pi_{ON})(s)$ where W_{ON} and W_{OFF} are the power consumed when the appliance is *ON* or *OFF*, respectively.
- network frequency stabilization: denoting by $[w(s)-w_{ref}]_\pm$ the positive or negative scalar part of the difference between the current frequency $w(s)$, and the reference frequency w_{ref}, frequency stabilization corresponds to a cost of type $\pi_{ON}[w(s)-w_{ref}]_+ + (1-\pi_{ON})[w(s)-w_{ref}]_-$ where w is the current frequency, w_{ref} the reference frequency. Intuitively, the network frequency w depends on the number of appliances in the ON position at time s. In general, for any scalar ξ, we denote by $[\xi]_\pm$ its positive or negative scalar part. To see why the above cost incentivizes frequency stabilization observe that the first term $\pi_{ON}[w(s)-w_{ref}]_+$ represents a penalty for all those agents that are in the ON state when $w(s) > w_{ref}$ (load exceeds nominal generation) while the second term $(1-\pi_{ON})[w(s)-w_{ref}]_-$ is a penalty for all those agents that are in the OFF state when $w(s) < w_{ref}$ (load is less than nominal generation).
- stabilization of the temperature around a comfortable value $x_{ref} := 0$. Note that taking $x_{ref} \neq 0$ is without loss of generality because i) we can always translate the axis to x_{ref} without compromising the modeling and solution approach.

Let us convexify the control set and consider the control of a single agent as the probability of setting the appliance to *ON*, thus we have $u(t) \in U := [0, 1]$ where U is the control set. It turns out that the dynamics (1) can be rewritten in the form $X' = f(X, u)$ where $f : \mathbb{R} \times U \to \mathbb{R}$ is the following affine dynamics:

$$\begin{cases} X'(s) = -\alpha X(s) + \sigma u(s) + c, & s > t, \\ X(t) = x, \end{cases} \tag{2}$$

where $x \in [X_{ON}, X_{OFF}]$, $t \in [0, T]$ are the initial state and the initial time, respectively, $\sigma := -\alpha(X_{OFF} - X_{ON})$, $c := \alpha X_{OFF}$. For sake of simplicity and without loss of generality we will take $X_{OFF} = -X_{ON}$. Indeed, we can always select lower and upper bounds of the temperature symmetric with respect to x_{ref}. In addition, note that the closed set $[X_{ON}, X_{OFF}]$ is invariant and that the two extremes are not reachable from any other interior point. Hence, it is not restrictive to assume that no appliances have the temperatures X_{ON} and X_{OFF}.

In light of the considerations provided above, and in order to introduce a macroscopic description of the game, consider a probability density function $m : [X_{ON}, X_{OFF}] \times [0, +\infty[\to [0, +\infty[$, $(x, s) \mapsto m(x, s)$ (in Sect. 3 we will weaken the regularity of such a density measure), which satisfies $\int_{X_{ON}}^{X_{OFF}} m(x, s)\, dx = 1$ for every s. Let us also define the mean temperature at time s as $\overline{m}(s) := \int_{\mathbb{R}} x m(x, s)\, dx$. Also, as any trajectory obtained from (2) can never reach X_{ON} and X_{OFF}, we can assume that $m(X_{ON}, s) = m(X_{OFF}, s) = 0$ for all s.

At every time s the network frequency $w(s)$ depends linearly on the mean temperature computed over all appliances, i.e., $w(s) - w_{ref} = -(\overline{m}(s) - \overline{m}_{ref})$, (the higher the mean temperature $\overline{m}(s)$, the lower the network frequency $w(s)$).

Observe that we simply assume that a discrepancy between the average (over appliances) and the desired temperature, $(\overline{m}(s) - \overline{m}_{ref})$ induces a discrepancy between demand and supply which in turn translates into a deviation of the network frequency from the nominal value. This mutual dependence is usually captured by a higher order dynamics which we decided to approximate using a steady-state relation. We do this as the spirit of the paper is to capture individuals' conflicting and common goals within a large population setting. Therefore, turning the problem into a higher order one would complicate the tractability without adding a value to the approach.

Then $w(s) = w_{\mathrm{ref}}$ implies $\overline{m}(s) = \overline{m}_{\mathrm{ref}}$. In the following, for sake of simplicity and without loss of generality, we take all nominal values equal to zero unless specified differently. We can do this as nonzero nominal values would only add nonzero constant terms to the computations. Then, let us take $\overline{m}_{\mathrm{ref}} = 0$. Also, for given scalars $q, k, h > 0$, take $W_{\mathrm{OFF}}, w_{\mathrm{ref}}, x_{\mathrm{ref}} = 0$, and denote $r := W_{\mathrm{ON}}$; then consider a running cost $g : \mathbb{R} \times U \times \mathbb{R} \to [0, +\infty[$, $(x, u, \overline{m}) \mapsto g(x, u, \overline{m})$ of the form

$$\begin{aligned} g(x, u, \overline{m}) &= u\big(r + qx^2 + h[\overline{m}]_+\big) + (1 - u)\big(qx^2 + k[\overline{m}]_-\big) \\ &= ru + qx^2 + h[\overline{m}]_+ u + k[\overline{m}]_-(1 - u). \end{aligned} \tag{3}$$

Observe that cost (3) includes three main terms. A first penalty term ru which accounts for minimization of power. Second, the term qx^2 which penalizes the deviation of the appliance's temperature from zero (the target value). Third, the term $h[\overline{m}]_+ u + k[\overline{m}]_-(1 - u)$ which accounts for the network stabilization in that it penalizes those appliances that are ON whenever $\overline{m} > 0$, the latter condition meaning that demand exceeds supply. Likewise, it penalizes those appliances that are OFF whenever $\overline{m} < 0$, i.e., whenever supply exceeds demand. Note that the presence in the cost of the coupled terms $h[\overline{m}]_+ u + k[\overline{m}]_-(1 - u)$ makes the model differ from the standard formulation of mean-field games in [13] where the terms depending on m and on the control are decoupled. In addition, here the control set is bounded.

Also consider a terminal cost $\Psi : \mathbb{R} \to [0, +\infty[$, $x \mapsto \Psi(x)$ to be yet designed.

2.2 Problem statement

Given a finite horizon $T > 0$ and an initial distribution of temperatures $m_0 : [X_{\mathrm{ON}}, X_{\mathrm{OFF}}] \to [0, +\infty[$, with mean $\overline{m}_0$, minimize over $\mathcal{U}$, subject to the controlled system (2), the cost functional

$$J\big(x, t, u(\cdot)\big) = \int_t^T g\big(X(s), u(s), \overline{m}(s)\big)\, ds + \Psi\big(X(T)\big),$$

where $\mathcal{U}$ is the set of all measurable functions $u(\cdot)$:$[0, +\infty[\to U$, and $\overline{m}(\cdot)$ is the time-dependent function describing the evolution of the mean of the distribution of temperatures if every one of the agents behaves optimally, i.e. minimizes J.

In the following, to make use of dynamic programming techniques based on Bellman equation, we consider the evolution of the mean $\overline{m}$ as a datum of the optimal control problem and look for a solution of the resulting fixed point problem. In other words, we consider a cost functional of the form

$$g_{\overline{m}(\cdot)} : \mathbb{R} \times U \times [0, +\infty[\to [0, +\infty[, \ (x, u, s) \mapsto g\big(x, u, \overline{m}(s)\big)$$

and restate the problem as follows. For any fixed evolution of $\overline{m}(\cdot)$, minimize, over all measurable controls $u(\cdot) \in \mathcal{U}$ and all corresponding trajectories of (2), the cost functional

$$J_{\overline{m}(\cdot)}\big(x, t, u(\cdot)\big) = \int_t^T g_{\overline{m}(\cdot)}\big(X(s), u(s), s\big)\, ds + \Psi\big(X(T)\big),$$

where $\overline{m}(\cdot)$ is given by the "mean evolution equation"

$$\overline{m}'(t) = -\alpha \overline{m}(t) + \sigma \overline{u}(t) + c, \quad \overline{m}(0) = \overline{m}_0. \tag{4}$$

Here, the "mean control" $\overline{u}(\cdot) : [0, +\infty[\to [0, 1]$ is the mean over the agents of all optimal feedback controls u^* at every time t. The mean evolution equation (4) can be obtained by averaging over all agents the terms appearing in the left-hand side of the first order *Kolmogorov–Fokker–Planck* equation $m_t + (f(x, u^*)m)_x = 0$ (see, e.g., [10]). The Kolmogorov equation describes the evolution of the distribution of the temperatures m when the temperature of each appliance follows the law $X' = f(X, u^*(X, t))$.

We recall that the value function $v = \inf_u J_{\overline{m}}$ of the optimal control problem is a function defined on $[X_{\text{ON}}, X_{\text{OFF}}] \times [0, T]$ as it is the solution of the corresponding Hamilton–Jacobi-Bellman equation for all $(x, t) \in [X_{\text{ON}}, X_{\text{OFF}}] \times [0, T]$ (see for instance [8]).

The problem results in the following mean-field game system (we denote by I, II, III, IV, V the five blocks of the system):

$$\begin{cases} -v_t(x,t) + \sup\limits_{u\in U}\{-f(x,u)v_x(x,t) - g_{\overline{m}(\cdot)}(x,u,t)\} = 0 \\ \quad \text{in } [X_{\text{ON}}, X_{\text{OFF}}]\times]0,T], \\ v(x,T) = \Psi(x) \quad \forall x \in [X_{\text{ON}}, X_{\text{OFF}}], \\ \\ u^*(x,t) = \operatorname*{argmax}\limits_{u\in[0,1]}\{-f(x,u)v_x(x,t) - g_{\overline{m}(\cdot)}(x,u,t)\}, \\ \\ m_t(x,t) + (f(x,u^*(x,t))m(x,t))_x = 0 \quad \text{in }]X_{\text{ON}}, X_{\text{OFF}}[\times]0,T[, \\ m(X_{\text{ON}},t) = m(X_{\text{OFF}},t) = 0 \quad \forall t \in [0,T], \\ m(x,0) = m_0(x) \quad \forall x \in [X_{\text{ON}}, X_{\text{OFF}}], \\ \displaystyle\int_{X_{\text{ON}}}^{X_{\text{OFF}}} m(x,t)\,dx = 1 \quad \forall t \in [0,T], \\ \\ \overline{u}(t) = \displaystyle\int_{X_{\text{ON}}}^{X_{\text{OFF}}} u^*(t,x)m(t,x)\,dx \quad \forall t \in [0,T], \\ \\ \overline{m}'(t) = -\alpha\overline{m}(t) + \sigma\overline{u}(t) + c, \\ \overline{m}(0) = \overline{m}_0. \end{cases} \tag{5}$$

Definition 1 Consider an initial distribution m_0 and the corresponding mean $\overline{m}_0$. By *solution of (5)* we mean any continuous function $\overline{m} : [0, T] \to \mathbb{R}$, such that $\overline{m}(0) = \overline{m}_0$, and that "solves" the following fixed point procedure. Given $\overline{m}$, solve I and take the solution $v_{\overline{m}(\cdot)}$. Use this to calculate $u^*_{\overline{m}(\cdot)}$ from II. With such $u^*_{\overline{m}(\cdot)}(\cdot)$, solve III to obtain the distribution $m_{\overline{m}(\cdot)}$, and calculate $\overline{u}_{\overline{m}(\cdot)}$ from IV. Use this to calculate a new function $M_{\overline{m}(\cdot)}(\cdot)$ from V. Require that $\overline{M}_{\overline{m}(\cdot)}(\cdot) = \overline{m}(\cdot)$.

It follows from the above definition that the problem has no solution if there exists no $\overline{m}$ for which the above procedure produces a fixed point.

We apply such a procedure in Sect. 4 to prove that, starting from a symmetric distribution, under a suitable choice of the terminal cost Ψ, there exists a solution with constant and null mean $\overline{m}$.

In the next section we elaborate more on the concept of weak solution introduced above.

Remark 1 As the interval $[X_{\text{ON}}, X_{\text{OFF}}]$ is invariant, there are no boundary conditions on v in (5) I. Also note that V is redundant. Indeed, $\overline{m}$, which satisfies V, is the mean of m which can be calculated from III. However, we prefer to include the ODE V because it plays a role in the sequel, and because our model is centered around the evolution of the mean $\overline{m}$. Furthermore, the distribution m is used in the calculation of the mean control $\overline{u}$. In the special case when the optimal feedback is linear in space, $u^*(t,x) = \gamma(t)x$, the mean control $\overline{u}$ can be calculated directly from the mean distribution $\overline{m}$:

$$\overline{u}(t) = \int_{X_{\text{ON}}}^{X_{\text{OFF}}} \gamma(t)xm(t,x)\,dx = \gamma(t)\overline{m}(t).$$

So, the equation $\overline{m}' = -\alpha\overline{m} + \sigma\gamma(t)\overline{m} + c$ can replace the Kolmogorov equation III.

In addition to this, note that given a solution of (5) as defined in Definition 1, the distribution m is, in general, determined from III provided that equation II yields a unique optimal control.

3 Weak Solutions in the System (5)

In this section, we address two main issues. First, a solution of the Bellman equation I does not necessarily have a spatial derivative and therefore we need to introduce the notion of viscosity solution (see for instance [8]). Second, even if the optimal feedback exists, it is often discontinuous and so the field $f(\cdot, u^*(\cdot,\cdot))$ is discontinuous too. We must then consider solutions m in the distributional sense (see for instance [13]). Moreover, in our particular case, as we will see in Sect. 4, we need to consider distributions m which are not absolutely continuous with respect to the Lebesgue measure, that is, distributions that cannot be defined through a function. Indeed, even starting from smooth data, the probability measure may be characterized by the formation of atomic parts (i.e. Dirac masses). Atomic parts of the distribution together with the discontinuity of the optimal field $f(\cdot, u^*(\cdot,\cdot))$ give rise, in general, to challenging aspects related to the definition itself of solutions.

To address the above issues, we next introduce a rigorous definition of weak solution.

Let $\mathcal{P}$ denote the set of positive probability measures on $[X_{\text{ON}}, X_{\text{OFF}}]$ endowed with the weak-star topology, that is, $\mu_n \to \mu$ in $\mathcal{P}$ if and only if, for every continuous function φ on $[X_{\text{ON}}, X_{\text{OFF}}]$,

$$\int_{X_{\text{ON}}}^{X_{\text{OFF}}} \varphi(x)\,d\mu_n \to \int_{X_{\text{ON}}}^{X_{\text{OFF}}} \varphi(x)\,d\mu,$$

where $\int(\cdot)d\mu$ stands for the integral with respect to the measure μ.

Let the initial datum m_0 belong to $\mathcal{P}$ and be of compact support in $]X_{\text{ON}}, X_{\text{OFF}}[$. Suppose that the optimal feedback u^* as in II is in L^1 and that it is defined everywhere in such a way that the corresponding trajectory $X' = f(X, u^*)$ is (optimal and) unique. Hence, a weak solution of the Kolmogorov equation III is a continuous function $m : [0,T] \to \mathcal{P}, t \mapsto m[t]$, such that, for every test function $\varphi \in C_c^1([X_{\text{ON}}, X_{\text{OFF}}] \times [0,T[)$, ($C_c^1$ is the space of continuously differentiable function with compact support)

$$\int_{X_{\text{ON}}}^{X_{\text{OFF}}} \varphi(x,0)\,dm_0 + \int_0^T \int_{X_{\text{ON}}}^{X_{\text{OFF}}} \big[\varphi_t(x,t) + f\big(x,u^*(x,t)\big)\varphi_x(x,t)\big]\,dm[t]\,dt = 0. \qquad (6)$$

Note that the null boundary condition is here taken into account because φ is not required to vanish at X_{ON} and at X_{OFF} and (6) does not contain (spatial-) boundary pieces.

We expect a solution m of the form

$$m[t] = \tilde{m}(\cdot, t) + \sum_{i=1}^{\ell} \gamma_i(t)\delta_{y_i(t)}, \tag{7}$$

where $\tilde{m} : [X_{\mathrm{ON}}, X_{\mathrm{OFF}}] \times [0, T]$ is an L^1 function (the continuous part of the solution), $\delta_{y_i(t)}$ is the Dirac mass concentrated on the point $y_i(t) \in [X_{\mathrm{ON}}, X_{\mathrm{OFF}}]$ with $y_i(\cdot)$ continuous, and $\gamma_i(\cdot)$ is a positive continuous function.

Note that, in general, a measure $m \in \mathcal{P}$ is not of the form as in (7), because singular (Cantor) parts may occur and the jump part (atomic part) may be more complicated than a finite sum of Dirac masses. However, as the model is one-dimensional, in the next section we consider initial data and solutions of the form (7).

Now, possible discontinuities of the integral of $f(\cdot, u^*(\cdot, \cdot))$ with respect to a Dirac mass may be an issue. Indeed the integral turns out to be the value of the integrand in the point of concentration of the Dirac mass.

A way to overcome the issue is the requirement that u^*, although discontinuous, be defined everywhere in such a way that the optimal trajectory exists for every times. By doing this, we obtain a unique value for u^* in the point of discontinuity, which guarantees that the trajectory is unique.

Once we have a solution m of III (recall that it must be a continuous function from $[0, T]$ to $\mathcal{P}$), the equation V is certainly satisfied by the continuous mean $\overline{m}(t)$ in a distributional sense, that is, for every $\varphi \in C_c^1([0, T[)$,

$$\int_0^T \overline{m}(t)\varphi_t(t)\,dt = -\overline{m}_0\varphi(0) - \int_0^T \big(-\alpha\overline{m}(t) + \sigma\overline{u}(t) + c\big)\varphi(t)\,dt.$$

To see this, for any test function $\varphi \in C_c^1([0, T[)$, let us set $\tilde{\varphi}(x, t) = x\varphi(t)$ and apply (6) with $\tilde{\varphi}$ as test function. Now, being $\overline{m}$ continuous and bounded, we see that $\overline{m}$ satisfies V in the usual integral sense:

$$\overline{m}(t) = \overline{m}_0 + \int_0^t \big(-\alpha\overline{m}(s) + \sigma\overline{u}(s) + c\big)\,ds \quad \forall t \in [0, T].$$

Hence, we can formally state the following definition of solution of (5) in a weak sense, which will be considered in the rest of the paper.

Definition 2 Consider an initial distribution $m_0 \in \mathcal{P}$ and the corresponding mean $\overline{m}_0$. A probability measure $m \in \mathcal{P}$ and corresponding continuous map $\overline{m} : [0, T] \to [0, +\infty[$, satisfying $\overline{m}(0) = \overline{m}_0$ are a solution of (5) if the corresponding unique locally Lipschitz continuous viscosity solution of I yields a solution with the following properties: (i) once substituted in II the solution yields a unique feedback control u^* which is defined almost everywhere, and (ii) whenever II is not defined, there exists a unique feedback control u^* that guarantees that the optimal trajectory $X' = f(X, u^*)$ exists for all times. With such u^* and m, there exists a distributional solution of III which, once entered into IV, defines $\overline{u}$ such that $\overline{m}$ is solution of V.

4 Looking for a Mean-Field Game Equilibrium

A stationary mean distribution $\overline{m}_0$ can be regarded as a mean-field game equilibrium if it is a fixed point of the fixed point procedure given by (5), in the sense of Definition 2. In this section we show that, starting from a symmetric initial distribution, we can design the terminal cost Ψ such that $\overline{m}_0 = 0$ is a mean-field equilibrium. We refer the reader to [8] for further details on the theory of viscosity solutions.

Theorem 1 *Let the initial distribution $m_0 \in \mathcal{P}$ have mean value equal to zero, and be symmetric with respect to the mean, and be of compact support in the open interval $]X_{\mathrm{ON}}, X_{\mathrm{OFF}}[$. In addition, let m_0 be absolutely continuous and still denote its density by m_0, which we also suppose to be continuous. Finally, assume that m_0 is separately C^1 in the open intervals $]X_{\mathrm{ON}}, 0[$ and $]0, X_{\mathrm{OFF}}[$. Then, there exists a terminal cost Ψ such that $\overline{m}_0 = 0$ is a mean-field equilibrium.*

Proof For $\overline{m}_0$ to be a mean-field equilibrium, we need that, given the corresponding mean-field optimal control $\overline{u}(\cdot)$, this is also stationary and equal to

$$\overline{u} \equiv \frac{\alpha \overline{m}_0 - c}{\sigma} = \frac{X_{\mathrm{OFF}} - \overline{m}_0}{X_{\mathrm{OFF}} - X_{\mathrm{ON}}},$$

where the second equality is obtained by replacing parameters σ and c by their expressions in terms of X_{OFF} and X_{ON}, namely, $\sigma := -\alpha(X_{\mathrm{OFF}} - X_{\mathrm{ON}})$, and $c := \alpha X_{\mathrm{OFF}}$. Let us suppose that $\overline{m}_0 = 0$ is a mean-field game equilibrium. The Hamilton–Jacobi part of (5) becomes

$$-v_t + \alpha v_x x - c v_x - q x^2 + [-\sigma v_x - r]_+ = 0 \quad \text{in } \mathbb{R} \times [0, T[, \;\; v(x, T) = \Psi(x). \tag{8}$$

Now, our goal is to find a terminal cost Ψ such that there exists a solution with $\overline{m} \equiv 0$. First, let us start by looking for a solution Ψ of the stationary equation

$$\alpha \Psi_x x - c \Psi_x - q x^2 + [-\sigma \Psi_x - r]_+ = 0 \quad \text{in } \mathbb{R}. \tag{9}$$

If Ψ^0 is solution and $-\sigma \Psi_x^0 - r \le 0$, then we get (recall that $c = \alpha X_{\mathrm{OFF}}$)

$$\Psi_x^0 = \frac{q x^2}{\alpha x - c} \qquad \text{in the half-open interval } [X_{\mathrm{ON}}, X_{\mathrm{OFF}}[$$

whose primitives are

$$\Psi^0(x) = \frac{q}{\alpha^3}\left(\frac{(\alpha x - c)^2}{2} + 2\alpha c x + c^2 \log(c - \alpha x)\right) + k \quad \text{in } [X_{\mathrm{ON}}, X_{\mathrm{OFF}}[. \tag{10}$$

Note that $\Psi_x^0 \le 0$ in the half-open interval $[X_{\mathrm{ON}}, X_{\mathrm{OFF}}[$ and so $-\sigma \Psi_x^0 - r < 0$ (recall $\sigma < 0$, $r > 0$). This means that Ψ^0 is a classical solution of (9) in the open interval $]X_{\mathrm{ON}}, X_{\mathrm{OFF}}[$. Analogously, if Ψ^1 is a solution with $-\sigma \Psi_x^1 - r > 0$, then we get

$$\Psi_x^1 = \frac{q x^2 + r}{\alpha x - \sigma - c} = \frac{q x^2 + r}{\alpha x + c} \qquad \text{in the half-open interval }]X_{\mathrm{ON}}, X_{\mathrm{OFF}}].$$

Note that Ψ_x^1 is positive and that $-\sigma \Psi_x^1 - r > 0$ everywhere and hence any primitive Ψ^1 is a solution of (9) in the half-open interval $]X_{\mathrm{ON}}, X_{\mathrm{OFF}}]$.

Now, with the same controlled dynamics and control set as above, we consider two optimal control problems in $[X_{\mathrm{ON}}, X_{\mathrm{OFF}}[\times [0, T]$ and $]X_{\mathrm{ON}}, X_{\mathrm{OFF}}] \times [0, T]$, respectively. The two problems consist in the minimization of the cost functional $\int_t^T g(x(s), u(s), 0)\, ds$, with terminal costs given by Ψ^0 and Ψ^1, respectively. The stationary functions Ψ^0, Ψ^1 themselves are then solutions of the corresponding final condition in the Hamilton–Jacobi equation (8). And so they are the value functions of the corresponding control problem.

To see this, observe that Ψ^0 and Ψ^1 are unbounded on their domains, and hence the uniqueness of the viscosity solution of the Hamilton–Jacobi equation with those terminal conditions is not straightforward. However, with respect, for instance, to the case of Ψ^0 in the half-open interval $[X_{ON}, X_{OFF}[$, we can see that our controlled dynamics is such that the state never exits through X_{ON} nor can reach X_{OFF} when starting from $x \in [X_{ON}, X_{OFF}[$. Given any $\overline{x} \in [X_{ON}, X_{OFF}[$ we can then consider the finite horizon problem with exit time from $[X_{ON}, \overline{x}]$ and exit cost and terminal cost both equal to Ψ^0. We then find that Ψ^0 is the unique viscosity solution of the corresponding boundary conditions problem for the Hamilton–Jacobi equation, and as such it is also the value function of the exit-time problem. By the arbitrariness of $\overline{x} \in [X_{ON}, X_{OFF}[$, Ψ^0 is then the value function of the final-time problem in the whole $[X_{ON}, X_{OFF}]$.

With regards to the control law, the feedback optimal controls for these problems are, respectively,

$$u_0(x,t) \equiv 0 \quad \text{and} \quad u_1(x,t) \equiv 1.$$

Now we consider the function on the closed interval $[X_{ON}, X_{OFF}]$

$$\Psi(x) = \begin{cases} \Psi^0(x), & \text{if } x < 0, \\ \Psi^1(x), & \text{if } x > 0, \\ 0 = \Psi^0(0) = \Psi^1(0), & \text{if } x = 0, \end{cases} \tag{11}$$

where the last line means that we have glued Ψ^0 and Ψ^1 in $x = 0$. We then consider the final-time problem in $[X_{ON}, X_{OFF}] \times [0, T]$ given by the running cost $g(\cdot, \cdot, 0)$ and by the continuous terminal cost Ψ. For every $0 \neq x \in [X_{ON}, X_{OFF}]$, we denote by $t^*(x)$ the time of arrival at $x = 0$ under the feedback control u_0 if $x < 0$ and u_1 if $x > 0$. Through direct computations, we have

$$t^*(x) = -\frac{1}{\alpha}\log\left(\frac{c}{c-\alpha x}\right), \quad \text{if } x < 0, \qquad t^*(x) = -\frac{1}{\alpha}\log\left(\frac{c}{c+\alpha x}\right), \quad \text{if } x > 0.$$

Then, we affirm that the value function of this problem is given by

$$v(x,t) = \begin{cases} \Psi(x), & \text{if } T - t < t^*(x), \\ \Psi(x) + \frac{r}{2}(T - t^*(x) - t), & \text{otherwise.} \end{cases}$$

To prove this, we first observe that such a function corresponds to the cost of the feedback control

$$u^*(x,t) = 0 \quad \text{for } x < 0, \qquad u^*(x,t) = 1 \quad \text{for } x > 0, \qquad u^*(x,t) = \frac{1}{2} \quad \text{for } x = 0. \tag{12}$$

To see this observe that (i) Ψ^0 and Ψ^1 are the value functions of the problems whose terminal costs are Ψ^0 and Ψ^1 and (ii) the corresponding controls are the ones considered above. Then v is continuous and bounded and satisfies the final condition. Hence, it is a viscosity subsolution of (8). Also, when differentiable, it satisfies the equation. On the other hand, in the points where it is not differentiable (i.e. the points $(0,t)$ and $(x, t^*(x))$, the superdifferential is empty. Indeed, Ψ is convex and non-differentiable at $x = 0$, and, for every $x \neq 0$, the function $t \mapsto v(x,t)$ is convex and non-differentiable at $t = t^*(x)$.

Being a subsolution, v must be less than or equal to the value function. At the same time, being the cost of a particular control, it must be greater than or equal to the value function. We conclude that v is the value function and the feedback control (12) is optimal.

Now, suppose that the initial distribution $m_0 \in \mathcal{P}$ has mean value equal to zero, and it is symmetric with respect to the mean, and that it is of compact support in the open interval $]X_{\text{ON}}, X_{\text{OFF}}[$. In addition, let m_0 be absolutely continuous, and still denote its density by m_0, which we also suppose to be continuous. Finally, assume that m_0 is separately C^1 in the open intervals $]X_{\text{ON}}, 0[$ and $]0, X_{\text{OFF}}[$. Then, considering the evolution of $X(t)$ with control (12), there is a solution m of the Kolmogorov equation III in (5) which remains zero-mean valued and symmetric. Hence, the mean control $\overline{u}$ is constant and equal to $1/2$ and stabilizes the mean to 0. So $\overline{m}_0 = 0$ is a mean-field equilibrium.

First, note that the optimal feedback control u^* (12) satisfies the conditions in Definition 2. For $x < 0$ consider the dynamics $f(x, 0)$ and let $\tilde{m}(x, t)$ be the corresponding solution of III in (5) (this procedure holds also if we include zero in the support of m_0). In other words, $\tilde{m}$ is the solution of the Kolmogorov equation in $[X_{\text{ON}}, 0[\times[0, T[$ with field $f(x, 0)$. Likewise, for $x > 0$, consider dynamics $f(x, 1)$ and let $\tilde{m}(x, t)$ be the corresponding solution of III in (5). Note that, by continuity, $\tilde{m}$ is well defined also for $x = 0$. Now take $\gamma(t) = 1 - \int_{X_{\text{ON}}}^{X_{\text{OFF}}} \tilde{m}(t, x)\,dx$ and define

$$m[t] = \tilde{m}(\cdot, t) + \gamma(t)\delta_0.$$

The first term takes into account the translation of the distribution towards the reference temperature $x = 0$, and the second one captures the accumulation of agents at $x = 0$. The function $m[\cdot]$ is continuous in time, and $\tilde{m}$ is symmetric with respect to its zero mean (note that $f(x, 0) = -f(x, 1)$). Recalling that $f(0, u^*(0, t)) = f(0, 1/2) = 0$, we can infer that such a symmetric and mean zero-valued distribution is a solution of (6). To see this observe that

$$\begin{aligned}
&\int_0^T \int_{X_{\text{ON}}}^{X_{\text{OFF}}} [\varphi_t(x, t) + f(x, u^*(x, t))\varphi_x(x, t)]\,dm[t]\,dt \\
&\quad= \int_0^T \int_{X_{\text{ON}}}^{X_{\text{OFF}}} [\varphi_t(x, t) + f(x, u^*(x, t))\varphi_x(x, t)]\,d\tilde{m}(x, t)\,dx\,dt + \int_0^T w(t)\varphi_t(0, t)\,dt \\
&\quad= \int_0^T \int_{X_{\text{ON}}}^{0} [\varphi_t(x, t) + f(x, 0)\varphi_x(x, t)]\,d\tilde{m}(x, t)\,dx\,dt \\
&\qquad+ \int_0^T \int_0^{X_{\text{OFF}}} [\varphi_t(x, t) + f(x, 1)\varphi_x(x, t)]\,d\tilde{m}(x, t)\,dx\,dt \\
&\qquad+ \int_0^T (1 - \int_{X_{\text{ON}}}^{X_{\text{OFF}}} \tilde{m}(x, t)\,dx)\varphi_t(0, t)\,dt \\
&\quad= -\int_{X_{\text{ON}}}^{X_{\text{OFF}}} \varphi(x, 0)\,dm_0 + (f(0, 0) - f(0, 1)) \int_0^T \tilde{m}(0, t)\varphi(0, t)\,dt \\
&\qquad+ \int_0^T (\frac{d}{dt} \int_{X_{\text{ON}}}^{X_{\text{OFF}}} \tilde{m}(x, t))\varphi(0, t)\,dx\,dt.
\end{aligned}$$

We can conclude by observing that, by definition of $\tilde{m}$ and by the fact that $\tilde{m}$ satisfies the Kolmogorov equation for $x < 0$ and for $x > 0$ independently, the term in the last line is exactly the opposite of the second one in the previous line. □

Remark 2 This calculation also extends to absolutely continuous initial distribution m_0 even if we relax some of the regularity assumptions introduced above. For instance, we could consider an initial distribution with a finite number of Dirac masses located symmetrically with respect to the origin.

5 Introducing a Thermostat

The mean-field equilibrium discussed above shows that all agents tend to the reference temperature $x = 0$ which also coincides with the reference mean. Note that, for $x = 0$, the optimal feedback implies $u = 1/2$, which is also the control that stabilizes the equation for the mean in 0. According to a possible interpretation of the control $u \in [0, 1]$ as a stochastic control, this means that the agents at $x = 0$ are in the state ON with probability $1/2$. The advantage of the above control law is that, at a macroscopic level, the system is stabilized (the mean is constant and null). This is due to the fact that the devices are not all in the ON or *OFF* state at the same time (we say that the devices are desynchronized). However, at a microscopic level, every single agent shows a fast switching *ON–OFF* control. It must be noted that the fast switching behavior is undesirable as well as troublesome. To overcome this issue, we then change the terminal cost Ψ in order to force the agents to avoid fast switchings while maintaining the desynchronization.

To this purpose, note that the fast switching behavior is due to the fact that, in the definition of Ψ in (11), we have only one threshold, $x = 0$, where the agents switch from Ψ^0 to Ψ^1 and back. Hence, the main idea is to split such threshold in two different thresholds, one determining the switches from Ψ^0 to Ψ^1 and the other one determining the switches from Ψ^1 to Ψ^0. This translates to inserting a hysteretic thermostatic rule in the mathematical model.

Let $\varepsilon > 0$ be fixed and denote by $z(\cdot) = h_\varepsilon[x(\cdot)](\cdot)$ the thermostat with thresholds $\varepsilon, -\varepsilon$ and switching between the values $z = 0, 1$ as a result of the evolution of the continuous scalar function $x(\cdot)$. The thermostat modifies the relation between state and control as illustrated in Fig. 1, left.

For the analytical definition and for the properties of such an operator (which acts between the space of continuous functions and the space of functions with bounded variation) see [34]. Now let Ψ^0 and Ψ^1 be as in (11), that is, $\Psi^0(0) = \Psi^1(0) = 0$. The idea is to consider the terminal cost Ψ^0 up to the upper threshold ε, and the terminal cost Ψ^1 up to the lower threshold $-\varepsilon$. This is illustrated in Fig. 1 right, where the continuous line is the graph of Ψ^0 and the dashed line is the graph of Ψ^1. To allow for the above behavior, we define below two "jumps on the threshold" of the terminal cost: the first condition characterizes the

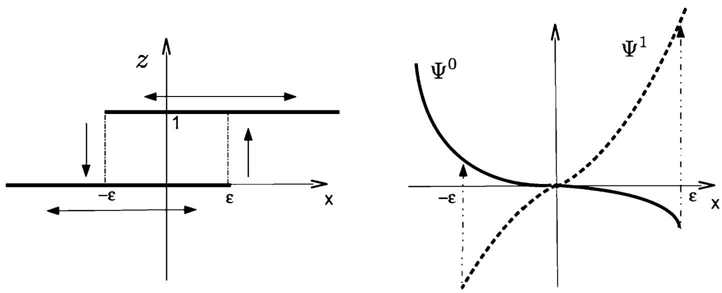

Fig. 1 Thermostat and associated terminal costs

switching from 0 to 1 and the second condition characterizes the switching from 1 to 0:

$$\xi_{01} = \Psi^1(\varepsilon) - \Psi^0(\varepsilon) > 0, \qquad \xi_{10} = \Psi^0(-\varepsilon) - \Psi^1(-\varepsilon) > 0.$$

Moreover, let $\eta_{01}(s) \in \mathbb{N}$ be the total number of switches of the thermostat from 0 to 1 from time zero to time s. Similarly, let $\eta_{10}(s) \in \mathbb{N}$ be the total number of switches of the thermostat from 1 to 0 from time zero to time s. The number of switches of z in a time interval $[t,s]$ is given by its total variation in that interval, $\mathrm{Var}_{[t,s]}(z)$. Indeed the total variation changes of a quantity equal to 1 at every switching instant. Hence we have, for $s \geq t$,

$$\eta_{10}(s) = \left[\frac{\mathrm{Var}_{[t,s]}(z)}{2} + \frac{1}{2}\big(1 - z(t)\big)\right], \qquad \eta_{01}(s) = \left[\frac{\mathrm{Var}_{[t,s]}(z)}{2} + \frac{1}{2}z(t)\right], \tag{13}$$

where $z(t) \in \{0,1\}$ is the output state of the thermostat at the initial time t; and $[\xi]$ indicates the integer part of $\xi \in \mathbb{R}$ (i.e. the largest integer not greater than ξ). The new state variable of the system is the 4-uple $(X, z, \eta_{01}, \eta_{10}) \in [X_{\mathrm{ON}}, X_{\mathrm{OFF}}] \times \{0,1\} \times \mathbb{N} \times \mathbb{N}$, which evolves according to the following dynamics subject to a control $u(\cdot) \in [0,1]$:

$$\begin{cases} X'(s) = -\alpha X(s) + \sigma u(s) + c, & s > t, \\ X(t) = \tilde{X}, & \\ z(s) = h_\varepsilon[X(\cdot)](s), & s > t, \\ z(t) = \tilde{z}, & \\ \eta_{10}(s) = \left[\dfrac{\mathrm{Var}_{[t,s]}(z)}{2} + \dfrac{1}{2}(1 - z(t))\right] + \tilde{\eta}_{10}, & s > t, \\ \eta_{01}(s) = \left[\dfrac{\mathrm{Var}_{[t,s]}(z)}{2} + \dfrac{1}{2}z(t)\right] + \tilde{\eta}_{01}, & s > t. \end{cases} \tag{14}$$

In order to guarantee that the initial point $(\tilde{X}, \tilde{z}, \tilde{\eta}_{01}, \tilde{\eta}_{10}) \in [X_{\mathrm{ON}}, X_{\mathrm{OFF}}] \times \{0,1\} \times \mathbb{N} \times \mathbb{N}$ is consistent and equal to the final point for the preceding evolution, some compatibility conditions have to be satisfied:

$$\begin{aligned} &\tilde{X} < -\varepsilon \implies \tilde{z} = 0, \qquad \tilde{X} > \varepsilon \implies \tilde{z} = 1, \\ &\tilde{z} = 0 \implies \tilde{\eta}_{10} - 1 \leq \tilde{\eta}_{01} \leq \tilde{\eta}_{10}, \\ &\tilde{z} = 1 \implies \tilde{\eta}_{01} - 1 \leq \tilde{\eta}_{10} \leq \tilde{\eta}_{01}. \end{aligned} \tag{15}$$

Now, we consider a final-time optimal control problem with the same running cost $g(X, u, \overline{m})$ as before, but with different terminal cost given by

$$\tilde{\Psi}\big(X(T), z(T), \eta_{01}(T), \eta_{10}(T)\big) = \Psi^{z(T)}\big(X(T)\big) - \xi_{01}\eta_{01}(T) - \xi_{10}\eta_{10}(T),$$

where Ψ^0, Ψ^1 are as before. The presence of the last two addenda in the right-hand side, is due to the fact that, for instance in the point $X = \varepsilon$, when the possible terminal cost drastically changes from Ψ^0 to Ψ^1, then the agent's cost suddenly increases of the value $\xi_{01} > 0$. Hence, those two addenda are useful to obtain a continuous terminal cost.

The final-time optimal control with switching terminal cost $\tilde{\Psi}$, can be studied, via dynamic programming and Hamilton–Jacobi approach, by considering it as the result of several coupled exit-time (non-switching) optimal control problems:

$$\begin{aligned}&[X_{\mathrm{ON}},\varepsilon]\text{ as exit set},\quad z(t)\equiv 0,\quad \eta_{01}\equiv\tilde\eta_{01}\in\mathbb{N},\quad \eta_{10}\equiv\tilde\eta_{10}\in\mathbb{N},\\&\left(\Psi^0-\xi_{01}\tilde\eta_{01}-\xi_{10}\tilde\eta_{10}\right)\text{ as terminal cost},\\&[-\varepsilon,X_{0FF}]\text{ as exit set},\quad z(t)\equiv 1,\quad \eta_{01}\equiv\tilde\eta_{01\in\mathbb{N}},\quad \eta_{10}\equiv\tilde\eta_{10}\in\mathbb{N},\\&\left(\Psi^1-\xi_{01}\tilde\eta_{01}-\xi_{10}\tilde\eta_{10}\right)\text{ as terminal cost}.\end{aligned}$$

In particular, this family of problems are mutually coupled by their exit cost, which is not a priori given, but instead depends on the value function itself evaluated for the other problems. Several examples of optimal control problems of this kind are studied, for instance, in [4], [5] and [6]. In particular, it is proved that the value function is characterized as the unique viscosity solutions of a system of Hamilton–Jacobi equations (one per every branch), suitably coupled via the boundary conditions.

When the evolution of the mean frequency $\overline{m}(\cdot)$ is given, then the final-time optimal control problem with terminal cost $\tilde\Psi$ and running cost $g(X,u,\overline{m}(\cdot))$ is then well-posed from a Hamilton–Jacobi point of view. This consideration extends also to the corresponding mean-field game.

In our particular case, supposing the mean frequency constant and equal to zero, it is not difficult to see (also thank to the results of the previous section) that an optimal feedback control is

$$u(x,z,\eta_{01},\eta_{10},t)=z. \tag{16}$$

Then, all the agents first converge to the interval $[-\varepsilon,\varepsilon]$ and then oscillate from $-\varepsilon$ to ε.

If the initial distribution m has null mean and it is symmetric with respect to its mean, and if the initial distribution of output initial states $\tilde z$, for states $X\in[-\varepsilon,\varepsilon]$, is also symmetric, then the evolution of the mean $\overline{m}(\cdot)$ is again constant (as in the case without thermostatic hysteresis), and so it is a mean-field equilibrium for the thermostatic mean-field game. With respect to the symmetric distribution of $\tilde z$ observe that for every initial time t and for every initial state $X\in[-\varepsilon,\varepsilon]$ we need to consider symmetric initial output state $\tilde z=z(X,t)\in\{0,1\}$; for the other states, $\tilde z$ is uniquely determined (and symmetric), see (15).

Finally, it is obvious that the feedback control (16) is ε-optimal for the mean-field equilibrium of the previous section.

6 Stability

The next theorem establishes that each individual best-response control is a 0-1 bang-bang-like function (Heaviside function).

Theorem 2 *Given $\overline{m}:[0,T]\to\mathbb{R}$ continuous and such that $\overline{m}(0)=\overline{m}_0$. The individual best-response control strategy of each single player is a function $u_{\overline{m}(\cdot)}:\mathbb{R}\times[0,T]\to[0,1]$, $(x,t)\mapsto u_{\overline{m}(\cdot)}(x,t)$ of the form*:

$$u_{\overline{m}(\cdot)}(x,t)=\begin{cases}1 & \text{if } -\sigma v_x(x,t)-r-h[\overline{m}(t)]_+ + k[\overline{m}(t)]_- > 0,\\ 0 & \text{otherwise}.\end{cases} \tag{17}$$

Proof From the first line of (5), and using the explicit expressions for $f(\cdot)$ and $g_{\overline{m}(\cdot)}(\cdot)$ as in (2) and (3) we obtain

$$\arg\sup_{u\in U}\{-f(x,u)v_x(x,t)-g_{\overline{m}(\cdot)}(x,u,t)\}$$
$$=\arg\sup_{u\in[0,1]}\{-\sigma v_x(x,t)u(t)-ru(t)-h[\overline{m}(s)]_+u(t)+k[\overline{m}(t)]_u(t))\}$$

from which we have the thesis. □

An immediate consequence of the above result applies to the case where $\overline{m}$ is stationary and equal to zero and as such we can drop the last two terms (which depends on $\overline{m}$) in the first line of (17). Note that the control strategy is now stationary (we drop explicit dependence on time t).

Corollary 1 *If $\overline{m}_0=0$ is stationary (a mean-field game equilibrium), each player's best-response is a function $u_{\overline{m}_0}:\mathbb{R}\to[0,1]$, $x\mapsto u_{\overline{m}_0}(x)$ of the form:*

$$u_{\overline{m}_0}(x)=\begin{cases}1, & \textit{if } -\sigma v_x(x,t)-r>0,\\ 0, & \textit{otherwise}.\end{cases} \tag{18}$$

Now, likewise in Sect. 4 where the terminal cost is convex, we here assume that the value function v is convex as well. Denote by $perc_m(x)$ the percentile of a given distribution m and let $\overline{perc}_m(x):=1-perc_m(x)$. Essentially, $\overline{perc}_m(x)$ indicates the percentage of players with state x greater than or equal to x according to the distribution m. We can alternatively write $\overline{perc}_m(x)=\int_{\xi\geq x}m[\xi](s)\,dx$.

Also, let us define the threshold (on state x) of the bang-bang control (18).

Definition 3 (Threshold) Let a stationary mean distribution $\overline{m}_0\geq 0$ be given. Also, let $S_{\overline{m}_0}:=inf_x\{x|v_x(x,t)=-\frac{r+h[\overline{m}_0]_+}{\sigma}\}$. The bang-bang control strategy (18) takes on the form:

$$u_{\overline{m}_0}(x)=\begin{cases}1, & \text{if } x>S_{\overline{m}_0},\\ 0, & \text{otherwise}.\end{cases}$$

The definition with $\overline{m}_0\leq 0$ is analogous with the sup rather than inf and $s[\overline{m}_0]_-$ in state of $h[\overline{m}_0]_+$.

According to (18) all appliances with temperature greater than or equal to $S_{\overline{m}_0}$ set to ON, while the rest set to OFF.

Theorem 3 *A mean distribution $\overline{m}_0$ is a mean-field equilibrium if*

$$\overline{perc}_{m_0}(S_{\overline{m}_0})=\frac{X_{\text{OFF}}-\overline{m}_0}{X_{\text{OFF}}-X_{\text{ON}}}, \tag{19}$$

where m_0 is the underlying probability distribution function.

Proof It holds $\overline{u}_{\overline{m}_0}=\int_{\mathbb{R}}u_{\overline{m}_0}(x)\,dx=\overline{perc}_{m_0}(S_{\overline{m}_0})$. Then (19) implies $\overline{u}_{\overline{m}_0}=\frac{X_{\text{OFF}}-\overline{m}_0}{X_{\text{OFF}}-X_{\text{ON}}}$ and this in turns implies $\overline{m}'(t)=0$ for all $t\geq 0$. □

Definition 4 (Bounds) Let a distribution $m[\cdot]$ and its mean $\overline{m}[\cdot]$ be given. We define the upper bound and lower bound as

$$\overline{m}_+:=\min_{0\leq\mu\leq X_{\text{OFF}}}\left\{\mu|\frac{X_{\text{OFF}}-\mu}{X_{\text{OFF}}-X_{\text{ON}}}\leq\overline{perc}_m(S_\mu)\right\};$$

Table 1 Simulation parameters

n	x_{min}	x_{max}	dt	$std(m_0)$	T	$\bar{m}_0$
10^3	-50	50	0.01	{3,5,7}	40	$\{0, -20, 20\}$

$$\overline{m}_- := \max_{X_{ON} \le \mu \le 0} \left\{ \mu | \frac{X_{OFF} - \mu}{X_{OFF} - X_{ON}} \ge \overline{perc}_m(S_\mu) \right\}.$$

Theorem 4 (Controlled invariant set) *Let an initial distribution m_0 and its mean $\overline{m}_0$ be given such that $\overline{m}_-(0) \le \overline{m}_0 \le \overline{m}_+(0)$. Let $m[\cdot](\cdot)$ and $\overline{m}[\cdot](\cdot)$ be the evolution of the distribution and its mean over the horizon $[0, T]$ according to $m_t + div(f(X, u)m) = 0$ and denote by $\overline{m}_\pm(t)$ the corresponding bounds. Then we have for every $0 \le t \le T$*

$$\overline{m}_-(t) \le \overline{m}(t) \le \overline{m}_+(t).$$

Proof It is sufficient to prove that if $\exists t$ such that $\overline{m}(t) = \overline{m}_+(t)$ then $\overline{m}'(t) \le 0$. Actually should such a t exist, from the continuity of $\overline{m}'(s)$ then $\overline{m}(s) \le \overline{m}_+(s)$ for every $0 \le s \le t$. To see that $\overline{m}(t) = \overline{m}_+(t)$ implies $\overline{m}'(t) \le 0$ observe that by the definition of $\overline{m}_+$ we have $\overline{u}_{\overline{m}_+}(t) = \overline{perc}_m(S_{\overline{m}_+}) \ge \frac{X_{OFF} - \overline{m}_+}{X_{OFF} - X_{ON}}$ which proves the thesis. The case $\overline{m}(t) = \overline{m}_-(t)$ implying $\overline{m}'(t) \ge 0$ can be proved similarly. □

7 Numerical Examples

Numerical studies show three main evolution plots for different initial distributions as summarized in Figs. 2, 3, 4. In a first example, the mean distribution $\overline{m}(t)$, initially at zero, remains at zero and the standard deviation $std(m(\cdot, t))$ decreases rapidly to zero. The second example shows the stabilizing effects of the bang-bang control: the mean distribution $\overline{m}(t)$, initially at -20, increases to zero and the standard deviation $std(m(\cdot, t))$ decreases rapidly to zero. In the third example, we visualize again the influence of the bang-bang control: both the mean distribution $\overline{m}(t)$ and the standard deviation $std(m(\cdot, t))$ decrease monotonically.

The numerical studies have been conducted using the algorithm displayed below and considering a number of players $n = 10^3$ and a discretized set of states $\mathcal{X} = \{x_{min}, x_{min} + 1, \dots, x_{max}\}$ where $x_{min} = 50$ (minimum temperature) and $x_{max} = 50$ (maximum temperature). The simulation parameters are listed in Table 1. We assume that the step size for the simulation is $dt = 0.01$. The horizon length (number of iterations) is $T = 40$, large enough to show convergence of the population regimes. With the above parameters' values, the dynamic equation (2) takes on the form

$$\begin{cases} dX(t) = (-\alpha X(t) + \sigma u(t) + c)\,dt, \\ X_0 \in \{x_{min}, x_{min} + 1, \dots, x_{max}\}. \end{cases} \tag{20}$$

As regards the initial distribution, we assume m_0 to be Gaussian with mean $\overline{m}_0$ equal to 0, -20, and 20 for Example 1, 2, and 3, respectively. For each example the standard deviation $std(m_0)$ is equal to 3, 5, and 7. To simplify the dependence of the threshold $S_{\overline{m}(t)}$ on $\overline{m}(t)$ we assume the relation $S_{\overline{m}(t)} = 0.5\overline{m}(t)$. We recall here that the exact relation is

$$S_{\overline{m}(t)} := inf_x \left\{ x | v_x(x, t) = -\frac{r + h[\overline{m}(t)]_+ - k[\overline{m}(t)]_-}{\sigma} \right\}.$$

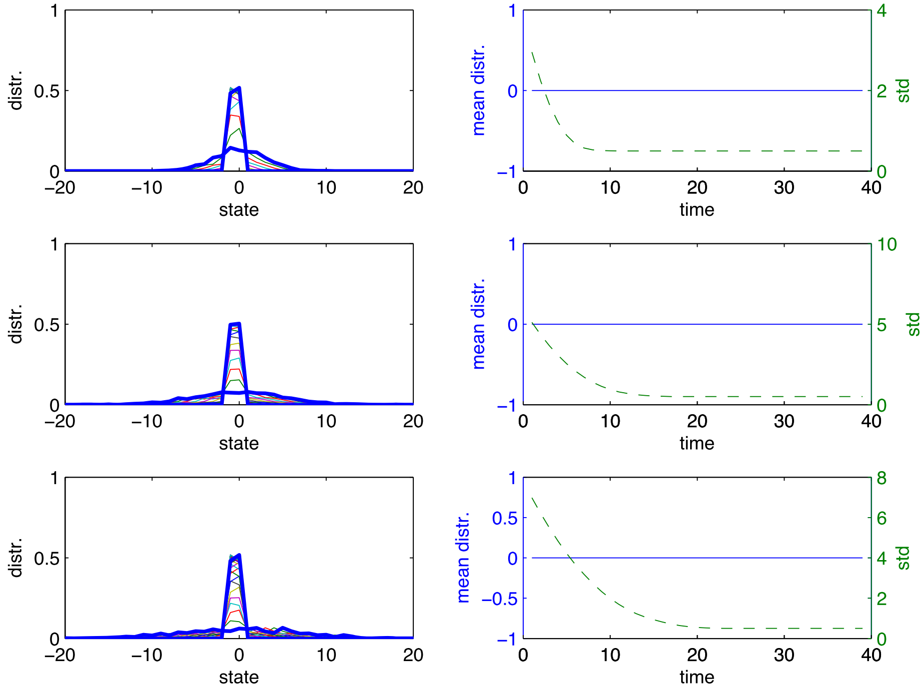

Fig. 2 Example 1: distribution at different times (*left*) and time plot of mean distribution and standard deviation (*right*)

However, we can always set the coefficients involved in the above relation so to approximate $S_{\overline{m}(t)} \approx 0.5\overline{m}(t)$. In Example 2 and 3, the coefficient 0.5 (we could take for it any value less than (1) plays a crucial role. Simulations carried out for coefficients values greater than 1, which we omit for sake of conciseness, have shown that zero is no longer a stable equilibrium for $\overline{m}(t)$.

Algorithm

Input: Set of parameters as in Table 1
Output: Distribution function $m(\cdot, t)$, mean $\overline{m}(t)$ and standard deviation $std(m(\cdot, t))$.

Initialize. Generate x_0 as n random samples from Gaussian distribution with mean $\overline{m}_0$ and standard deviation $std(m_0)$,
for time $t = 0, 1, \ldots, T-1$ **do**
  if $t > 0$, **then** compute distribution $m(\cdot, t)$, mean distribution $\overline{m}(t)$, standard deviation $std(m(\cdot, t))$,
  end if
  compute threshold $S_{\overline{m}(t)}$,
  for player $i = 1, 2, \ldots, n$ **do**
    compute bang-bang control $u(t)$ based on $S_{\overline{m}(t)}$,
    compute new state $X(t+1)$ by executing (20),
  end for
end for
STOP

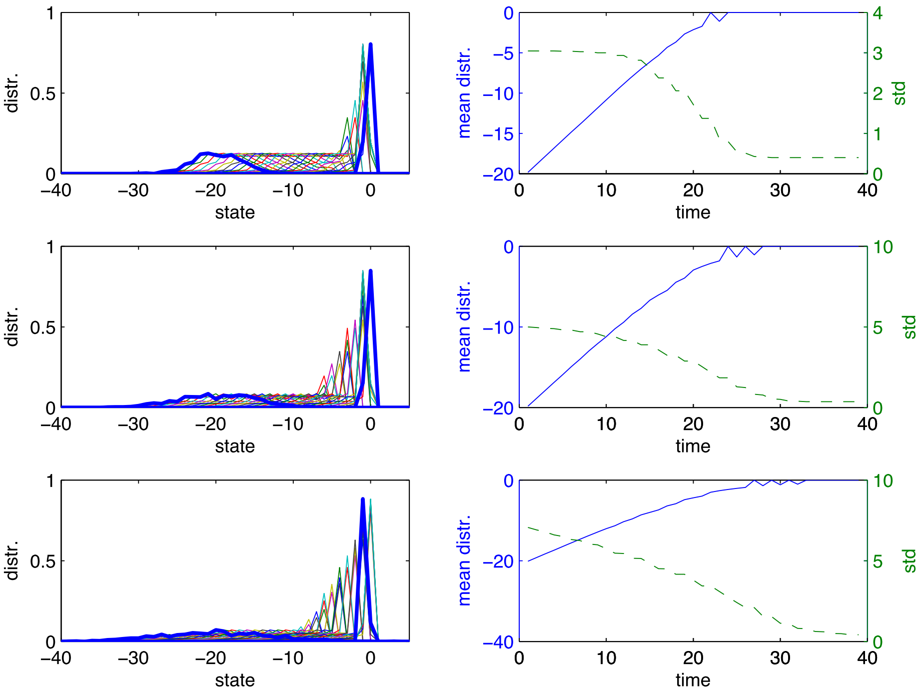

Fig. 3 Example 2: distribution at different times (left) and time plot of mean distribution and standard deviation (right)

Figure 2, left, from top to bottom, shows the distribution evolution $m(\cdot, t)$ vs. the state $x(t)$ at different times. The initial distribution m_0 has mean zero $\overline{m}_0 = 0$ and standard deviation $std(m_0) = 3$ (top), $std(m_0) = 5$ (middle), $std(m_0) = 7$ (bottom). The graphics on the right column display the time plot $\overline{m}(t)$ (solid line and y-axis labeling on the left) and the evolution of the standard deviation $std(m(\cdot, t))$ (dashed line and y-axis labeling on the right). Note that the mean distribution $\overline{m}(t)$ is fixed to zero and at approximately $t = 8$ (top), $t = 10$ (middle), and $t = 20$ (top), the standard deviation $std(m(\cdot, t))$ decreases to zero, which means that all the appliances have reached the reference temperature. In correspondence to this, the mains frequency has reached the nominal value.

Example 2 shows the stabilizing effects of the bang-bang control. Indeed, the standard deviation $std(m(\cdot, t))$ as well as sparsity decrease with time and the mean distribution $\overline{m}(t)$, initially at -20, increases to zero. This is summarized in Fig. 3, left. From top to bottom, the figure shows the distribution evolution $m(\cdot, t)$ vs. the state $x(t)$ at different times. Again, initial standard deviation increases from top to bottom, and in particular is $std(m_0) = 3$ (top), $std(m_0) = 5$ (middle), $std(m_0) = 7$ (bottom). The graphics on the right column display the time plot $\overline{m}(t)$ (solid line and y-axis labeling on the left) and the evolution of the standard deviation $std(m(\cdot, t))$ (dashed line and y-axis labeling on the right). Note that, at approximately $t = 20$ (top) ($t = 30$ in the plot below), the mean distribution $\overline{m}(t)$ reaches zero as well as the standard deviation $std(m(\cdot, t))$.

Example 3 highlights once more the stabilizing effects of the bang-bang control, see Fig. 4, left. From top to bottom, the figure displays the distribution evolution $m(\cdot, t)$ vs. the state $x(t)$ at different times. As before, the initial standard deviation increases from top to bottom, and takes on the values $std(m_0) = 3$ (top), $std(m_0) = 5$ (middle), $std(m_0) = 7$

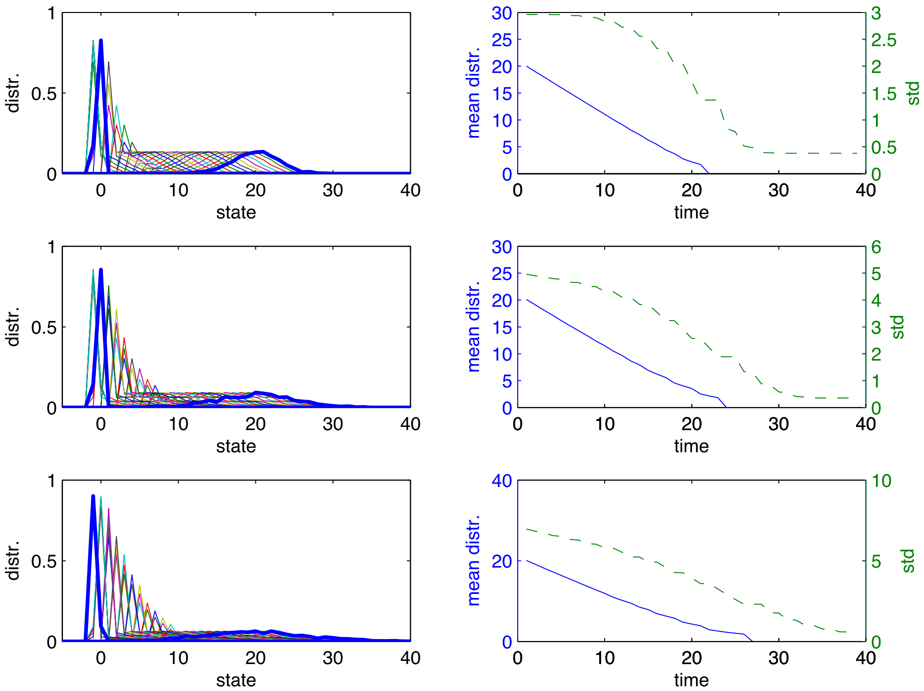

Fig. 4 Example 3: distribution at different times (*left*) and time plot of mean distribution and standard deviation (*right*)

(bottom). The graphics on the right column displays the time plot $\overline{m}(t)$ (solid line and y-axis labeling on the left) and the evolution of the standard deviation $std(m(\cdot, t))$ (dashed line and y-axis labeling on the right). Note that both the mean distribution $\overline{m}(t)$ and the standard deviation $std(m(\cdot, t))$ decrease monotonically to zero.

8 Conclusions

This paper shows that the theory of mean-field games captures interesting phenomena in dynamic demand management in power grids. Mean-field games have been used to improve the network resilience in the case where part of the regulation is shifted to the consumer side. As such we have considered a large population of electrical appliances (the agents) and have shown that an opportune design of the terminal penalty leads the agents to desynchronize their functioning thus reducing systems frequency oscillations.

Acknowledgements This research was supported by the 2012 "Research Fellow" Program of the Dipartimento di Matematica, Università di Trento and by the PRIN 20103S5RN3 "Robust decision making in markets and organization".

References

1. Achdou Y, Camilli F, Capuzzo Dolcetta I (2012) Mean field games: numerical methods for the planning problem. SIAM J Control Optim 50:77–109

2. Achdou Y, Capuzzo Dolcetta I (2010) Mean field games: numerical methods. SIAM J Numer Anal 48:1136–1162
3. Angeli D, Kountouriotis P-A (2012) A stochastic approach to "dynamic-demand" refrigerator control. IEEE Trans Control Syst Technol 20(3):581–592
4. Bagagiolo F (2001) An infinite horizon optimal control problem for some switching system. Discrete Contin Dyn Syst, Ser B 1(4):443–462
5. Bagagiolo F (2005) Optimal control of finite horizon type for a multidimensional delayed switching system. Discrete Contin Dyn Syst, Ser B 5(2):239–264
6. Bagagiolo F, Bauso D (2011) Objective function design for robust optimality of linear control under state-constraints and uncertainty. ESAIM Control Optim Calc Var 17:155–177
7. Bardi M (2012) Explicit solutions of some linear-quadratic mean field games. Netw Heterog Media 7:243–261
8. Bardi M, Capuzzo Dolcetta I (1997) Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations. Birkhäuser, Boston
9. Başar T, Olsder GJ (1999) Dynamic noncooperative game theory. SIAM series in classics in applied mathematics. SIAM, Philadelphia
10. Bauso D, Tembine H, Başar T (2012) Robust mean field games with application to production of an exhaustible resource. In: Proceedings of 7th IFAC symposium on robust control design, Aalborg, Denmark
11. Bauso D, Zhu Q, Başar T (2012) Mixed integer optimal compensation: decompositions and mean-field approximations. In: Proceedings of 2012 American control conference, Montreal, Montreal, Canada
12. Callaway DS, Hiskens IA (2011) Achieving controllability of electric loads. Proc IEEE 99(1):184–199
13. Cardaliaguet P (2012) Notes on mean field games. P.-L. Lions' lectures, Collège de France. https://www.ceremade.dauphine.fr/~cardalia/MFG100629.pdf
14. Couillet R, Perlaza SM, Tembine H, Debbah M (2012) Electrical vehicles in the smart grid: a mean field game analysis. IEEE J Sel Areas Commun 30(6):1086–1096
15. Eto JH, Nelson-Hoffman J, Torres C, Hirth S, Yinger B, Kueck J, Kirby B, Bernier C, Wright R, Barat A, Watson DS (2007) Demand response spinning reserve demonstration. Ernest Orlando Lawrence Berkeley Nat Lab, Berkeley, CA, LBNL Energy, Analysis Dept, 62761
16. Gellings C, Chamberlin J (1988) Demand-side management: concepts and methods. Fairmont Press, Lilburn
17. Gueant O, Lasry JM, Lions PL (2010) Mean field games and applications. Paris-Princeton lectures. Springer, Berlin, pp 1–66
18. Huang MY, Caines PE, Malhamé RP (2003) Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. In: IEEE conference on decision and control, HI, USA, December, pp 98–103
19. Huang MY, Caines PE, Malhamé RP (2006) Large population stochastic dynamic games: closed loop Kean–Vlasov systems and the Nash certainty equivalence principle. Commun Inf Syst 6(3):221–252
20. Huang MY, Caines PE, Malhamé RP (2007) Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria. IEEE Trans Autom Control 52(9):1560–1571
21. Jovanovic B, Rosenthal RW (1988) Anonymous sequential games. J Math Econ 17:77–87
22. Lachapelle A, Salomon J, Turinici G (2010) Computation of mean field equilibria in economics. Math Models Methods Appl Sci 20:1–22
23. Lasry J-M, Lions P-L (2006) Jeux à champ moyen. I. Le cas stationnaire. C R Math 343(9):619–625
24. Lasry J-M, Lions P-L (2006) Jeux à champ moyen. II. Horizon fini et controle optimal. C R Math 343(10):679–684
25. Lasry J-M, Lions P-L (2007) Mean field games. Jpn J Math 2:229–260
26. Ma Z, Callaway DS, Hiskens IA (2013) Decentralized charging control of large populations of plug-in electric vehicles. IEEE Trans Control Syst Technol 21(1):67–78
27. Louzada Pinto JC (2011) Nash equilibrium stability in stochastic control and mean field games. M.S. thesis, advisor P. Cardaliaguet, CEREMADE. http://www-public.it-sudparis.eu/~louzada/Publications/memoire_proba.pdf
28. Mathieu JL, Koch S, Callaway DS (2013) State estimation and control of electric loads to manage real-time energy imbalance. IEEE Trans Power Syst 28(1):430–440
29. Pesenti R, Bauso D (2011) Mean field linear quadratic games with set up costs. In: Proceedings of the int conf on NETwork games, COntrol and OPtimization (NetGCooP 2011), Paris, 12–14 Oct 2011
30. Selten R (1970) Preispolitik der Mehrprodktenunternehmung in der statischen Theorie. Springer, Berlin
31. Tembine H (2011) Mean field stochastic games. Notes
32. Tembine H, Zhu Q, Başar T (2011) Risk-sensitive mean-field stochastic differential games. In: Proceedings of 2011 IFAC world congress, Milan, Italy, August 29–September 2

33. Tembine H, Le Boudec JY, ElAzouzi R, Altman E (2009) Mean field asymptotic of Markov decision evolutionary games. In: International IEEE conference on game theory for networks (Gamenets 2009)
34. Visintin A (1994) Differential models of hysteresis. Springer, Heidelberg
35. Weintraub GY, Benkard C, Van Roy B (2005) Oblivious equilibrium: a mean field approximation for large-scale dynamic games. Advances in neural information processing systems. MIT Press, Cambridge
36. Zhu Q, Tembine H, Başar T (2011) Hybrid risk-sensitive mean-field stochastic differential games with application to molecular biology. In: Proceedings of conference on decision and control, Orlando, FL, Dec 14–Dec 16, 2011
37. Zhu Q, Başar T (2011) A multi-resolution large population game framework for smart grid demand response management. In: Proceedings of intl conference on network games, control and optimization (NETGCOOP 2011), Paris, France, Oct 12–14, 2011