A Multi-Resolution Large Population Game Framework for Smart Grid Demand Response Management

Quanyan Zhu, Tamer Basar

A Multi-Resolution Large Population Game Framework for Smart Grid Demand Response Management

Quanyan Zhu and Tamer Başar

***Abstract*—Dynamic demand response (DR) management is becoming an integral part of power system and market operational practice. Motivated by the smart grid DR management problem, we propose a multi-resolution stochastic differential game-theoretic framework to model the players' intra-group and inter-group interactions in a large population regime. We study the game in both risk-neutral and risk-sensitive settings, and provide closed-form solutions for symmetric mean-field responses in the case of homogenous group population, and characterize the symmetric mean-field Nash equilibrium using the Hamilton-Jacobi-Bellman (HJB) equation together with the Fokker-Planck-Kolmogorov (FPK) equation. Finally, we apply the framework to the smart grid DR management problem and illustrate with a numerical example.**

I. INTRODUCTION

Smart grid is a visionary user-centric system that will elevate the conventional power grid system to one which functions more cooperatively, responsively, and economically [1], [2]. In addition to the incumbent function of delivering electricity from suppliers to consumers, smart grid will also provide information and intelligence to the power grid to enable grid automation, active operation, and efficient demand response. A reliable and efficient communication and networking infrastructure will connect the functional elements within the smart grid [3], [4].

In electricity grids, demand response (DR) is a mechanism for achieving energy efficiency through managing customer consumption of electricity in response to supply conditions, e.g., having end users reduce their demand at critical times or in response to market prices. Modern smart grid technologies such as smart metering devices are enabling demand response in everyday operations of power systems and markets [3], [8]. In the future smart grid, the two-way communications between energy provider and end users enabled by advanced communication infrastructure (e.g., wireless sensor networks and power line communications) and protocols will greatly enhance demand response capabilities of the whole system. In contrast to the current simple time-of-use (TOU) pricing (e.g. peak time vs. off-peak time) [10], it can be envisaged that a more dynamic, real-time adaptation to market prices would not only enable consumers to save more energy and money, as well as manage their usage preferences more flexibly, but also facilitate the grid move closer towards its optimal operating point [5], [6].

Q. Zhu and T. Başar are with Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1308 West Main St., Urbana, IL, 61801, USA. E-mail:{zhu31,basar1}@illinois.edu; The research was partially supported by an AFOSR MURI Grant (FA9550-10-1-0573).

Demand response and distributed energy storage can be seen as distributed energy resources and are the main drivers of smart grid. Demand-side participation in the provision of the power system security not only benefits the consumers but also provides the system operator with more options to maintain a balance between supply and demand. Demand response during periods of peak hours and high prices tend to reduce the load and hence reduce the likelihood of involuntary emergency load disconnections [9]. Unlike the supply-side management of power production, the demand side often involves a large population of customers and DR management needs to take into account aggregate DRs distributed across the distribution network [7].

Motivated by the smart grid DR problem, in this paper, we propose a two-level multi-resolution stochastic differential game-theoretic framework to capture DRs of the consumers at different levels in a large population regime. Power consumers or players in the game are separated into different groups which correspond to different geographical areas or zones in ISO/RTO territory. The interactions between different groups or zones are coupled through some macroscopic quantities such as prices in power markets, and the state of each group or zone is represented by an aggregator that lumps the microscopic behaviors of the players within the group.

In addition to the multi-resolution feature of the framework, we study the game in a large population regime at both macroscopic inter-group level and microscopic intra-group level. In the recent literature, large population games have been studied within different frameworks. In [15], [16], the large population behavior is studied within the deterministic linear-quadratic differential game framework by taking the limit as the number of players goes infinity, in the closed form solution of the game. In [14], [11], [12], a mean-field approach is used to study large population stochastic differential games by coupling a forward Fokker-Planck-Kolmogorov (FPK) equation together with a backward Hamilton-Jacobi-Bellman (HJB) equation.

In this paper, we use both methodologies to study the two-level multi-resolution game in which the inter-group interactions are analyzed using a mean-field approach, while the intra-group interactions are analyzed using a limiting approach. We investigate the game in both risk-neutral and risk-sensitive settings. Risk-sensitive large population games have been studied in [11], [12]. In [11], the authors have studied a class of risk-sensitive mean-field stochastic differential games with an exponential in the long-term cost function. In [12], the class of risk-sensitive games in [11] is extended to hybrid dynamics with continuous-time stochastic dynamics driven by disturbances (of Brownian motion type) and event-driven random switching. In this paper, we study the multi-resolution features of the same class of risk-sensitive mean-field games with exponentiated cost functionals as in [11].

The paper is organized as follows. In Section II, we describe the problem formulation and the risk-neutral stochastic differential game model. In Section III, we characterize players' symmetric mean-field responses in the case of homogenous group population and study two special cases of scaling factors. In Section IV, we formulate the risk-sensitive version of the stochastic differential game and study the risk-sensitive mean-field response. In Section V, we characterize the mean-field equilibrium of the multi-resolution game framework in risk-neutral and risk-sensitive cases. In Section VI, we apply the game-theoretic model to distributed automatic DR management in the smart grid, and illustrate with a numerical example and simulations. We conclude in Section VII and identify some future work.

II. Problem Formulation

In this section, we propose a game-theoretic framework for networked systems with a large number of interacting agents. We consider a system consisting of a set of M groups $\mathcal{M} = \{m_1, m_2, \cdots, m_M\}$, with each group $k \in \mathcal{M}$ having a set $\mathcal{N}_k$ of N_k members or players, where $\mathcal{N}_k := \{1, 2, \cdots, N_k\}$. Members in group m_k share a group state $x_k \in \mathbb{R}$ which is influenced by the control actions $u_i^k \in \mathcal{U}_i \subseteq \mathbb{R}, m_k \in \mathcal{M}, i = 1, 2, \cdots, N_k$, taken by the group members. Let $\mathbf{u}^k = [u_1^k, u_2^k, \cdots, u_{N_k}^k]$ be the vector of control actions in group m_k. The dynamics of group state x^k evolve according to

$$dx^k(t) = \left(\frac{1}{f(N_k)} \sum_{i=1}^{N_k} u_i^k(t) \right) dt + \sigma^k e^{-\alpha^k t} dB(t), \quad (1)$$

with $x^k(0) = x_0^k$ as the initial condition of group m_k, where $B(t)$ is the Wiener process; $\sigma^k, \alpha^k \in \mathbb{R}_+$ and $f(N_k)$ is a function of the number of players; for example, when $f(N_k) = 1$, the group dynamics are controlled by the aggregate effect of the actions of the players; when $f(N_k) = N_k$, the group state x^k is driven by the average of the control efforts in the group. Each player $i \in \mathcal{N}_k$ in group m_k seeks a control strategy γ_i^k to minimize his integral cost functional

$$J_i^k(\gamma_i^k, \gamma_{-i}^k, \gamma^{-k}) =$$

$$\mathbb{E}\left[\frac{1}{2} \int_0^T e^{-r_i^k t} \left(q_i^k (x^k(t))^2 + (u_i^k)^2(t) \right) dt + \frac{1}{2} q_{i,T}^k (x_T^k)^2 \right], \quad (2)$$

where $r_i^k \in \mathbb{R}_+$ is the discount factor, $q_i^k \in \mathbb{R}_+$ is a weighting parameter on the state cost, $q_{i,T}^k = q_i^k(T)$ and $x_T^k = x^k(T)$. In the cost functional (2), we have used the short-hand notations $\gamma_i^k := \{\gamma_j^k, j \neq i\}$ and $\gamma^{-k} := \{\gamma_j^{k'}, k' \neq k, k' \in \mathcal{M}, j \in \mathcal{N}_{k'}\}$ to denote the set of control strategies of other members in the same group k and the set of control strategies of all players in different groups, respectively.

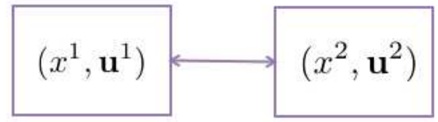

Fig. 1. Illustration of Multi-resolution Games: Upper Layer

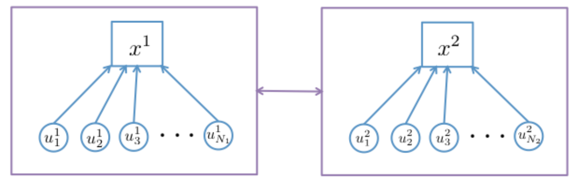

Fig. 2. Illustration of Multi-resolution Games: Bottom Layer

If $r_i^k = r^k$ and $q_i^k = q^k$ for all $i = 1, 2, \cdots, N_k$, then the game in group m_k is completely symmetric in that players share the same parameters. The cost parameter $q_i^k(m^M(t))$ depends on the mean of the group population state $m^M(t)$ defined by

$$m^M(t) = \frac{1}{M} \sum_{k=1}^{M} x^k. \quad (3)$$

In [11] and [12], it has been shown that, under some regularity conditions, the mean state value converges to $m(t)$ as the number of groups becomes infinite, and it can be characterized by a probability density function (to be discussed later).

A state-feedback strategy of player i in group m_k is a mapping $\gamma_i^k : [0, T] \times \mathbb{R}^M \to \mathcal{U}_i^k \subseteq \mathbb{R}$, which depends on agent i's own group state x^k as well as the state of other groups $x^{-k} := \{x^{k'}, k \neq k', m_k, m_{k'} \in \mathcal{M}\}$. Denote by Γ_i^k the set of feasible feedback control laws of agent i in group k. In the mean-field control paradigm where the number of groups M is large, the dependence of the states of other groups x^{-k} will be through the distribution of the group states. A mean-field state-feedback strategy of player i in group m_k is a mapping $\bar{\gamma}_i^k : [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathcal{U}_i^k$, which depends on player i's own group state x^k as well as the mean group state $m(t)$.

The stochastic differential game defined by (1), (2) and (3) constitutes a multi-resolution game framework, or simply a multi-resolution game. A player i's interactions with other players are two-fold. First, player i's state x^k is coupled with members in the game group m_k. Second, the player's cost functional is influenced by the mean group state $m(t)$, and hence is coupled with players in the other groups. Therefore, we can adopt a layered perspective toward solving the game. At the upper inter-group layer, illustrated in Fig. 1, the interactions are among the groups. At the lower intra-group layer, illustrated in Fig. 2, the interactions are within the groups where each player attempts to use minimum control effort to regulate his group state, with an intention to rely on other group members. Such interactions determine some global parameters (for example, price, temperature, health, etc.) through the mean state value, which group members use to achieve their goal at the bottom layer with minimum cost.

In this paper, we study the mean-field Nash equilibrium of the multi-resolution game. The user strategies $(\gamma_i^{k*}, i \in \mathcal{N}_k, m_k \in \mathcal{M})$ constitute a noncooperative feedback Nash equilibrium [13] if they satisfy the following set of inequalities.

$$J_i^k(\gamma_i^{k*}, \gamma_{-i}^{k*}, \gamma^{-k*}) \leqslant J_i(\gamma_i^k, \gamma_{-i}^{k*}, \gamma^{-k*}), \quad \forall \gamma_i^k \in \Gamma_i^k,\ i \in \mathcal{N}_k,\ m_k \in \mathcal{M} \tag{4}$$

The feedback Nash equilibrium is strongly-time consistent if the properties described by inequalities (4) hold for every game in the same form with starting time $t, t \in [0, T)$.

The user strategies $(\gamma_i^{k*}, m^*(t), i \in \mathcal{N}_k, m_k \in \mathcal{M})$ constitute a mean-field feedback Nash equilibrium if they satisfy the following set of inequalities:

$$J_i^k(\gamma_i^{k*}, \gamma_{-i}^{k*}, m^*(t)) \leqslant J_i(\gamma_i^k, \gamma_{-i}^{k*}, m^*(t)), \quad \forall \gamma_i^k \in \Gamma_i^k,\ i \in \mathcal{N}_k,\ m_k \in \mathcal{M}, \tag{5}$$

where $m^*(t)$ is the limit of (3) as $M \to \infty$ at the equilibrium strategy $\gamma_i^{k*}, i \in N_k, m_k \in \mathcal{M}$, i.e.,

$$m^*(t) = \limsup_{M \to \infty} \frac{1}{M} \sum_{k=1}^{M} x^{k*}(t), \tag{6}$$

where x^{k^*} is the state driven by the equilibrium strategy $\gamma_i^{k*}(\cdot), i \in \mathcal{N}_k, m_k \in \mathcal{M}$. The existence of the mean field limit requires the system to satisfy the asymptotic indistinguishability conditions of [17] under some regularity conditions. For an i.i.d. initial condition and under the optimal controls γ_i^{k*} the solution to the state dynamics generates an indistinguishable sequence and the weakly convergence of the population profile to a certain measure μ is equivalent to the μ-chaoticity condition in [17].

III. Symmetric Mean-Field Response

Provided that the mean group population state $m(t)$ is given and the agents in group m_k are homogenous, the N_k-person non-cooperative differential game in group m_k has symmetric mean-field response $\{\gamma_i^k, i = 1, 2, \cdots, N_k\}$, where $\gamma_i^k = \gamma^k \in \overline{\mathcal{U}}_i^k$. By fixing the other players' optimal strategies at γ^{k*}, an agent i in group m_k seeks an optimal control which is a minimizer in HJB equation:

$$r^k V^k(x^k, m(t), t) - \partial_t V^k(x^k, m(t), t) =$$

$$\min_{u_i \in \mathcal{U}_i^k} \Big\{ \frac{1}{2}\left(q_i^k(m(t))(x^k(t))^2 + (u_i^k(t))^2\right) + \partial_{x^k} V^k(x^k, m(t), t) \frac{1}{f(N_k)} \left[u_i^k + (N_k - 1)\gamma^{k*}\right] + \frac{1}{2}(\sigma^k)^2 e^{-2\alpha^k t} \partial_{x^k x^k} V^k(x^k, m(t), t) \Big\}$$

$$V^k(x^k, m(T), T) = q_{i,T}^k. \tag{7}$$

Note that $V_i^k = V^k$ for every agent i in group m_k. The right-hand side of the HJB equation is minimized by

$$u_i^k = \gamma_i^{k*}(x, t) = -\frac{1}{f(N_k)} \partial_{x^k} V^k(x^k, m(t), t).$$

In the following, we assume, without much loss of generality, that $q_{i,T}^k = 0, i \in \mathcal{N}_k, m_k \in \mathcal{M}$, and examine two cases where $f(N_k) = 1$ and $f(N_k) = N_k$ for every $m_k \in \mathcal{M}$.

A. The case where $f(N_k) = 1$

When $f(N_k) = 1, m_k \in \mathcal{M}$, the group dynamics are driven by the aggregate control effort from their group members. The HJB equation (7) can be written as

$$r^k V^k(x^k, m(t), t) - V^k(x^k, m(t), t) =$$

$$\frac{q}{2}(x^k)^2 - \frac{(2N_k - 1)}{2}[\partial_{x^k} V^k(x^k, m(t), t)]^2 + \frac{(\sigma^k)^2}{2} e^{-2\alpha^k t} V_{x^k, x^k}^k(x^k, m(t), t).$$

$$V^k(x^k, m(T), T) = q_{i,T}^k. \tag{8}$$

Let $V^k(x^k, m(t), t) = A^k(m(t), t)(x^k)^2 + B^k(m(t), t)$ and $A^k(T) = B^k(T) = 0$. We arrive at the following equations:

$$r^k A^k(t) - \dot{A}^k(t) = \frac{q^k(m)}{2} - 2(2N_k - 1)(A^k)^2(t), \tag{9}$$

$$r^k B^k(t) - \dot{B}^k(t) = (\sigma^k)^2 e^{-2\alpha^k t} A^k(t), \tag{10}$$

which admit a unique solution for A^k and B^k

$$A^k(t) = \frac{1 - \exp(C^k(t - T))}{(C^k - r^k)\exp(C^k(t - T)) + C^k + r^k}, \tag{11}$$

where $C^k = \sqrt{(r^k)^2 + 8N_k - 4}$, and

$$B^k(t) = \exp(r^k t) \int_t^T \exp(-(r^k + 2\alpha^k)s) A^k(s) ds. \tag{12}$$

The mean-field optimal response γ_i^{k*} is given by

$$\gamma_i^{k*}(m(t), x^k, t) = -V_{x^k}^k(x^k, t) = -2A^k(m(t), t)x^k(t).$$

The stochastic differential equation hence becomes

$$dx^k = -2N_k A^k(m(t), t)x^k dt + \sigma^k e^{-\alpha^k t} dB(t),\ x(0) = x_0.$$

B. The case where $f(N_k) = N_k$.

When $f(N_k) = N_k, m_k \in \mathcal{M}$, the group dynamics are controlled by the average control effort from their N_k members. Using the HJB equation (7) and assuming the quadratic form of the value functions as in the case where $f(N_k) = 1$, we arrive at a set of ODEs that characterize the value functions as follows.

$$r^k A^k - \dot{A}^k = \frac{q^k(m)}{2} - \left[2 + \frac{4(N_k-1)}{N_k^2}\right](A^k)^2, \quad (13)$$

$$r^k B^k - \dot{B}^k = (\sigma^k)^2 e^{-2\alpha^k t} A^k. \quad (14)$$

The solutions to the ODEs are in the same form as in (11) and (12). As $N_k \to \infty$, the ODE (13) becomes

$$r^k A^k(t) - \dot{A}^k(t) = \frac{q^k(m)}{2} - 2(A^k)^2(t)$$

and its solution is given by (11) with $C^k = \sqrt{(r^k)^2 + 4q^k}$. Hence, the group dynamics as $N_k \to \infty$ become

$$dx^k(t) = -2A^k(m(t),t)x^k dt + \sigma^k e^{-\alpha^k} dB(t), m_k \in \mathcal{M},$$

which is independent of the size of the group.

IV. Risk-Sensitive Mean-Field Symmetric Response

In this section, we study the risk-sensitive stochastic differential game framework where each player i in group m_k minimizes the risk-sensitive cost functional or the exponentiated integral cost [11], [12]

$$\tilde{J}_i^k = \delta_i \log \mathbb{E}\left\{\exp\left(\frac{1}{\delta_i}\right)\left[\frac{1}{2}\int_0^T e^{-r_i^k t}\left(q_i^k (x^k)^2(t) + (u_i^k)^2(t)\right)dt + q_{i,T}^k (x_T^k)^2\right]\right\}, \quad (15)$$

where $\delta_i > 0$ is the risk-sensitivity index for player i and $x_T^k = x^k(T)$ and $q_{i,T}^k = q_i^k(m(T),T)$. Similar to the previous section, we can characterize the risk-sensitive symmetric equilibrium for a group of homogenous players. Let $\tilde{\gamma}_i : [0,T]\times\mathbb{R}\times\mathbb{R} \to \mathcal{U}_i$ be the mean-field risk-sensitive strategy of a player i in group m_k, which is assumed to be piecewise continuous in t and Lipschitz continuous in x^k and m. Let $\tilde{\Gamma}_i^k$ be the set of all admissible state-feedback controls. Suppose that the players in the group are homogenous, i.e., $q_i^k = q^k, \alpha_i^k = \alpha^k, r_i^k = r^k, \delta_i = \delta^k$, for $i = 1, 2, \cdots, N$ in group m_k. Due to symmetry, by fixing the control policy of other $N_k - 1$ players at $\tilde{\gamma}^*$, the HJB equation for the risk-sensitive stochastic control problem is

$$r^k V^k(x,m,t) - \partial_t V^k(x,m,t) = \min_{u^k \in \mathcal{U}^k}\left\{\frac{q}{2}(x^k)^2 + \frac{1}{2}(u^k)^2 + \partial_{x^k}\frac{V^k}{f(N_k)}(u^k + (N_k-1)\tilde{\gamma}) + \frac{(\sigma^k)^2 e^{-2\alpha^k t}}{2\delta^k}(\partial_x V^k)^2 + \frac{1}{2}(\sigma^k)^2 e^{-2\alpha^k t}\partial_{x^k,x^k}V^k\right\}, \quad (16)$$

and the terminal condition

$$V^k(x, m(T), T) = q_T^k, \quad (17)$$

where V^k is the value function. The mean-field optimal response is obtained as

$$u_i^k = -\frac{1}{f(N)}\partial_{x^k} V^k. \quad (18)$$

Assume that $q_{i,T}^k = 0, \alpha^k = 0$, and the value function V^k is quadratic, i.e., $V^k = A^k(t)(x^k)^2 + B^k(t)$ with $A^k(T) = B^k(T) = 0$. Then, we arrive at the following set of differential equations that characterize the value function:

$$r^k\tilde{A}^k(t) - \dot{\tilde{A}}^k(t) = \frac{1}{2}q^k - \eta^k(\tilde{A}^k)^2(t),\ A^k(T) = 0, \quad (19)$$
$$r^k\tilde{B}^k(t) - \dot{\tilde{B}}^k(t) = \frac{1}{2}(\sigma^k)^2 e^{-2\alpha t}\tilde{A}(t),\ B^k(T) = 0,$$

where

$$\eta^k = 4\left(\frac{1}{2} + \frac{1}{f^2(N_k)}(N_k - 1) - \frac{(\sigma^k)^2}{2\delta^k}\right).$$

The closed form solution to (19) is

$$\tilde{A}^k(t) = \frac{1 - \exp(\tilde{C}^k(t-T))}{(\tilde{C}^k - r^k)\exp(\tilde{C}^k(t-T)) + \tilde{C}^k + r^k}, \quad (20)$$

where $\tilde{C}^k = \sqrt{(r^k)^2 + 2\eta^k q^k}$. The risk-sensitive mean-field response is

$$u_i^k = \tilde{\gamma}_i^k(m(t), x_i^k) = -2\tilde{A}^k(m(t),t)x_i^k, \quad (21)$$

which results in the group m_k dynamics

$$dx_i^k(t) = -\frac{N_k\tilde{A}^k(m(t),t)}{f(N_k)}x_i^k + \sigma^k e^{-\alpha^k t}dB(t), \quad \forall i \in \mathcal{N}_k. \quad (22)$$

Remark 1. *From this result, we can observe the relation between the risk-sensitive control and its risk-neutral counterpart. The symmetric risk-sensitive equilibrium control is equivalent to the symmetric risk-neutral equilibrium control with N_{eq}^k players. For the case where $f(N_k) = 1$ for every $m_k \in \mathcal{M}$,*

$$N_{eq}^k = N^k - \frac{(\sigma^k)^2}{2\delta^k},$$

and for the case $f(N_k) = N_k$ for every $m_k \in \mathcal{M}$,

$$N_{eq}^k = \frac{N^k}{1 + \frac{(\sigma^k)^2}{2\delta^k}N^k}.$$

The intuition is that in the risk-sensitive case, the players behave more conservatively and design their optimal strategies by believing there are actually fewer number of players in the group, which means that they have to spend their control effort to regulate the state without much reliance on other players.

V. Mean-Field Equilibrium

In this section, we study the large population regime where $M \to \infty$. Assume that all groups are homogenous or identical, i.e., $N_k = N, \sigma^k = \sigma$, and $\alpha^k = \alpha$ for all $m_k \in \mathcal{M}$. We let $P^n(x,t)$ be the current population profile, i.e., $P^n(x,t) = \frac{1}{M}\sum_{k=1}^{M} \mathbf{1}_{\{x^k(t)=x\}}$. At each time t, $P^n(x,t)$ has the support in the finite set $\{0, \frac{1}{M}, \frac{2}{M}, \cdots, 1\}$ for every $x \in \mathcal{X}$. It was shown in [18] that the process $P^n(x,t) \to p(x,t)$ in law if the transition kernel associated with the dynamics (1) satisfies the regularity conditions B0-B3' in [18]. Then, $P^n(x,0) \to p_0(x)$ and at the limit $p(x,t)$ is characterized by the forward Fokker-Planck-Kolmogorov (FPK) equation:

$$\partial_t p(x,t) + \partial_x p(x,t)\left(-\frac{2N}{f(N)}A(m(t),t)x(t)\right) = \frac{\sigma^2}{2}e^{-2\alpha t}\partial^2_{xx}p(x,t) \quad (23)$$

with initial state distribution $p(x,0) = p_0(x)$. Hence $m(t) = \mathbb{E}_{p(x,t)}[x(t)]$.

From the definition of Nash equilibrium in (5) and (6), we can characterize the mean-field equilibrium using (7) and (23). Note that the HJB equation (7) depends on $m(t)$ while the FPK equation (23) yields $m(t)$ that depends on the control variables.

Theorem 1. *Under some regularity conditions as identified above, the symmetric mean-field feedback Nash equilibrium γ^{k*} for homogeneous groups solves the HJB equation (7) coupled with the FPK equation (23).*

The mean-field Nash equilibrium characterized in Theorem 1 corresponds to the case where the group size is finite while the number of groups M is large. Hence the equilibrium depends on the group size. In the following corollary, we see that it is also possible to characterize the equilibrium where both the group size and the number of groups are large.

Corollary 1. *Consider the multi-resolution game described in (1) and (2) with the scaling factor $f(N_k) = N_k$ for all $m_k \in \mathcal{M}$. Under regularity conditions, the symmetric feedback mean-field equilibrium γ^{k*} is characterized by the set of ODEs (9) and (10) together with (23) as $N_k \to \infty$.*

We call the equilibrium obtained in Corollary 1 a *scale-free* Nash equilibrium in that it is independent of the population within and across groups.

In the risk-sensitive case, the symmetric risk-sensitive Nash equilibrium can be characterized in the same manner as in Theorem 1.

Theorem 2. *Under some regularity conditions delineated above, the symmetric risk-sensitive mean-field feedback Nash equilibrium $\tilde{\gamma}^{k*}$ for homogeneous groups solves the HJB equation (16) coupled with the following FPK equation (24).*

$$\partial_t p(x,t) + \partial_x p(x,t)\left(-\frac{2N}{f(N)}\tilde{A}(m(t),t)x(t)\right) = \frac{\sigma^2}{2}e^{-2\alpha t}\partial^2_{xx}p(x,t) \quad (24)$$

with initial state distribution $p(x,0) = p_0(x)$, where $\tilde{A}$ is defined in (20).

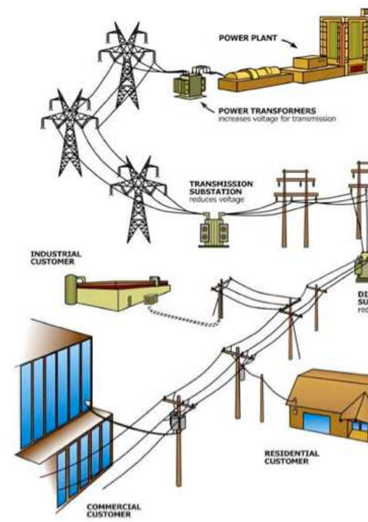

Fig. 3. The process of power generation, transmission and distribution in the electric power grid.

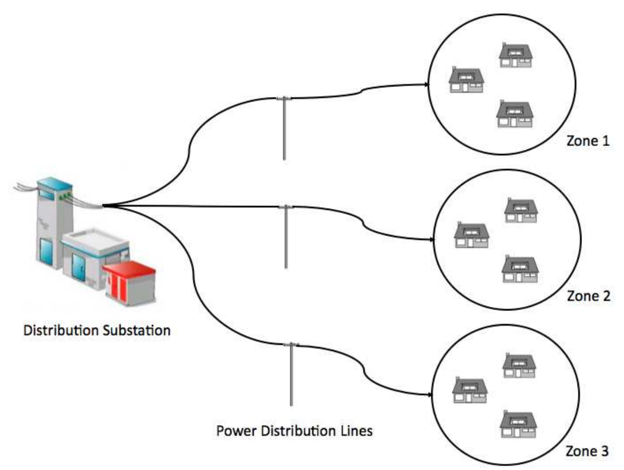

Fig. 4. The geographical description of power distribution from distribution substation to users in different residential zones.

VI. Smart Grid Application: Distributed Automatic Demand Management

In this section, we illustrate the game-theoretic framework with an application in smart grid demand response management. Let $m_k \in \mathcal{M}$ be the index for each power supply zone in the entire power distribution network and $\mathcal{N}_k := \{1, 2, \cdots, N_k\}$ be the set of N_k users in the zone

m_k. The users manage their power consumptions using automatic power management devices, which has a built-in power consumption policy that optimally decides the demand. Fig. 3 illustrates the procedure of power generation, transmission and distribution in electric power grid. The power is finally delivered from the distribution substation to industrial users, commercial users, and residential users. We can see that there is heterogeneity in the types of users. Industrial and commercial users are the major sources of major consumption and they require different types of quality of service from residential customers. In this example, we focus on the power distribution to residential customers, who are geographically located in different zones. Fig. 4 illustrates the distribution process from the substation to the customers. Since residential users often share similar power usage patterns, we can assume symmetry among the users in the zones. In addition, we assume similar geographical characteristics of the zones and hence we can assume the homogeneity between the zones.

Let d_i^k be the power demand of user i in zone m_k. The goal of each user is to optimize his utility and also meet the energy demand with the supply. The total energy demand in zone m_k is described by state $\tilde{x}^k(t)$ whose dynamics evolve according to

$$\dot{\tilde{x}}(t) = \sum_{i=1}^{N_k} d_i^k + \sigma^k dB(t). \tag{25}$$

Each user i in zone m_k optimizes its power usage by effectively allocating power to appliances. Let $\mathcal{L}_i^k = \{l_1^k, l_2^k, \cdots, l_{l_i}^k\}$ be the set of L_i^k appliances for user i in zone m_k such as washer and dryer, refrigerator, air-conditioner, refrigerator, PHEV, etc. For each appliance $l \in \mathcal{L}_i^k$, we define an energy consumption scheduling vector $\mathbf{v}_i^k(t) = [v_{i,1}^k, v_{i,2}^k, \cdots, v_{i,L_i^k}^k] \in \mathbb{R}_+^{L_i^k}$, where $v_{i,j}^k$ denotes the power consumption that is scheduled for appliance $l_j^k \in \mathcal{L}_i^k$ of user i. Each user i seeks to allocate power to its appliances by maximizing their total utilities subject to the power demand constraint d_i^k at time t. This is described by the utility optimization problem (UOP) as follows.

$$\max_{\mathbf{v}_i^k \in \mathbb{R}_+^{L_i^k}} \quad W_i^k(\mathbf{u}_i^k) := \exp\left[\zeta_i^k \sum_{l\in\mathcal{L}_i^k} a_{i,l}^k(t)\ln(v_{i,l}^k(t))\right]$$
$$\text{s. t.} \quad \sum_{l\in\mathcal{L}_i^k} v_{i,l}^k = d_i^k,$$

where $\zeta_i^k \geqslant 0$ is a sensitivity parameter with respect to the demand, and $a_{i,l}^k \in \mathbb{R}_+, l \in \mathcal{L}_i^k$ is a parameter that indicates the importance and time-sensitivity of each appliance. Without loss of generality, we assume that $\sum_{l\in\mathcal{L}_i^k} a_{i,l}^k = 1$ and $a_{i,l}^k \in [0,1], l \in \mathcal{L}_i^k$. The objective function $W_i^k : \mathbb{R}_+^{L_i^k} \to \mathbb{R}_+$ can be equivalently written into the form of products

$$W_i^k(\mathbf{v}_i) = \prod_{l\in\mathcal{L}_i^k} (v_{i,l}^k)^{a_{i,l}^k \zeta_i^k}(t).$$

Hence, the objective function is analogous to the product function in Nash bargaining problems [19] and the outcome of its solution embodies proportional fairness [20]. The utility α_i^k of a household i as a function of the demand d_i^k is given by $\alpha_i(d_i) = \max_{\mathbf{v}_i\in\mathcal{U}_i} W_i(\mathbf{v}_i^k)$, where

$$\mathcal{U}_i^k := \left\{ \mathbf{v}_i^k \in \mathbb{R}_+^{L_i} : v_{i,l}^k \geqslant 0, l \in \mathcal{L}_i^k, \sum_{l\in\mathcal{L}_i^k} v_{i,l}^k = d_i^k \right\}$$

is the feasible set of UOP, which is convex and compact.

Due to the monotonicity of the exponential function, (UOP) is related to the following equivalent utility optimization problem (EUOP):

$$\text{(EUOP)} \;\; \hat{\alpha}_i^k(d_i) := \max_{\mathbf{v}_i^k\in\mathcal{U}_i} \hat{W}_i^k(\mathbf{v}_i) := \sum_{l\in\mathcal{L}_i^k} a_{i,l}^k(t)\ln(v_{i,l}^k(t)),$$

where $\hat{W}_i^k : \mathbb{R}_+^{L_i} \to \mathbb{R}$ and $W_i^k = \exp(\hat{W}_i^k)$, $\alpha_i^k = \exp(\hat{\alpha}_i^k)$. The optimization problem (EUOP) is a convex problem. At every time t, we form a Lagrangian $\mathcal{J}_i^k$ as follows:

$$\begin{aligned}\mathcal{J}_i^k(t) \;\; &= \;\; \sum_{l\in\mathcal{L}_i^k} \ln v_i^k(t) - \lambda_i^k(t)\left(\sum_{l\in\mathcal{L}_i^k} v_{i,l}^k(t) - d_i^k(t)\right), \\ & \quad\;\; i \in \mathcal{N}_k, \end{aligned} \tag{26}$$

where $\lambda_i^k(t) \in \mathbb{R}_+$ is the Lagrange multiplier. Using the constraint (26), we obtain the optimal solution as follows.

$$v_{i,l}^k(t) = a_{i,l}^k(t)d_i^k(t), \quad \lambda_i^k(t) = 1/d_i^k(t). \tag{27}$$

The optimal value of EUOP is achieved at

$$\begin{aligned}\hat{\alpha}_i^k \;\; &= \;\; \sum_{l\in\mathcal{L}_i^k} a_{i,l}^k \ln(a_{i,l}^k d_i^k) \\ &= \;\; \sum_{l\in\mathcal{L}_i^k} a_{i,l}^k \ln a_{i,l}^k + a_{i,l}^k \ln d_i^k.\end{aligned} \tag{28}$$

Hence, the optimal value of (UOP) is obtained as

$$\alpha_i^k = \bar{a}_i^k(d_i^k)^{\zeta_i^k}, \tag{29}$$

where $\bar{a}_i^k = \exp\left(\zeta_i^k \sum_{l\in\mathcal{L}_i^k} a_{i,l}^k \ln a_{i,l}^k\right) \in \mathbb{R}_+$. By Jensen's inequality, we can obtain a lower bound on $\bar{a}_i^k$, i.e., $\bar{a}_i^k \geqslant \sum_{l\in\mathcal{L}_i^k} a_{i,l}^k \exp(a_{i,l}^k\zeta_i^k) \geqslant \min_{l\in\mathcal{L}_i^k} \exp(a_{i,l}^k\zeta_i^k)$. In addition, the sensitivity of the optimal value $\hat{\alpha}_i^k$ with respect to the demand d_i is given by

$$SN_i^k = \left|\frac{d\alpha_i^k}{dd_i^k}\right| = \zeta_i^k \bar{a}_i^k (d_i^k)^{\zeta_i^k - 1}.$$

Hence, the parameter ζ_i^k controls the sensitivity of the optimal value function.

From UOP, the optimal utility of user i is polynomial in the demand d_i^k. Illustrated in Fig. 5, the lower level UOP can be interfaced with the higher level large population game model described in Section II. We obtain the cost functional of user i in zone m_k, which is composed of the cost terms and the utility payoff term as follows.

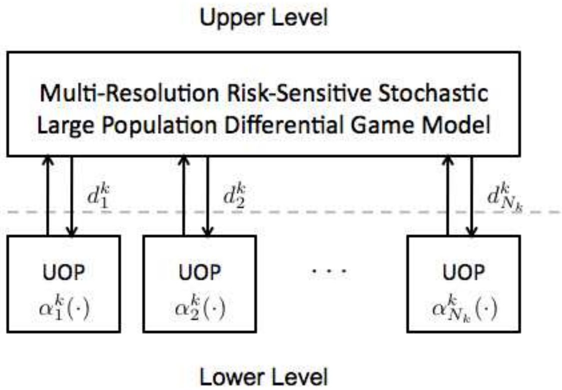

Fig. 5. Aggregated game-theoretic framework with an upper-level multi-resolution risk-sensitive large population stochastic differential game model and a lower-level UOP.

$$\tilde{J}_i^k = \delta_i \log \mathbb{E}\left\{\exp\left(\frac{1}{\delta_i}\right)\left[\frac{1}{2}\int_0^T e^{-r_i^k t}\left((\tilde{x}^k(t)-s^k(t))\right.\right.\right.$$
$$\left.\left.\left. +(p(m(t),t)d_i^k(t))^2 - \alpha_i^k(d_i^k)^2\right) dt\right]\right\}, \tag{30}$$

where $s^k(t)$ is the energy supply to zone m_k at time t; $p(m,t)$ is the price of the power at time t, which is dependent on the mean-field energy consumption in the network m; $\alpha_i^k(\cdot)$ is the optimal value as a function of the demand. For convenience, we can choose $\zeta_i^k = 2$ for every $i \in \mathcal{N}_k, m_k \in \mathcal{M}$. Hence, $\alpha_i^k(d_i^k) = \bar{a}_i^k(d_i^k)^2, \bar{a}_i^k \in \mathbb{R}_{++}$ is the utility achieved by user i in zone m_k by consuming d_i^k amount of power with $p^2 \geqslant \bar{a}_i^k$. Assume that $s^k(t) = S^k$ is constant over the horizon from $t=0$ to $t=T$. Then, we can let $x^k(t) = \tilde{x}^k(t) - S^k$ and the game framework (25) and (30) can be transformed to the one in Section III. The market price p is related to the mean-field energy demand m. We assume that the relation is in the form of

$$p = \sqrt{\beta_1 + \frac{m}{\beta_2(S)}},$$

where $\beta_1 > 0$ and β_2 is an increasing function in the total supply $S = \sum_{m_k \in \mathcal{M}} S^k$. The price increases as the power demand increases for a fixed supply and it decreases as the supply increases for a given demand. The symmetric equilibrium policy can be obtained as in (9) with

$$q^k(m) = \frac{\beta_2}{m + \beta_2(\beta_1 - \bar{a}^k)},$$

where $\bar{a}_i^k = \bar{a}^k, i \in \mathcal{N}_k, m_k \in \mathcal{M}$.

In the following simulation, we set $\bar{a}^k = 1$, $\beta_1 = 1.2, \beta_2 = 5$, $f(N_k) = 1$, $r_i^k = 0.1, \delta_i^k = 10$, for all $i \in \mathcal{N}_k$ and $m_k \in \mathcal{M}$. In Fig. 6, we show the density evolution of the homogenous group state x^k with initial condition $p(x,0) \sim \mathcal{N}(1,1)$. In Fig. 7 and Fig. 8, we show the mean and the variance, respectively, of energy demand in the distribution network. It is easy to observe that the mean of the distribution approaches 0 with initial mean value at 1. The variance of distribution decreases from $t=0$ to $t=1.45$ and then starts to increase onward. In Fig. 9, we show the symmetric risk-sensitive mean-field equilibrium control gain of each user. The magnitude of the control gain starts with a higher value of 0.15 and then decrease smoothly to 0. The sharp decrease close to the terminal time $t=T$ is due to the transversality condition (17). Most of the time from $t=4$ to $t=9.5$, the magnitude of the control effort stays roughly at 0.115.

VII. Conclusion

In this paper, we have introduced a multi-resolution stochastic differential game framework to capture the macroscopic and microscopic interactions among a large population of players. We have studied the stochastic differential game in both risk-neutral and risk-sensitive settings. We have applied the risk-sensitive game framework to the smart grid demand response management problem. We have illustrated the mean-field equilibrium solution with a numerical example. One direction for future work would be to generalize the two-level multi-resolution game to a multiple-hierarchy game. In addition, it would be also interesting to study two Stackelberg game scenarios. One is where the inter-group interactions can lead the intra-group interactions, and the other one is where a third party, for example, power regulatory organizations, can determine the price at a higher hierarchical level by optimizing the equilibrium social welfare.

References

[1] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power and Energy Magazine*, vol. 7, no. 2, pp. 52–62, March 2009.

[2] T. F. Garrity , "Getting smart," *IEEE Power and Energy Magazine*, vol. 6, no. 2, pp. 38–45, March 2009.

[3] C. W. Gellings, *The Smart Grid: Enabling Energy Efficiency and Demand Response*, CRC Press, August 21, 2009.

[4] Z. Han, E. Hossain and V. Poor (Eds.), *Smart Grid Communications and Networking*, Cambridge University Press, 2012.

[5] Department of Energy, "The Smart Grid: An Introduction," Office of Electricity Delivery and Energy Reliability, http://energy.gov/node/235057.

[6] J. B. Cruz, Jr. and X. Tan, *Dynamic Noncooperative Game Models for Deregulated Electricity Markets*, Nova Publishers, New York NY, 2009.

[7] J. Medina, N. Muller and I. Roytelman, "Demand response and distribution grid operations: opportunities and challenges," *IEEE Trans. on Smart Grid*, vol. 1, no. 2, September 2010.

[8] G. Strbac, "Demand side management: benefit and challenges," *Energy Policy*, vol. 36, 2008, pp. 4419–4426.

[9] D. S. Kirschen, "Demand-side view of electricity markets," *IEEE Trans. on Power Systems*, vol. 18, no. 2, May 2003.

[10] M. H. Albadi and E. F. El-Saadany, "Demand response in electricity markets: An overview," IEEE Power Engineering Society General Meeting, 2007, pp. 1–5.

[11] H. Tembine, Q. Zhu and T. Başar, "Risk-sensitive mean-field stochastic differential games," In Proc. of IFAC World Congress, Milan, Italy, Aug. 28 - Sept. 2, 2011.

[12] Q. Zhu, T. Hamidou and T. Başar, "Hybrid risk-sensitive mean-field stochastic differential games with application to molecular biology," in Proc. of 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, Florida, Dec. 12 - 15, 2011.

[13] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM Classics in Applied Mathematics, 2nd ed., 1998.

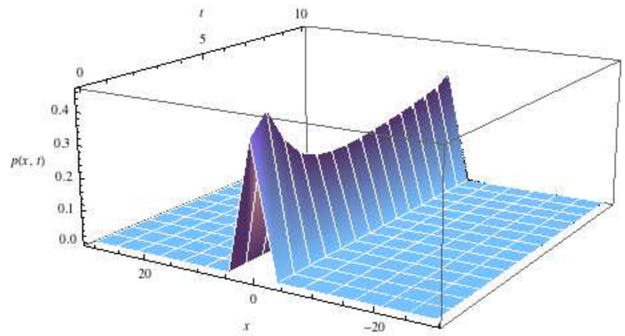

Fig. 6. Evolution of the probability density function $p(x, t)$.

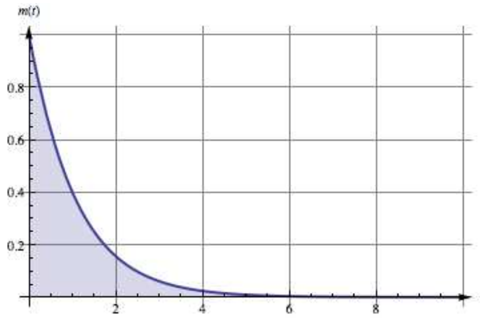

Fig. 7. The mean energy demand $m(t)$.

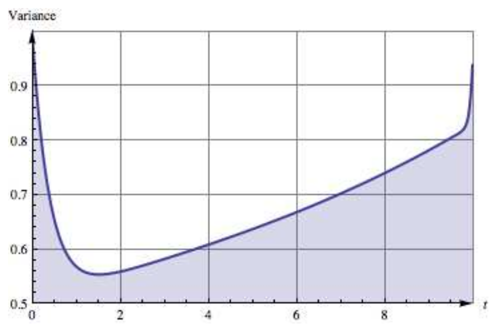

Fig. 8. The variance of energy demand.

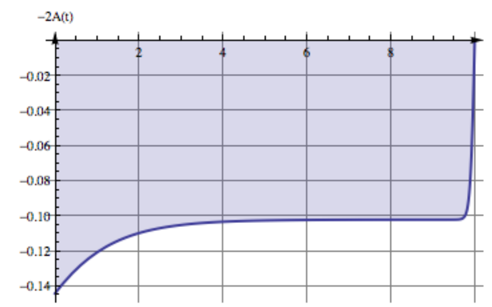

Fig. 9. The mean-field equilibrium control gain $-2A(t)$.

[14] J.-M. Lasry and P.-L. Lions, Mean field games, *Japanese Journal of Mathematics*, vol. 2, 2007, pp. 229–260.

[15] Q. Zhu and T. Başar, "Price of anarchy and price of information in linear quadratic differential games," In Proc. of American Control Conference (ACC), Baltimore, Maryland, June 30 - July 2, 2010, pp. 782–787.

[16] T. Başar and Q. Zhu, "Prices of anarchy, information, and cooperation in differential games," *J. Dynamic Games and Applications*, vol. 1, no. 1, March 2011, pp. 50–73.

[17] Y. Tanabe, "The propagation of chaos for interacting individuals in a large population," *Mathematical Social Sciences*, vol. 51, 2006, pp. 125–152.

[18] H. Tembine, "Mean field stochastic games: convergence, Q/H-learning and optimality," in Proc. of American Control Conference (ACC), San Francisco, CA, USA, pp. 2413–2428.

[19] H. Yaïche, R. R. Mazumdar, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp.667-678, Oct. 2000.

[20] F. P. Kelly, A.K. Maulloo and D.K.H. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237-252, 1998.