Mean Field Games on Prosumers

Baar, Wouter; Bauso, Dario

Published in:
Proceedings of the Conference on Decision and Control 2019

DOI:
10.1109/CDC40024.2019.9030034

2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

Mean Field Games on Prosumers

Wouter Baar[1] and Dario Bauso[2]

***Abstract*— In the realm of dynamic demand, prosumers are agents that can produce and consume goods. In this paper, we study a large population of prosumers and the strategy of each prosumer depends on the average behavior of the population. Every prosumer optimizes his own objective function and we formulate the problem as a first order mean field game. We study the corresponding linear quadratic optimal control problem, which gives us a mean field equilibrium. Finally, a connection to the Bass model is studied. A numerical experiment covering our findings concludes the paper.**

***Index Terms*— Mean field games, dynamic demand, prosumers, best response.**

I. INTRODUCTION

In this article we will study dynamic demand by means of mean field games on a large population of prosumers. A prosumer is an agent that can produce and consume goods. The behavior of the prosumer (to produce or to consume) depends on the behavior of the other players.

In mean field games, the underlying structured network interconnecting the players does not play a role. Since the number of players in a mean field game is usually very large (tending to infinity), it is not of interest to study *which* player does what, but *how many* players do what. Put differently, when an agent decides his next strategy, it does not consider the strategies of its neighboring players, but he bases his strategy on the average behavior of the population.

In the model proposed in this paper, there is a continuous demand and in order to meet the demand, the prosumer can produce goods at certain time periods. However, we assume that one cannot produce indefinitely and at all times (for example, machines need maintenance, or the worker has ended his shift, et cetera.) It is thus of interest to know when the prosumer should switch on to produce, and when to switch off and consume. The cross-coupling mean field term we develop models different kinds of smart pricing policies, aiming at shifting demand to off-peak periods of time.

The model in this paper is as follows. Each prosumer is characterized by three state variables: the inventory level, the probability of being in the mode of producer, and the probability of being in the mode of consumer. The dynamics of these variables are captured in differential equations. Furthermore, every agent is equipped with a finite horizon cost functional, which involves consumption costs and error costs. The error costs are deviations from the mean behavior. Every prosumer wants to minimize its own cost functional, and by using the error as a cross-coupling term, the mean field game framework comes into play.

We will now present a small literature overview. Mean field games were formulated by Lasry and Lions in [1] and [2], and independently by Huang, Caines and Malhamé in [3], [4] and [5]. For more information on mean field games in general, we refer the reader to [6] and [7]. Mean field games in combination with dynamic marketing are also presented in [8], where the situation is applied to electrical power grids. More information on the Bass model can be found in [9].

Regarding notation, we denote by $\mathbb{R}, \mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ the real numbers, the real n-dimensional vector space and the set of all real $n \times m$ matrices, respectively. For a set of measured values x_i, $i = 1, \dots, n$, the mean is denoted by $\bar{x}$. Finally, $0_{n\times n}$ denotes the $n \times n$ zero matrix.

The organization of the paper is as follows. The underlying model will be presented in Section II and first order mean field games are introduced in Section III. Next, we analyze the mean field game for prosumers and we find a mean field equilibrium in Section IV. In Section V a connection to the Bass model will be presented and in Section VI a numerical experiment is run. Section VII concludes our paper with directions for future research.

II. POPULATION OF PROSUMERS

We will now design the mean field game for the case of prosumers. This is done with the aim of incentivizing cooperation among the prosumers by designing cost functionals, one per each prosumer.

Consider a large population of prosumers and a finite time horizon window $[0, T]$. Each prosumer is characterized by the following continuous states: the inventory level $x(t)$, the probability of being in the mode of *producer* $y_1(t) \in [0, 1]$, and the probability of being in the mode of *consumer* $y_2(t) \in [0, 1]$ at time $t \in [0, T]$. Put differently, if $y_1(t) = 1$ the prosumer is producing, while if $y_2(t) = 1$ the prosumer is in consuming mode.

The inventory level of each prosumer evolves according to the following stock-based differential equation

$$\dot{x}(t) = \begin{cases} p - d - \lambda x(t) & \text{if } y_1(t) = 1, \\ -d - \lambda x(t) & \text{if } y_2(t) = 1, \end{cases} \quad t \in [0, T), \quad (1)$$

where the rates p and d are given positive scalars representing the production and demand rates, respectively, and λ is some positive constant.

[1]Wouter Baar is with the ENgineering and TEchnology institute Groningen (ENTEG), Faculty of Science and Engineering, Groningen, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands w.baar@rug.nl

[2]Dario Bauso is with the Jan C. Willems Center for Systems and Control, ENgineering and TEchnology institute Groningen (ENTEG), Faculty of Science and Engineering, Groningen, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands and Dipartimento dell'Innovazione Industriale e Digitale (DIID), Università di Palermo, 90128 Palermo, Italy d.bauso@rug.nl

In accordance with [8], we set the problem in a stochastic framework in which each prosumer is in either the mode of *producer* or it is in the mode of *consumer*, with probabilities $y_1 \in [0,1]$ and $y_2 \in [0,1]$, respectively. The expected inventory level $x(t)$ then evolves according to the following differential equation

$$\dot{x}(t) = (p - d - \lambda x(t))y_1(t) + (-d - \lambda x(t))y_2(t). \quad (2)$$

The modes $y_1(t)$ and $y_2(t)$ are characterized by evolving dynamics in the form of a differential equation as well. To formulate our problem as a control problem we assume that $y_1(t)$ and $y_2(t)$ can be controlled by input variables $u_1(t)$ and $u_2(t)$, as follows:

$$\begin{aligned} \dot{y}_1(t) &= u_1(t) - u_2(t), \\ \dot{y}_2(t) &= u_2(t) - u_1(t). \end{aligned} \quad (3)$$

For a mean-field game formulation, we need to consider a probability density function $m(\cdot)$ that describes the density of the prosumers

$$\begin{aligned} m : \mathbb{R} \times [0,1] \times [0,1] \times [t,T] \to [0,\infty), \\ (x, y_1, y_2, t) \mapsto m(x,y,t). \end{aligned}$$

which satisfies $\int_{\mathbb{R}} \int_{[0,1]\times[0,1]} m(x,y,t)dxdy = 1$ for every t. Note that y is short hand notation for (y_1, y_2), so by $\int_{[0,1]\times[0,1]} m(\cdot)dy$ we simply mean $\int_{[0,1]} \int_{[0,1]} m(\cdot)dy_1dy_2$. Also note that in fact you integrate over a triangle (and not the unit square), since $y_1 + y_2 \leq 1$, because a prosumer cannot be a producer and consumer simultaneously. We assume that m is simply zero outside the set of feasible modes. The function m is a probability density function and describes the distribution of the prosumers, i.e., the way in which the prosumers are distributed over the states (inventory level and mode). Let us now define

$$\begin{aligned} m_p(t) &= \int_{\mathbb{R}} \int_{[0,1]\times[0,1]} y_1 m(x,y,t)dxdy, \\ m_c(t) &= \int_{\mathbb{R}} \int_{[0,1]\times[0,1]} y_2 m(x,y,t)dxdy, \end{aligned} \quad (4)$$

which are the producer and consumer distributions.

To capture the mismatch between producers and consumers, we define the *error* as the discrepancy between the percentage of producers and consumers:

$$e(t) = pm_p(t) - dm_c(t), \quad (5)$$

at every time t. We now consider the running cost $c(\cdot)$ below, which depends on the distribution $m(x,y,t)$ through the error $e(t)$:

$$\begin{aligned} c(\cdot) &= c(x(t), y(t), u(t), m(x,y,t)) \\ &= \tfrac{1}{2}qx(t)^2 + \tfrac{1}{2}r_1u_1^2(t) + \tfrac{1}{2}r_2u_2^2(t) + y_1(t)Se(t) \\ &\quad + y_1(t)W \ , \end{aligned} \quad (6)$$

where q, r_1, r_2, S and W are propitious positive scalars.

In (6), the term $\frac{1}{2}qx(t)^2$ penalizes the deviation of the inventory level of the prosumer from the nominal value, which we set to zero. Setting the nominal inventory to a nonzero value would simply imply a translation of the origin. The terms $\frac{1}{2}r_1u_1^2(t)$ and $\frac{1}{2}r_2u_2^2(t)$ introduce costs on the inputs. A positive error $e(t) > 0$ means that there are more producers than consumers. Since we are minimizing the costs, $y_1 \to 0$ so the prosumer will not produce. If there is more consumed then produced, so the error is negative, $y_1 \to 1$ so the prosumer will produce. In any case, the term $y_1(t)Se(t)$ penalizes the producers when supply exceeds demand and penalizes the consumers when demand exceeds supply. Finally, the term $y_1(t)W$ represents the producing costs, i.e., whenever a prosumer starts producing, the cost is W. In addition to this, the total cost includes the running cost $c(\cdot)$ plus some additional terminal cost $g : \mathbb{R} \to [0, +\infty)$, $x(T) \mapsto g(x(T))$, to be yet designed.

We will now specify the relation between $y_1(t)$ and $y_2(t)$. To simplify our dynamics, we assume that $y_2(t) = 1 - y_1(t)$, that is, the probability of being in the mode of consumer is simply the complement of the probability of being in the mode of producer. By imposing $y_1(t) + y_2(t) = 1$, the variable $y_2(t)$ becomes redundant and the above model simplifies as follows

$$\dot{x}(t) = \begin{cases} p - d - \lambda x(t) & \text{if } y_1(t) = 1, \\ -d - \lambda x(t) & \text{if } y_1(t) = 0, \end{cases} \quad t \in [0,T), \quad (7)$$

and the complete dynamics are governed by

$$\begin{aligned} \dot{x}(t) &= (p - d - \lambda x(t))y_1(t) + (-d - \lambda x(t))y_2(t) \\ &= (p - d - \lambda x(t))y_1(t) + (-d - \lambda x(t))(1 - y_1(t)) \\ &= py_1(t) - d - \lambda x(t), \\ \dot{y}_1(t) &= u_1(t) - u_2(t). \end{aligned} \quad (8)$$

With the constraint $y_2(t) = 1 - y_1(t)$, we note that $m(\cdot) = m(x, y_1, t)$ and

$$\begin{aligned} m_p(t) &= \int_{\mathbb{R}} \int_0^1 y_1 m(x, y_1, t)dxdy_1, \\ m_c(t) &= 1 - m_p(t). \end{aligned} \quad (9)$$

For the error we now have that $e(t) = (p + d)m_p(t) - d$.

The problem we consider now is the following. Given a finite horizon $T > 0$ and an initial distribution $m_0 : \mathbb{R} \times [0,1] \to [0, +\infty)$, subject to the dynamics in system (8), we want to minimize the following cost functional

$$J(x,y,t,u) = \int_0^T c(x(t), y_1(t), u(t), m(x, y_1, t))dt + g(X(T)),$$

over all possible inputs $u(\cdot)$. Here, $m(\cdot)$ is the time-dependent function describing the evolution of the mean of the distribution of the state of the prosumer.

III. First-order mean field game

We will first review the concept of first-order mean field game. In a first-order mean field game, the microscopic dynamics is deterministic and the resulting mean-field game involves only the first derivatives of the value function and of the density function. Consider a generic cost and dynamics

$$J(X(t),u(t),t) = \inf_{u\in\mathbb{R}^\bullet} \int_0^T c(X(t),u(t),m)dt + g(X(T)),$$
$$\dot{X}(t) = F(X(t),u(t)) \quad (10)$$

where $c(\cdot)$ is the running cost, $g(X(T))$ is the terminal penalty, and where $u(\cdot)$ is any state feedback closed loop control. Let now $v(X,t)$ be the value function, i.e., the optimal value of $J(X,U,t)$. Then from [2] it is well known that the problem results in the following mean field game system

$$\begin{cases} \partial_t v(X,t) + \partial_X v(X,t)^T \cdot F(X,u^*) + c(X,u^*,m) = 0, \\ v(X,T) = g(X(T)), \\ u^*(X(t)) = \arg\min_{u\in\mathbb{R}^\bullet} \left\{\partial_X v(X,t)^T \cdot F(X,u) + c(X,u,m)\right\} \end{cases} \quad (11)$$

$$\begin{cases} \partial_t m(X,t) + \mathrm{div}(F(X,u^*)) \cdot m(X,t) = 0 \\ m(X,0) = m_0(X). \end{cases} \quad (12)$$

The first partial differential equation in (11) is the Hamilton-Jacobi-Bellman equation which returns the value function $v(X,t)$, once we fix the distribution $m(X,t)$. This partial differential equation has to be solved backwards with the boundary condition $v(X,T) = g(X(T))$ on the final time T, represented by the second line in (11). In the third line of (11) we have the optimal closed-loop control $u^*(X(t))$ as minimizer of the Hamiltonian function in the right-hand side. The partial differential equation in (12) represents the transport equation of the function m immersed in a vector field $F(X,u^*(X))$. It returns the distribution $m(X,t)$ once we fixed the optimal closed-loop control $u^*(X(t))$, and consequently the vector field $F(X,u^*(X))$. Such a partial differential equation has to be solved forward with the boundary condition $m(X,0) = m_0(X)$ at the initial time, represented by the last line of (12). Finally, once we know $m(X,t)$, it can be put into the running cost $c(X,u,m)$ and the error can be obtained from (5) with (9). We remark that $\bar{X}(t)$ is given by

$$\begin{bmatrix}\bar{x}(t) \\ \bar{y}_1(t)\end{bmatrix} = \begin{bmatrix}\int_{\mathbb{R}} \int_0^1 x m(x,y_1,t)dxdy_1 \\ \int_{\mathbb{R}} \int_0^1 y_1 m(x,y_1,t)dxdy_1\end{bmatrix} = \begin{bmatrix}\bar{x}(t) \\ m_p(t)\end{bmatrix},$$

and hence, a mean field equilibrium solution is any pair $(v(X,t),\bar{X}(t))$ that satisfies (11) and (12).

We will now proceed as follows. The goal of the paper is to find a mean field equilibrium in the case of prosumers. To do so, we look at the mean field best response u^* to a given distribution $m(\cdot)$. We note that the problem we consider is of the form (10), if we plug in the running cost (6) together with the dynamics (8), where $X(t) = [x(t), y_1(t)]^T$ and $u(t) = [u_1(t), u_2(t)]^T$.

IV. Mean field equilibrium

Considering the problem of minimizing the generic cost in (6) subject to dynamics as displayed in (8), we note that we have a linear-quadratic optimization problem of the form (10), namely:

$$\inf_{u\in\mathbb{R}^2} \int_0^T \left(\tfrac{1}{2}X^TQX + \tfrac{1}{2}u^TRu^T + L^TX\right)dt + g(X(T)),$$
$$\dot{X}(t) = AX(t) + Bu(t) + C, \quad (13)$$

where we recall that $X(t) = [x(t), y_1(t)]^T$ and where $x(t)$ denotes the inventory level and $y_1(t)$ denotes the probability of being in the mode of consumer. Furthermore we have

$$A = \begin{bmatrix}-\lambda & p \\ 0 & 0\end{bmatrix}, \quad B = \begin{bmatrix}0 & 0 \\ 1 & -1\end{bmatrix}, \quad C = \begin{bmatrix}-d \\ 0\end{bmatrix},$$

and also

$$Q = \begin{bmatrix}q & 0 \\ 0 & 0\end{bmatrix}, \quad R = \begin{bmatrix}r_1 & 0 \\ 0 & r_2\end{bmatrix}, \quad L = \begin{bmatrix}0 \\ Se(t) + W\end{bmatrix}.$$

The linear quadratic problem posed in (13) is easily solved by the use of calculus of variations [10], and we are now in a position to state and prove the main result.

Theorem 1: A mean field equilibrium for (11) with (12) is given by

$$v(X,t) = \tfrac{1}{2}X^TP(t)X + \Psi^T(t)X + \chi(t),$$
$$\dot{\bar{X}}(t) = [A - BR^{-1}B^TP]\bar{X}(t) - BR^{-1}B^T\Psi + C, \quad (14)$$

where

$$\bar{\Psi}(t) = \int_{\mathbb{R}} \int_0^1 \Psi(t) m(x,y,t) dxdy_1,$$

and P, Ψ, χ can be found by solving the following Riccati equations

$$\dot{P} + PA + A^TP - PBR^{-1}B^TP + Q = 0,$$
$$\dot{\Psi} + A^T\Psi + PC - PBR^{-1}B^T\Psi + L = 0, \quad (15)$$
$$\dot{\chi} + \Psi^TC - \tfrac{1}{2}\Psi^TBR^{-1}B^T\Psi = 0,$$

with boundary condition

$$P(T) = \phi, \ \Psi(T) = 0, \ \chi(T) = 0, \ g(X(T)) = \tfrac{1}{2}X^T\phi X.$$

Furthermore, the mean field equilibrium strategy is

$$u^*(X,t) = -R^{-1}B^T(PX + \Psi).$$

Proof: As candidate function $v(\cdot)$ that minimizes $J(\cdot)$ we take

$$v(X,t) = \tfrac{1}{2}X^TP(t)X + \Psi^T(t)X + \chi(t),$$

where P is an unknown symmetric matrix, Ψ is an unknown vector, and χ is an unknown scalar function. Then we have

$$\partial_t v(X,t) = \tfrac{1}{2}X^T\dot{P}(t)X + \dot{\Psi}^T(t)X + \dot{\chi}(t),$$
$$\partial_X v(X,t) = PX + \Psi.$$

Now, the optimal state feedback is given by

$$u^*(X,t) = \arg\min_{u\in\mathbb{R}^2} \{\partial_X v(X,t)^T \cdot F(X,u) + c(X,u,m)\},$$

and with substitution of (13) we find

$$u^*(X,t) = \arg\min_{u\in\mathbb{R}^2}\{(PX+\Psi)^T\cdot(AX+Bu+C) + \tfrac{1}{2}X^TQX + \tfrac{1}{2}u^TRu + L^TX\}.$$

Finding u^* is straightforward, we can just take the gradient in the above expression and set it equal to zero to find the minimum, which is

$$u^* = -R^{-1}B^T(PX+\Psi).$$

The corresponding Hamilton-Jacobi equation is now

$$\partial_t v(X,t) + \partial_X v(X,t)^T\cdot F(X,u^*(X)) + c(X,u^*(X),m) = 0.$$

And substituting the expression for u^* and rewriting we find

$$\begin{aligned}0 &= \tfrac{1}{2}X^T\dot{P}(t)X + \dot{\Psi}^T(t)X + \dot{\chi}(t) + (PX+\Psi)^T\cdot\\ &\quad\cdot(AX+Bu^*+C) + \tfrac{1}{2}X^TQX + \tfrac{1}{2}(u^*)^TRu^* + L^TX\\ &= \tfrac{1}{2}X^T\dot{P}(t)X + \dot{\Psi}^T(t)X + \dot{\chi}(t) + (X^TP+\Psi^T)\cdot\Big(AX\\ &\quad - BR^{-1}B^T(PX+\Psi)+C\Big) + \tfrac{1}{2}X^TQX\\ &\quad + \tfrac{1}{2}(X^TP+\Psi^T)BR^{-1}B^T(PX+\Psi) + L^TX.\end{aligned}$$

After performing some algebraic manipulations we find

$$\begin{aligned}0 &= \tfrac{1}{2}X^T\dot{P}(t)X + \dot{\Psi}^T(t)X + \dot{\chi}(t) + \tfrac{1}{2}X^TPAX\\ &\quad + \tfrac{1}{2}X^TA^TPX - X^TPBR^{-1}B^TPX\\ &\quad - \Psi^TBR^{-1}B^TPX + X^TPC + \Psi^TAX\\ &\quad - \Psi^TBR^{-1}B^TPX - \Psi^TBR^{-1}B^T\Psi + \Psi^TC\\ &\quad + \tfrac{1}{2}X^TQX + \tfrac{1}{2}X^TPBR^{-1}B^TPX + \Psi^TBR^{-1}B^TPX\\ &\quad + \tfrac{1}{2}\Psi^TBR^{-1}B^T\Psi + L^TX.\end{aligned}$$

And grouping the terms yields

$$\begin{aligned}0 &= X^T\Big(\tfrac{1}{2}\dot{P} + \tfrac{1}{2}PA_i + \tfrac{1}{2}A_i^TP - PB_iR^{-1}B_i^TP + \tfrac{1}{2}Q\\ &\quad + \tfrac{1}{2}PB_iR^{-1}B_i^TP\Big)X + X^T\Big(\dot{\Psi}(t) - PB^TR^{-1}B^T\Psi\\ &\quad + A^T\Psi + PC - PB_iR^{-1}B^T\Psi + PBR^{-1}B^T\Psi + L\Big) +\\ &\Big(\dot{\chi}(t) + \Psi^TC - \Psi^TB_iR^{-1}B_i^T\Psi + \tfrac{1}{2}\Psi^TB_iR^{-1}B_i^T\Psi\Big).\end{aligned}$$

Now, if the terms between brackets in the above expression are zero, then at least the right hand side equals zero. So in order to have a sufficient condition for a solution the expressions between brackets need to be zero. These are precisely the required Riccati equations. ■

To practically implement the above result, we make the following remark. Often a discretized version of Theorem 1 is considered, which is useful in numerical simulations, of which we will give an example in Section VI. In the spirit of receding horizon, at every iteration, instead of solving the Riccati equations (15) for a finite horizon window, one solves the Riccati equations corresponding to the infinite horizon control problem

$$\begin{aligned}PA + A^TP - PBR^{-1}B^TP + Q &= 0,\\ A^T\Psi + PC - PBR^{-1}B^T\Psi + L &= 0,\\ \Psi^TC - \tfrac{1}{2}\Psi^TBR^{-1}B^T\Psi &= 0.\end{aligned} \tag{16}$$

So, at every time instance the above equations are solved and consequently the optimal control input u^* is computed. This is fed through giving us a new state and we repeat the procedure. We will also use this algorithm when we perform the numerical experiment in Section VI. However, before doing so, we elaborate in the next section on the link between the optimal control law u^* and the well-known Bass model.

V. Link to Bass Model

The Bass model is a commonly used diffusion model and is often used to study for example the way in which a disease spreads over a population, or the way in which consumers adapt to a new technology, [9]. To understand the link with the Bass model, we note that the optimal control law $u^* = -R^{-1}B^T(PX+\Psi)$ can be rewritten as

$$u^* = \begin{bmatrix}u_1^*\\ u_2^*\end{bmatrix} = \begin{bmatrix}\frac{1}{r_1}P_2x + \frac{1}{r_1}P_3y_1 + \frac{1}{r_1}\Psi_2\\ \frac{1}{r_2}P_2x + \frac{1}{r_2}P_3y_1 + \frac{1}{r_2}\Psi_2\end{bmatrix}, \tag{17}$$

where

$$P = \begin{bmatrix}P_1 & P_2\\ P_2 & P_3\end{bmatrix},\quad \Psi = \begin{bmatrix}\Psi_1\\ \Psi_2\end{bmatrix},$$

are the solutions to the following equations

$$\begin{aligned}PA + A^TP - PBR^{-1}B^TP + Q &= 0,\\ \left(A^T - PBR^{-1}B^T\right)\Psi &= -L - PC.\end{aligned}$$

Now, we remark that the optimal controls u_1^* and u_2^* of (17) are of the form

$$\begin{aligned}u_1^* &= \alpha_{1,1}x + \alpha_{2,1}y_1 + c_1,\\ u_2^* &= \alpha_{1,2}x + \alpha_{2,2}y_1 + c_2.\end{aligned} \tag{18}$$

This is also depicted in Figure 1 below. We now highlight

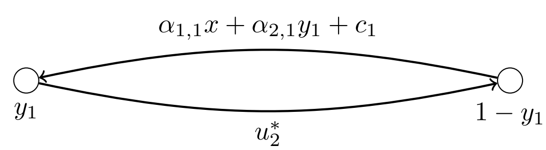

Fig. 1: Transition rates from producer, y_1, to consumer, y_2, and vice versa.

that this shows strong connections to the Bass model. The idea of the Bass model is that we have a large population and each individual can be classified as a susceptible or infected player. Let y be the fraction of infected people, so $1-y$ denotes the fraction of the population that is susceptible. In the Bass model, the rate of change of y depends on two factors. First of all, susceptible agents can spontaneously become infected, which is represented by a parameter c. Secondly, the process of imitation or the word-of-mouth effect takes place. Susceptible agents become infected due to a factor proportional to the fraction of infected individuals, quantified by αy. In a nutshell, the Bass model is then described by the following differential equation:

$$\dot{y} = (c+\alpha y)(1-y). \tag{19}$$

We now remark that the dynamics of mode y_1, i.e., the probability of being a producer, shows striking similarities with the Bass model. In fact, the optimal control law u_1^* from (17) is an adjusted Bass model where

- The term $\frac{1}{r_1}\Psi_2$ accounts for spontaneously infection.
- The term $\frac{1}{r_1}P_3 y_1$ is the imitation factor or word-of-mouth rate.
- The term $\frac{1}{r_1}P_2 x$ is a correction term, which takes into account the inventory level $x(t)$. This term can be interpreted as a disturbance.

We remark that in the original Bass model, the word-of-mouth factor induces crowd seeking behavior, since a higher word-of-mouth rate implies that more people will become infected. However, in the adjusted Bass model with which we deal here, the word-of-mouth factor corresponds to crowd averse behavior, since an agent will be in the mode of consumer if a large fraction of the population is in the mode of producer.

An argument for the presence of the correction term is the following. Consider a single prosumer characterized by its two states: the inventory level x and the mode y_1. Then, the evolution of y_1 is not solely governed by the global interest given by the general Bass model. At a local level the agent also has a local objective, governed by the running costs function (6), so the correction term based on the state x comes into play.

Finally, we remark that we have an explicit solution to the infinite time horizon Riccati equations, so we can establish a priori the optimal control law u^*. Equivalently one can say that the model parameters in the adjusted Bass model are known. All this is explored in the following theorem.

Theorem 2: Consider the optimal control law $u^* = -R^{-1}B^T(PX + \Psi)$, which is of the form

$$\begin{aligned} u_1^* &= \alpha_{1,1}x + \alpha_{2,1}y_1 + c_1, \\ u_2^* &= \alpha_{1,2}x + \alpha_{2,2}y_1 + c_2. \end{aligned}$$

Here, the coefficients (which are the parameters of the Bass model) are given by

$$\begin{aligned} \alpha_{1,i} &= \tfrac{1}{r_i}P_2, \quad \alpha_{2,i} = \tfrac{1}{r_i}P_3, \\ c_i &= \frac{-pdP_1 + \lambda(Se+W) - \lambda dP_2}{r_i\left(\frac{1}{r_1}+\frac{1}{r_2}\right)(\lambda P_3 + pP_2)}, \end{aligned}$$

where

$$\begin{aligned} P_1 &= \frac{\dfrac{-\left(\frac{1}{r_1}+\frac{1}{r_2}\right)^3 P_3^4}{4p^2} + q}{2\lambda}, \quad P_2 = \frac{P_3^2 r}{2p}, \\ P_3 &= \frac{2p}{\left(\frac{1}{r_1}+\frac{1}{r_2}\right)} \pm \frac{2\sqrt{q}p}{\left(\frac{1}{r_1}+\frac{1}{r_2}\right)^{\frac{3}{2}}}. \end{aligned}$$

Proof: From the optimal control law $u^* = -R^{-1}B^T(PX + \Psi)$, we obtain

$$\begin{aligned} u_1^* &= \tfrac{1}{r_1}P_2 x + \tfrac{1}{r_1}P_3 y_1 + \tfrac{1}{r_1}\Psi_2, \\ u_2^* &= \tfrac{1}{r_2}P_2 x + \tfrac{1}{r_2}P_3 y_1 + \tfrac{1}{r_2}\Psi_2. \end{aligned}$$

Here, the values of P_2, P_3, Ψ_2 can be found by solving the equations

$$\begin{aligned} PA + A^TP - PBR^{-1}B^TP + Q = 0, \\ \left(A^T - PBR^{-1}B^T\right)\Psi = -L - PC, \end{aligned}$$

since these equations have as solution

$$P = \begin{bmatrix} P_1 & P_2 \\ P_2 & P_3 \end{bmatrix}, \quad \Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix}.$$

So we inspect

$$\begin{aligned} PA + A^TP - PBR^{-1}B^TP + Q = 0 \\ \begin{bmatrix} -\lambda P_1 & pP_1 \\ -\lambda P_2 & pP_2 \end{bmatrix} + \begin{bmatrix} -\lambda P_1 & -\lambda P_2 \\ pP_1 & pP_2 \end{bmatrix} \\ - \begin{bmatrix} P_2^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right) & P_2P_3\left(\frac{1}{r_1}+\frac{1}{r_2}\right) \\ P_2P_3\left(\frac{1}{r_1}+\frac{1}{r_2}\right) & P_3^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right) \end{bmatrix} + \begin{bmatrix} q & 0 \\ 0 & 0 \end{bmatrix} = 0_{2\times 2}. \end{aligned}$$

This yields the following set of equalities

$$\begin{aligned} -2\lambda P_1 - P_2^2 r + q = 0, \\ pP_1 - \lambda P_2 - P_2P_3 r = 0, \\ 2pP_2 - P_3^2 r = 0, \end{aligned}$$

where we simplified the notation by introducing $r = \left(\frac{1}{r_1}+\frac{1}{r_2}\right)$. From the first and third equations, we obtain

$$P_1 = \frac{-\dfrac{r^3P_3^4}{4p^2} + q}{2\lambda}, \quad P_2 = \frac{P_3^2 r}{2p}, \tag{20}$$

and substituting these equations in the second equation yields

$$\frac{P_3^2 r}{2} + \frac{\dfrac{r^3P_3^4}{4p^2} - q}{2} - \frac{P_3^3 r^2}{2p} = 0.$$

After multiplication by $8p^2$ we obtain that either $P_3 = 0$ or

$$4p^2 r + r^3P_3^2 - 4qp^2 - 4pr^2P_3 = 0.$$

This is a quadratic solution in P_3 and has as solution

$$P_3 = \frac{4pr^2 \pm \sqrt{16p^2r^4 - 4r^3(4p^2r - 4qp^2)}}{2r^3} = \frac{2p}{r} \pm \frac{2\sqrt{q}p}{r\sqrt{r}}.$$

This can now be plugged back in (20) to find expressions for P_1 and P_2. To find Ψ, we inspect

$$\begin{aligned} (A^T - PBR^{-1}B^T)\Psi = -L - PC, \\ \begin{bmatrix} -\lambda & -P_2\left(\frac{1}{r_1}+\frac{1}{r_2}\right) \\ p & -P_3\left(\frac{1}{r_1}+\frac{1}{r_2}\right) \end{bmatrix} \begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix} = \begin{bmatrix} P_1 d \\ P_2 d - (Se+W) \end{bmatrix}. \end{aligned}$$

We write down the solution of this system of linear equations directly, by using the adjugate of a matrix

$$\begin{aligned} \Psi_1 &= \frac{-P_3 - P_2(Se+W) + P_2^2 d}{\lambda P_3 + pP_2}, \\ \Psi_2 &= \frac{-pdP_1 + \lambda(Se+W) - \lambda dP_2}{\left(\frac{1}{r_1}+\frac{1}{r_2}\right)(\lambda P_3 + pP_2)}. \end{aligned}$$

Note that both Ψ_1 and Ψ_2 depend on the error e, so here the cross-coupling comes into play. This completes the exposition. ∎

VI. Numerical experiment

We start with a population of $n = 200$ prosumers and we run the simulation up to $T = 40$. The production and demand rates are $p = 5$ and $d = 3$, and we take $\lambda = 1$. Finally, as penalties on the input we take $r_1 = r_2 = 10$ and the producing costs are set to $W = 10$.

To run the simulation, we apply a discretized version of (14). At every iteration, for every agent, we compute the best response u^* for the infinite horizon control problem, which is then fed through to compute the next state. Equivalent to solving the Riccati equation using the build-in Matlab command `care`, we can also use the direct result established in Theorem 2. We now have

$$X(t+1) = \left(\left[A - BR^{-1}B^T P\right] X(t) - BR^{-1}B^T\Psi + C\right) dt + X(t).$$

In the first simulation, we take a high penalty on the

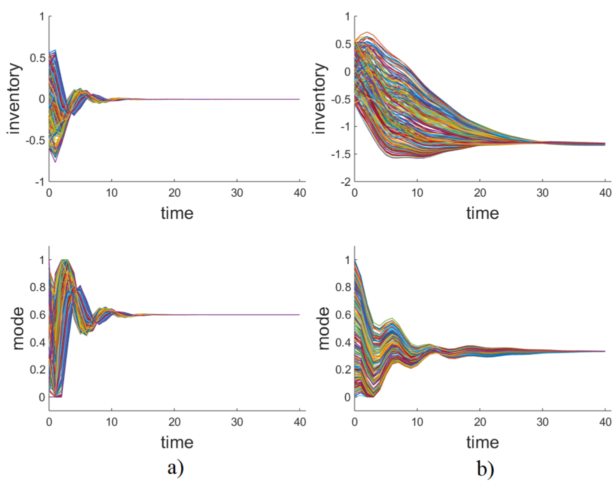

Fig. 2: a) time plots of the inventory level $x(t)$ and the mode $y_1(t)$ for a large q and small S. Fig. 2: b) time plots of the inventory level $x(t)$ and the mode $y_1(t)$ for a small q and large S.

inventory level, $q = 100$, and a relative low penalty on the error term, $S = 10$. Due to the significant penalty on the inventory level, we expect that the inventory levels are pushed to zero. This can also be observed in Figure 2a, where we have plotted the results of our simulation. Here, having no inventory is the desired case, since a shortage in inventory means that you could not match the demand and one has a lower profit, while on the other hand a positive inventory level means storage costs, and this is also undesirable. By introducing a high costs q on the inventory level $x(t)$ in the running cost function we penalize for these cases.

Furthermore we note that the mode y_1, i.e., the probability of being in the mode of producer, converges to 0.60. This is a sensible result, since the dynamics are governed by $\dot{x} = py_1 - d - \lambda x$ and this expression is equal to zero if and only if $y_1 = \frac{d}{p}$.

We now perform a second simulation with a high penalty on the error, $S = 100$, while $q = 10$. The results are shown in Figure 2b. We expect that for this case the error is pushed to zero, due to the high penalty. Since $e = (p + d)\bar{y}_1 - d$, we note that $e = 0$ if and only if $\bar{y}_1 = \frac{d}{p+d} = 0.375$. We observe from the plotted results in Figure 2b that the mode y_1 indeed converges to this value. Furthermore we note that the inventory level does not converge to zero, since the error is penalized more urgently. We have $\dot{x} = py_1 - d - \lambda x$ and since y_1 converges to 0.375, we have that the inventory level converges to $x = \frac{py_1 - d}{\lambda} = -1.125$. This can also be seen from the simulation results.

VII. Conclusion

In this paper we studied mean field games on prosumers. A prosumer is an agent that can produce or consume, and its behavior depends on the average behavior of the population. Every prosumer is characterized by three states: the inventory level, the probability of being in the mode of producer, and the probability of being in the mode of consumer. We simplified the problem by taking the probability of being a consumer equal to the complement of the probability of being a producer. We formulated the problem as a first order mean field game, and we found a mean field game equilibrium. We then established a clear connection to the Bass model.

In future research, we would like to add deterministic and stochastic disturbances to the dynamics, and study robustness properties. Finally we would like to study the case where the probability of being a consumer is not the complement of the probability of being a producer. In such a case, we add a third state corresponding to being uncommitted (idle), the prosumer is then not producing nor consuming.

References

[1] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions, "Mean field games and applications," in *Paris-Princeton lectures on mathematical finance 2010*, pp. 205–266. Springer, 2011.

[2] Jean-Michel Lasry and Pierre-Louis Lions, "Mean field games," *Japanese journal of mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[3] Minyi Huang, Peter E. Caines, and Roland P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: Centralized and nash equilibrium solutions," *IEEE Conference on Decision and Control, HI, USA, December,*, pp. 98–103, 2003.

[4] Minyi Huang, Peter E. Caines, and Roland P. Malhamé, "Large population stochastic dynamic games: Closed loop kean-vlasov systems and the nash certainty equivalence principle," *ommunications in Information and Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[5] Minyi Huang, Peter E. Caines, and Roland P. Malhamé, "Large population cost-coupled lqg problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.

[6] Diogo A. Gomes and Joào Saúde, "Mean field games models – a brief survey," *Dynamic Games and Applications*, vol. 4, no. 2, pp. 110–154, 2014.

[7] Dario Bauso, *Game Theory with Engineering Applications*, vol. 30, SIAM, Philadelphia, PA, USA, 2016.

[8] Fabio Bagagiolo and Dario Bauso, "Mean-field games and dynamic demand management in power grids," *Dynamic Games and Applications*, vol. 4, no. 2, pp. 155–176, 2014.

[9] Matthew O. Jackson, *Social and Economic Networks*, Princeton University Press, 1st edition edition, 2010.

[10] Arjan J. van der Schaft, *Calculus of Variations*, Lecture Notes, University of Groningen, 2019.