Proceedings Paper:
Bauso, D., Mylvaganam, T. and Astolfi, A. (2014) Approximate solutions for crowd-averse robust mean-field games. In: 2014 European Control Conference, ECC 2014. 2014 European Control Conference (ECC), June 24-27, 2014, Strasbourg, France. Institute of Electrical and Electronics Engineers , 1217 - 1222. ISBN: 9783952426913.

https://doi.org/10.1109/ECC.2014.6862413

Approximate solutions for crowd-averse robust mean-field games

D. Bauso and T. Mylvaganam and A. Astolfi

***Abstract*— We consider a population of dynamic agents (players). The state of each player evolves according to a linear stochastic differential equation driven by a Brownian motion and under the influence of a control and an adversarial disturbance. Every player minimizes a cost functional which involves quadratic terms on state and control plus a cross-coupling mean-field term measuring the congestion resulting from the collective behavior, which motivates the term "crowd-averse". For this game we first illustrate the paradigm of robust mean-field games. Second, we provide a new approximate solution approach based on the extension of the state space and prove the existence of equilibria and their stability properties. Third, we provide a bound for the approximation introduced by the solution method. Simulations illustrating the approximate solution are presented.**

I. INTRODUCTION

We illustrate the robust mean-field game approach on a population of dynamic agents that wish to regulate their state to zero. Each agent's state evolves according to a linear stochastic differential equation (SDE) driven by a Brownian motion and under the influence of a control and an adversarial disturbance. The control minimizes a cost functional which involves quadratic terms on state and control plus a cross-coupling mean-field term involving the control of the single player and the average control computed over all players. Such a term allows the redistribution of the control load away from peak "hours" thus reducing congestion, from which the term "crowd-averse". Indeed every player pays a cost from controlling its own system when the population as a whole has a high average control. Based on the provided mean-field game formulation we analyze both the microscopic evolution of each player and the macroscopic evolution of the system as a whole.

The work of D. Bauso has been 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".

Dario Bauso is with Dipartimento di Ingegneria Chimica, Gestionale, Informatica e Meccanica, Università di Palermo, Italy, and Dipartimento di Matematica, Università di Trento (as research fellow), Via Sommarive 14, 38050 Povo-Trento, Italy. D. Bauso is currently academic visitor at the Department of Engineering Science, University of Oxford, United Kingdom. email: *dario.bauso@unipa.it*

T. Mylvaganam is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK (Email: thulasi.mylvaganam06@imperial.ac.uk).

A. Astolfi is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK and the Dipartimento di Ingegneria Civile e Ingegneria Informatica, Università di Roma "Tor Vergata", Via del Politecnico, 1 00133 Roma, Italy (Email: a.astolfi@ic.ac.uk).

Highlights of contributions. We highlight three main contributions. First, we establish a robust mean-field system for the considered game under adversarial disturbances. Second, we provide a new approximate solution approach based on the extension of the state space in the same spirit as [14], [15]. The method allows to prove the existence of equilibria and their stability properties. Third, we provide a bound for the approximation introduced by the solution method.

Related literature on mean-field games. Mean-field games were formulated by Lasry and Lions in [10] and independently by M.Y. Huang, P. E. Caines and R. Malhamé in [7], [8]. The mean-field theory of dynamical games is a modeling framework at the interface of differential game theory, mathematical physics, and H_∞-optimal control that tries to capture the mutual influence between a crowd and its individuals. Mean-field games arise in several application domains such as economics, physics, biology, and network engineering (see [1], [5], [6], [8], [9], [13], [17]).

From a mathematical point of view the mean-field approach leads to a system of two partial differential equations (PDEs). The first PDE is the Hamilton-Jacobi-Bellman equation. The second PDE is the Fokker-Planck equation which describes the density of the players [10], [16]. Explicit solutions in terms of mean-field equilibria are not common unless the problem has a linear-quadratic structure, see [3]. In this sense, a variety of solution schemes has been recently proposed based on discretization and/or numerical approximations [1]. More recently, robustness and risk-sensitivity have been brought into the picture of mean-field games [4], [16], where the first PDE is now the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation.

The paper is organized as follows. In Section II we formulate the problem. In Section III we provide some motivations. In Section IV we derive the mean-field game. In Section V we introduce the approximate solution approach and study equilibria and stability properties. In Section VI we carry out some numerical studies. Finally in Section VII we provide some conclusions.

Notation We denote by $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space. We let $\mathcal{B}$ be a finite-dimensional Brownian motion defined on this probability space. Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be its natural filtration augmented by all the $\mathbb{P}-$null sets (sets of measure-zero with respect to $\mathbb{P}$). We use ∂_x and ∂^2_{xx} to denote the first and second partial derivatives with respect to x, respectively.

II. Problem set-up

Consider a game with an infinite number of homogeneous players. For each player let x_0 be its initial state, which is realized according to the probability distribution m_0. The state of the player at time t, denoted by $x_t \in \mathbb{R}$, evolves according to a controlled stochastic process over a finite horizon $T > 0$, i.e.

$$dx_t = [\alpha x_t + \beta u_t]dt + \sigma [x_t d\mathcal{B}_t + \zeta_t dt], \quad (1)$$

where $u_t \in \mathbb{R}$ is the control input, $\mathcal{B}_t \in \mathbb{R}$ is a Brownian motion, which is independent of the initial state x_0, and independent across players and time. The constants $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$ and $\sigma \in \mathbb{R}$ are parameters, and $\zeta_t \in \mathbb{R}$ is an adversarial disturbance.

To introduce a macroscopic description of the game consider probability density functions on the state and control spaces:

$$\begin{cases} m : \mathbb{R} \times [0, +\infty[\to [0, +\infty[,\ (x,t) \mapsto m(x,t) \\ \int_{\mathbb{R}} m(x,t)dx = 1 \text{ for every } t, \end{cases}$$

and

$$\begin{cases} z : \mathbb{R} \times [0, +\infty[\to [0, +\infty[,\ (u,t) \mapsto z(u,t), \\ \int_{\mathbb{R}} z(u,t)du = 1 \text{ for every } t. \end{cases}$$

Define now the average state and control distributions at time t as

$$\begin{cases} \bar{m}_t := \int_{\mathbb{R}} x m(x,t)dx, \\ \\ \bar{z}_t := \int_{\mathbb{R}} u z(u,t)du. \end{cases}$$

Finally we introduce a cost functional with penalty on the final state $g(\cdot)$, stage cost function $c(\cdot)$, and quadratic penalty on the unknown disturbance:

$$\begin{aligned} &J(x_0, u, m, \zeta) = \mathbb{E}\Big(g(x_T) \\ &+ \int_0^T c(x_t, u_t, \bar{z}_t)dt - \gamma^2 \int_0^T |\zeta_t|^2 dt\Big). \end{aligned} \quad (2)$$

Players wish to stabilize their state to zero, and therefore we can select the stage cost

$$c(x_t, u_t, \bar{z}_t, \zeta_t) = \frac{h}{2}\bar{z}_t u_t^2 + \left[\frac{a}{2}x_t^2 + \frac{b}{2}u_t^2\right],$$

with $h \geq 0$. The term $\frac{h}{2}\bar{z}_t u_t^2$ represents a cross-term coupling the control of each player and the average control of the population; $\frac{a}{2}x_t^2$, with $a > 0$, is the cost of a non-zero state, and $\frac{b}{2}u_t^2$, with $b > 0$, accounts for a penalty on the control energy. The penalty on the final state $g(x_T)$ is, in general, convex with minimum in zero, thus penalizing non-zero states at the end of the horizon.

Note that the mean of the state is generated by

$$\frac{d}{dt}\bar{m}_t = \alpha \bar{m}_t + \beta \bar{z}_t + \sigma \bar{\zeta}_t.$$

Considering deterministic disturbance ζ_t, and using indistinguishability, we find that the mean of the average control evolves according to:

$$\bar{z}_t = \frac{1}{\beta}\left(\frac{d}{dt}\bar{m}_t\right) - \frac{\alpha}{\beta}(\bar{m}_t) - \frac{\sigma}{\beta}\bar{\zeta}_t.$$

A relation between $\frac{d}{dt}\bar{m}_t$ and $\bar{m}_t$ is yet to be introduced. However, we will see later that both $\frac{d}{dt}\bar{m}_t$ and $\bar{\zeta}_t$ can be approximated by linear functions in $\bar{m}_t$ and therefore we can rewrite

$$\bar{z}_t = \hat{k}\bar{m}_t, \quad (3)$$

for some $\hat{k} \in \mathbb{R}$. The above preamble leads to the following robust mean-field game problem.

Problem 1: (**Robust mean-field problem**) Let $\mathcal{B}$ be a one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is the natural filtration generated by $\mathcal{B}$. Let x_0 be independent of $\mathcal{B}$ and with density $m_0(x)$. Let m^*_t be the optimal mean-field trajectory. The robust mean-field problem in $\mathbb{R}$ and $(0, T]$ is given by

$$\begin{cases} \inf_{\{u_t\}_t} \sup_{\{\zeta_t\}_t} J(x, u, m^*, \zeta) \\ dx_t = [\alpha x_t + \beta u_t + \sigma \zeta_t]\, dt \quad + \quad \sigma x_t d\mathcal{B}_t. \end{cases}$$

III. Motivations

We provide three different interpretations of the problem. The first is an example of inventory control. The second is the description of a congestion control problem in networked controlled systems or power grids. Finally the third is an example from economics and describes an oil production applications.

Example 1: **(Inventory control with shared set-up costs [13])** In multi-retailer inventory control equation (1) describes the evolution of the inventory over time. The control is the reordered quantity and the disturbance is the unknown market demand. A classical scenario is where the transportation cost is shared among all retailers who reorder at a given time instant, called active retailers. Then a certain level of coordination of the retailers' replenishment strategies may lead to individual costs reduction. Thus the cross mean-field term in the objective function (2) accounts for the reduced cost when orders are placed jointly. The other two terms are usually the costs of reordering and shortage or the holding costs on inventory. Clearly, we can generalize the framework to any application where multiple players share a service facility as airport facilities or telephone systems, drilling for oil, cooperative farming, and fishing (see also the references on cost-sharing games in [13]).

Example 2: **(Dynamic demand management in power grids [2], [12])** Players are electrical appliances, say for instance heating or cooling appliances, and their state is their temperature at a given time. Each single appliance can be in one of the two states ON or OFF. The dynamics (1) describe the time evolution of the temperature of each appliance. Each single controller is given a cost function that accounts for i) the energy consumption, which is captured by the penalty on the control, ii) the deviation of the mains frequency from the nominal value, represented by the cross-term, and iii) the deviation of the agent's temperature from the reference value, described by the penalty on the state. With respect to goal ii), the cross mean-field term incentivizes the appliances 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.

Example 3: **(Oil production [4], [6])** Suppose we have a finite number of oil producers, and let the state be the stock of raw material available at a given time. Let the control be the produced oil quantity by a single producer and the adversarial disturbance be a cautious disturbance parameter reflecting the taxation or inflation on the produced quantity. Equation (1) is widely used in stock market models as it describes the variation of the reserve at time t given the current reserve and the consumed resource quantity. The term $\sigma_t\zeta_t$ is intended to capture the negative and uncertain influence of taxation, or inflation, on the production. The cost functional, $\hat{k}\bar{m}_t$ is the sale's price of oil and the cross-term is related to the income collected from producing and selling the quantity u_t; $\frac{a}{2}(x_t)^2$ accounts for a production energy consumed, $a > 0$ and bu_t^2 is a known linear taxation on production. The penalty on the final state $g(x_T)$ can be assumed quadratic in the reserve, so that unexploited reserve at the end of the horizon is penalized.

IV. The resulting mean-field games

Let $v_t(x)$ be the (upper) value of the robust optimization problem under worst-case disturbance starting at time t from state x. Let the corresponding Hamiltonian be given by

$$H(x,p,m) = \inf_u \left\{ c(x,u,m) + p(\alpha x + \beta u) \right\},$$

where p is the co-state. Then the mean-field system associated to the robust mean-field game introduced in Problem 1 is given by

$$\begin{cases} \partial_t v_t + H(x, \partial_x v_t, m_t) + \left(\frac{\sigma}{2\gamma}\right)^2 (\partial_x v_t)^2 \\ \qquad + \frac{1}{2}\sigma^2 x^2 \partial^2_{xx} v_t = 0, \ \text{in } \mathbb{R} \times [0,T[, \\ \\ v_T(x) = g(x), \ \text{in } \mathbb{R}, \\ \\ m_0(x) = d(x) \text{ in } \mathbb{R}, \\ \\ \partial_t m_t + \partial_x \left(m_t \partial_p H(x, \partial_x v_t, m)\right) \\ + \frac{\sigma^2}{2\gamma^2} \partial_x (m_t \partial_x v_t) - \frac{1}{2}\sigma^2 \partial^2_{xx}\left[x^2 m_t\right] = 0, \\ \ \text{in } \mathbb{R} \times [0,T[, \end{cases} \tag{4}$$

where d is the initial population state distribution and g the terminal payoff. Any solution of the above system of equations is referred to as *worst-disturbance feedback mean-field equilibrium*. We are ready to specialize the results obtained above to the case of a crowd-averse system.

Theorem 1: The mean-field system associated to the robust mean-field game for the crowd-averse system is described by the equations:

$$\begin{cases} \partial_t v_t + \left[-\frac{\beta^2}{2(b+h\bar{z}_t)} + \left(\frac{\sigma}{2\gamma}\right)^2 \right] (\partial_x v_t)^2 \\ +\alpha x_t \partial_x v_t + \frac{a}{2} {x_t^*}^2 + \frac{1}{2}\sigma^2 x^2 \partial^2_{xx} v_t = 0, \\ \ \text{in } \mathbb{R} \times [0,T[, \\ \\ v_T(x) = \phi |x|^2, \ \text{in } \mathbb{R}, \\ \\ \partial_t m_t + \partial_x \Big[m_t \Big(\alpha x_t + \beta \frac{-\partial_x v_t \beta}{b + h\bar{z}_t} \\ + \frac{\sigma^2}{2\gamma^2} \partial_x v_t \Big) \Big] + \frac{\sigma^2}{2\gamma^2} \partial_x (m_t \partial_x v_t) \\ -\frac{1}{2}\sigma^2 \partial^2_{xx} \left[x^2 m_t\right] = 0, \ \text{in } \mathbb{R} \times [0,T[, \\ \\ m_0(x) = d(x) \text{ in } \mathbb{R}, \end{cases} \tag{5}$$

where $d(x)$ is a given function. Furthermore, the optimal control and worst disturbance are

$$\begin{cases} u_t^* = \frac{-\beta}{b + h\bar{z}_t} \partial_x v_t, \\ \\ \zeta_t^* = \frac{\sigma}{2\gamma^2} \partial_x v_t. \end{cases} \tag{6}$$

Proof: We first prove condition (6). To this end write the Hamiltonian as:

$$\begin{aligned} H(x_t, \partial_x v_t, m_t) = \inf_u \Big\{ \tfrac{h}{2} \bar{z}_t u_t^2 + \Big[\tfrac{a}{2} x^2 \\ + \tfrac{b}{2} u_t^2 \Big] + \partial_x v_t (\alpha x_t + \beta u) \Big\} = 0. \end{aligned} \tag{7}$$

Differentiating with respect to u gives

$$(b + h\bar{z}_t) u_t + \partial_x v_t \beta = 0, \tag{8}$$

which yields (6).

We now prove (5). First note that the second and last equations are the boundary conditions and derive straightforwardly from Bellman equations and the evolution of the state.

To prove the first equation, which is a PDE corresponding to the HJBI equation, replace u in the Hamiltonian (7) by its expression (6), i.e.

$$
\begin{aligned}
H(x_t, \partial_x v_t, m_t) &= \frac{a}{2} {x_t^*}^2 + \frac{b}{2} {u_t^*}^2 + \frac{h}{2} \bar{z}_t {u_t^*}^2 \\
&\quad + \partial_x v_t \alpha x_t + \partial_x v_t \beta u_t^* \\
&= -\frac{\beta^2}{2(b + h\bar{z}_t)} (\partial_x v_t)^2 + \alpha x_t \partial_x v_t + \frac{a}{2} {x_t^*}^2 .
\end{aligned}
$$

Using the above expression of the Hamiltonian in the HJBI equation in (4), we obtain the HJBI in (5).

To prove the third equation, which is a PDE representing the FPK equation, we simply bring (6) into the FPK equation in (4), and this concludes the proof. ∎

The significance of the above result is that to find the optimal control input we need to solve the two coupled PDEs in (5) in v and m with given boundary conditions (the second and last conditions). This is usually done by iteratively solving the HJBI equation for fixed m and by entering the optimal u obtained from (6) in the FPK equation in (5), until a fixed point in v and m is reached.

Note that since the Bellman equation depends explicitly on the mean of the mean-field and not on the other moments, one can reduce the mean-field system to a lower dimensional system. The reduced mean-field system associated to the robust mean-field game for the problem under study is

$$
\begin{cases}
\partial_t v_t + \left[-\frac{\beta^2}{2(b+h\bar{z}_t)} + \left(\frac{\sigma}{2\gamma}\right)^2 \right] (\partial_x v_t)^2 \\
+\alpha x_t \partial_x v_t + \frac{a}{2} {x_t^*}^2 + \frac{1}{2}\sigma^2 x^2 \partial^2_{xx} v_t = 0, \\
\text{ in } \mathbb{R} \times [0, T[, \\
v_T(x) = \phi |x|^2, \text{ in } \mathbb{R}, \\
\frac{d}{dt} \bar{m}_t = \alpha \bar{m}_t + \beta \bar{u}_t^* + \sigma \bar{\zeta}_t^*, \text{ in } [0, T[, \\
\bar{m}_0 = \bar{d} > 0,
\end{cases}
\tag{9}
$$

where $\bar{u}_t^* = \bar{z}_t$ is the mean of the optimal individual state feedback control.

V. Mean-field equilibrium and stability

In this section we study the problem in the extended state space involving both the state of the player and the average state distribution. The main idea is illustrated in Fig. 1. In the mean-field system (9) the gradient $\partial_x v_t$ is parametrized in the average distribution $\bar{m}_t$, which evolves according to a nonlinear differential equation. Then, we replace the dynamics of $\bar{m}_t$ with two linear dynamics on the new variables $\hat{m}_t$ and $\tilde{m}_t$ (dashed and dotted trajectories) that upper and lower bound the nonlinear dynamics of $\bar{m}_t$ (solid).

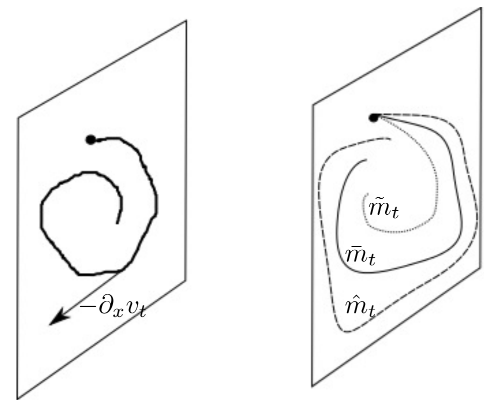

Fig. 1. Extended state space: the gradient $\partial_x v_t$ depends on $\bar{m}_t$, which is upper and lower bounded by $\hat{m}_t$ (dashed) and $\tilde{m}_t$ (dotted) respectively.

In the extended state space, the state variable evolves according to the equations

$$
\begin{cases}
dx_t = [\alpha x_t + \beta u_t] dt + \sigma \left[x_t d\mathcal{B}_t + \zeta_t dt \right], \\
\frac{d}{dt} \bar{m}_t = \alpha \bar{m}_t + \beta \bar{u}_t^* + \sigma \bar{\zeta}_t^*,
\end{cases}
\tag{10}
$$

which can be rewritten in matrix form as

$$
\begin{aligned}
\begin{bmatrix} dx_t \\ d\bar{m}_t \end{bmatrix} &= \Big(\alpha \begin{bmatrix} x_t \\ \bar{m}_t \end{bmatrix} + \beta \begin{bmatrix} u_t^* \\ \bar{u}_t^* \end{bmatrix} \\
&\quad + \sigma \begin{bmatrix} \zeta_t^* \\ \bar{\zeta}_t^* \end{bmatrix} \Big) dt + \begin{bmatrix} \sigma x_t d\mathcal{B}_t \\ 0 \end{bmatrix} .
\end{aligned}
\tag{11}
$$

For this system we introduce an assumption on the rate of convergence of the state $\bar{m}_t$.

Assumption 1: There exists θ such that

$$
\frac{d}{dt} \bar{m}_t = \alpha \bar{m}_t + \beta \bar{u}_t^* + \sigma \bar{\zeta}_t^* \geq -\theta \bar{m}_t, \text{ for all } t \in [0, T].
$$

The above assumption implies that there exists a variable $\tilde{m}_t$ which approximates the average distribution from below and that evolves according to

$$
\begin{cases}
\frac{d}{dt} \tilde{m}_t = -\theta \tilde{m}_t, \quad \text{for all } t \in [0, T], \\
\tilde{m}_0 = \bar{m}_0.
\end{cases}
\tag{12}
$$

By substituting the current average distribution $\bar{m}_t$ by its estimate $\tilde{m}_t$ the extended state dynamics takes the form

$$
\begin{aligned}
\begin{bmatrix} dx_t \\ d\tilde{m}_t \end{bmatrix} &= \Big(\begin{bmatrix} \alpha & 0 \\ 0 & -\theta \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{m}_t \end{bmatrix} \\
&\quad + \begin{bmatrix} \beta \\ 0 \end{bmatrix} u_t^* + \begin{bmatrix} \sigma \\ 0 \end{bmatrix} \zeta_t^* \Big) dt + \begin{bmatrix} \sigma x_t d\mathcal{B}_t \\ 0 \end{bmatrix} .
\end{aligned}
\tag{13}
$$

Given the above dynamics we summarize the problem at hand as

$$\begin{cases} \inf_{\{u_t\}_t} \sup_{\{\zeta_t\}_t} \int_0^T \Big[\frac{s}{2}\tilde{m}_t u_t^2 + \frac{q}{2}\tilde{m}_t^2 \\ \qquad\qquad + \Big(\frac{a}{2}x_t^2 + \frac{b}{2}u_t^2 - \gamma^2\zeta_t^2\Big)\Big]dt \\ \begin{bmatrix} dx_t \\ d\tilde{m}_t \end{bmatrix} = \Big(\begin{bmatrix} \alpha & 0 \\ 0 & -\theta \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{m}_t \end{bmatrix} + \begin{bmatrix} \beta \\ 0 \end{bmatrix} u_t^* \\ \qquad + \begin{bmatrix} \sigma \\ 0 \end{bmatrix} \zeta_t^* \Big) dt + \begin{bmatrix} \sigma x_t d\mathcal{B}_t \\ 0 \end{bmatrix}, \end{cases}$$

where $s = 2h\hat{k}$ by equation (3). Reformulating the problem in terms of the extended state

$$X_t = \begin{bmatrix} x_t \\ \tilde{m}_t \end{bmatrix},$$

yields the linear quadratic problem:

$$\begin{cases} \inf_{\{\tilde{u}_t\}_t} \sup_{\{\zeta_t\}_t} \int_0^T \Big[\frac{1}{2}(X_t^T \tilde{Q} X_t + R\tilde{u}_t^2 - \Gamma)\zeta_t^2\Big] dt \\ dX_t = \Big(\tilde{A}X_t + B\tilde{u}_t + C\zeta_t\Big)dt + Cx_t d\mathcal{B}_t, \end{cases}$$

where

$$\tilde{Q} = \begin{bmatrix} a & 0 \\ 0 & q \end{bmatrix}, \quad R = b + s\tilde{m}_t, \quad \Gamma = 2\gamma^2,$$

$$\tilde{A} = \begin{bmatrix} \alpha & 0 \\ 0 & -\theta \end{bmatrix}, \quad B = \begin{bmatrix} \beta \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} \sigma \\ 0 \end{bmatrix}.$$

The idea is therefore to consider a new value function $\mathcal{V}_t(x, \tilde{m})$ (in compact form $\mathcal{V}_t(X)$) in the extended state space which satisfies

$$\begin{cases} \partial_t \mathcal{V}_t(X) + H(X, \partial_X \mathcal{V}_t(X)) + \Big(\frac{\sigma}{2\gamma}\Big)^2 |\partial_x \mathcal{V}_t(X)|^2 \\ + \frac{1}{2}\sigma^2 x^2 \partial_{xx}^2 \mathcal{V}_t(X) = 0, \text{ in } \mathbb{R}^2 \times [0, T[, \\ \mathcal{V}_T(X) = g(x) \text{ in } \mathbb{R}^2. \end{cases}$$

Assume that $\mathcal{V}_T(X)$ is given by the quadratic form

$$\mathcal{V}_t(X) = [x_t \quad \bar{m}_t] \underbrace{\begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix}}_{P(t)} \begin{bmatrix} x_t \\ \bar{m}_t \end{bmatrix},$$

where the matrix $P(t)$ is the solution of the differential Riccati equation

$$\begin{array}{l} \dot{P}(t) + P(t)\tilde{A} + \tilde{A}^T P(t) \\ -2P(t)(BR^{-1}B^T - C\Gamma^{-1}C^T)P(t) \\ \quad +\tilde{Q}/2 + W = 0, \end{array} \tag{14}$$

where

$$BR^{-1}B^T - C\Gamma^{-1}C^T = \begin{bmatrix} \frac{\beta^2}{b+s\tilde{m}} - \frac{1}{2\gamma^2}\sigma^2 & 0 \\ 0 & 0 \end{bmatrix},$$

$$W = \begin{bmatrix} \sigma^2 P_{11} & 0 \\ 0 & 0 \end{bmatrix}.$$

Note that in the stationary case the above differential equation simplifies to

$$\begin{array}{l} P\tilde{A} + \tilde{A}^T P - 2P(BR^{-1}B^T - C\Gamma^{-1}C^T)P \\ +\tilde{Q}/2 + W = 0. \end{array} \tag{15}$$

Let P be the solution of the differential Riccati equation (14), then the optimal control is given by

$$\begin{array}{l} \tilde{u}_t = -2R^{-1}B^T P X_t \\ = -\frac{2}{b+s\bar{m}_t}[\beta \; 0] \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix} \begin{bmatrix} x_t \\ \bar{m}_t \end{bmatrix} \\ = -\frac{2}{b+s\bar{m}_t}\beta(P_{11}(t)x_t + P_{12}(t)\bar{m}_t), \end{array} \tag{16}$$

and the worst disturbance is

$$\begin{array}{l} \tilde{w}_t = 2\Gamma^{-1}C^T P X_t \\ = \frac{1}{\gamma^2}[\sigma \; 0] \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix} \begin{bmatrix} x_t \\ \bar{m}_t \end{bmatrix} \\ = \frac{1}{\gamma^2}\sigma(P_{11}(t)x_t + P_{12}(t)\bar{m}_t). \end{array} \tag{17}$$

We are then in the position to establish the following result, which provides a lower bound for the value function in (9) when $\sigma = 0$.

Theorem 2: Let $\sigma = 0$. Then $\mathcal{V}_t(X)$ approximates $v(x)$ from below, i.e.,

$$\mathcal{V}_t(X) \leq v_t(x), \quad \forall X, x, t. \tag{18}$$

Furthermore, the approximation error is upper bounded by

$$\begin{array}{l} v_t(x) - V_t(x) \\ \leq s\Big(\frac{2\beta(P_{11}+P_{12})}{b}\Big)^2 \bar{m}_0^3 \Big[e^{-3\kappa t} - e^{(-\theta-2\kappa)t}\Big]. \end{array} \tag{19}$$

Proof: The main idea is to approximate the mean distribution $\bar{m}_t$ from below by $\tilde{m}_t$ and from above by $\hat{m}_t$. In other words we wish the following condition to hold:

$$\tilde{m}_t \leq \bar{m}_t \leq \hat{m}_t, \quad \text{for all } t \in [0, T]. \tag{20}$$

The above is true if we consider the following dynamics:

$$\begin{cases} \frac{d}{dt}\bar{m}_t = \Big(\alpha - \frac{2\beta(P_{11}+P_{12})}{b+s\bar{m}_t}\Big)\bar{m}_t, \\ \frac{d}{dt}\tilde{m}_t = \Big(\alpha - \frac{2\beta(P_{11}+P_{12})}{b}\Big)\bar{m}_t := -\theta\tilde{m}_t, \\ \frac{d}{dt}\hat{m}_t = \Big(\alpha - \frac{2\beta(P_{11}+P_{12})}{b+s\hat{m}_0}\Big)\hat{m}_t := -\kappa\hat{m}_t, \\ \bar{m}_0 = \hat{m}_0 = \tilde{m}_0\,, \end{cases} \tag{21}$$

with

$$\begin{cases} \theta = -\alpha + \frac{2\beta(P_{11}+P_{12})}{b}, \\ \kappa = -\alpha + \frac{2\beta(P_{11}+P_{12})}{b+s\hat{m}_0}. \end{cases} \tag{22}$$

Then, for the approximation error we have

$$\begin{array}{rcl} e(t) & := & v_t(x) - \mathcal{V}_t(X) \\ & \leq & \int_0^T s\tilde{u}_\tau^2(\bar{m}_\tau - \tilde{m}_\tau)d\tau \\ & \leq & \int_0^T s\tilde{u}_\tau^2 \bar{m}_0(\frac{\bar{m}_\tau}{\bar{m}_0} - \frac{\tilde{m}_\tau}{\bar{m}_0})d\tau \\ & \leq & \int_0^T s\tilde{u}_\tau^2 \bar{m}_0(\frac{\hat{m}_\tau}{\bar{m}_0} - \frac{\tilde{m}_\tau}{\bar{m}_0})d\tau \end{array} \tag{23}$$

for any $\bar{m}_t$, $\tilde{m}_t$, and $\hat{m}_t$ satisfying (20). Now, from (21)-(22), the above inequalities can be rewritten as

$$e(t) \leq \int_0^t s\tilde{u}_\tau^2 \bar{m}_0 \Big[e^{-\kappa\tau} - e^{-\theta\tau}\Big] d\tau \qquad (24)$$

from which, after differentiating with respect to t and substituting $\tilde{u}_t$ by the expression in (16), we obtain

$$\begin{aligned} \dot{e}(t) \quad &\leq s\Big(\tfrac{2\beta(P_{11}+P_{12})}{b}\tilde{m}_t\Big)^2 \bar{m}_0 \Big[e^{-\kappa\tau} - e^{-\theta\tau}\Big] \\ &\leq s\Big(\tfrac{2\beta(P_{11}+P_{12})}{b}\Big)^2 \hat{m}_t^2 \bar{m}_0 \Big[e^{-\kappa\tau} - e^{-\theta\tau}\Big] \\ &\leq s\Big(\tfrac{2\beta(P_{11}+P_{12})}{b}\Big)^2 \Big(\tfrac{\hat{m}_t}{\bar{m}_0}\Big)^2 \bar{m}_0^3 \Big[e^{-\kappa\tau} - e^{-\theta\tau}\Big] \\ &\leq s\Big(\tfrac{2\beta(P_{11}+P_{12})}{b}\Big)^2 \bar{m}_0^3 \Big[e^{-3\kappa t} - e^{(-\theta-2\kappa)t}\Big] \end{aligned}$$

which proves the claim. ■

A. *Exponential asymptotic stability*

In this section we show that the stochastic differential equation describing the closed-loop system has an exponentially and asymptotically stable equilibrium. To see this from (16)-(17) rewrite the dynamics for x_t in (10) as

$$\begin{aligned} dx_t &= [\alpha x_t + \beta u_t^* + \sigma\zeta_t^*]\, dt + \sigma x_t d\mathcal{B}_t \\ &= \Big[\alpha x_t + (-\tfrac{2\beta^2}{b+s\bar{m}_t} + \tfrac{\sigma^2}{\gamma^2})(P_{11}(t)x_t + P_{12}(t)\bar{m}_t)\Big]\, dt \\ &\quad + \sigma x_t d\mathcal{B}_t \quad t \in (0,T],\ x_0 \in \mathbb{R}. \end{aligned}$$

and consider the following assumption.

Assumption 2: There exists $\kappa > 0$ such that

$$\begin{aligned} -\kappa x_t \quad \geq \quad & \Big[\alpha + (-\tfrac{2\beta^2}{b+s\bar{m}_t} + \tfrac{\sigma^2}{\gamma^2})P_{11}(t)\Big]x_t \\ & + \Big[(-\tfrac{2\beta^2}{b+s\bar{m}_t} + \tfrac{\sigma^2}{\gamma^2})P_{12}(t)\Big]\bar{m}_t. \end{aligned} \qquad (25)$$

With the above assumption we can perform the analysis within the framework of stochastic stability theory [11]. To this end consider the infinitesimal generator

$$\mathcal{L} = \frac{1}{2}\sigma^2 x_t^2 \frac{d^2}{dx_t^2} - \kappa x_t \frac{d}{dx_t}. \qquad (26)$$

Then

$$\begin{aligned} & \frac{1}{2}\mathbb{E}\left(dx_t^2 \frac{d^2}{dx_t^2}\right) + \mathbb{E}\left(dx_t \frac{d}{dx_t}\right) \\ &= \frac{1}{2}\left[\mathbb{E}\left(\kappa^2 x_t^2 dt^2\right) + \mathbb{E}\left(\sigma^2 x_t^2 d\mathcal{B}_t^2\right)\right. \\ &\qquad \left. + \mathbb{E}\left(-2\kappa x_t dt \sigma x_t d\mathcal{B}_t\right)\right] \frac{d^2}{dx_t^2} \\ &\quad + \left[\mathbb{E}\left(-\kappa x_t dt\right) + \mathbb{E}\left(\sigma x_t d\mathcal{B}_t\right)\right] \frac{d}{dx_t}. \end{aligned}$$

Now, recalling that for a Brownian motion $\mathbb{E}d\mathcal{B}_t = 0$ and $\mathbb{E}d\mathcal{B}_t^2 \to 0$, and ignoring the second-order terms (in dt^2) we obtain (26).

Consider the Lyapunov function $V(x) = x^2$, then the stochastic derivative of $V(x)$ is obtained by applying the infinitesimal generator to $V(x)$, which yields

$$\begin{aligned} \mathcal{L}V(x_t) \quad &= \quad \lim_{dt \to 0} \frac{\mathbb{E}V(x_{t+dt}) - V(x_t)}{dt} \\ &= \quad [\sigma^2 - 2\kappa]x_t^2. \end{aligned}$$

Proposition 5.1 ([11]): Let Assumption 2 hold. If $V(x) \geq 0$, $V(0) = 0$ and $\mathcal{L}V(x) \leq -\eta V(x)$ on $Q_\epsilon := \{x : V(x) \leq \epsilon\}$ for some $\eta > 0$ and for arbitrarily large ϵ, then the origin is asymptotically stable "with probability one", and

$$P_{x_0}\Big\{\sup_{T \leq t < +\infty} x_t^2 \geq \lambda\Big\} \leq \frac{V(x_0)e^{-\psi T}}{\lambda}$$

for some $\psi > 0$.

From the above theorem we have the following result, which establishes exponential stochastic stability of the mean-field equilibrium.

Corollary 5.1: Let Assumption 2 hold. If $[\sigma^2 - 2\kappa] < 0$ then $\lim_{t\to\infty} x_t = 0$ almost surely and

$$P_{x_0}\Big\{\sup_{T \leq t < +\infty} x_t^2 \geq \lambda\Big\} \leq \frac{V(x_0)e^{-\psi T}}{\lambda}$$

for some $\psi > 0$.

B. *Mean-field equilibrium*

Let Assumption 2 hold. We can approximate the mean-field equilibrium, which is captured by the evolution of $\bar{m}_t$ over the horizon $(0,T]$, as follows:

$$\tfrac{d}{dt}\bar{m}_t \quad \leq \quad -\kappa\bar{m}_t, \quad t \in (0,T],\ m_0 \in \mathbb{R} \times [0,T],$$

which yields the upper bound for $\bar{m}_t$:

$$\bar{m}_t \leq \bar{m}_0 e^{-\kappa t}, \quad t \in (0,T],\ x_0 \in \mathbb{R}.$$

Essentially, the inequality above describes converging linear dynamics which upper bound the time evolution of $\bar{m}_t$, for all $t \in (0,T]$. As a result

$$\begin{aligned} \tfrac{d}{dt}\bar{m}_t &\leq \Big[\alpha + (-\tfrac{2\beta^2}{b+s\bar{m}_t} + \tfrac{\sigma^2}{\gamma^2})(P_{11}(t) + P_{12})\Big]\bar{m}_t \\ & \quad t \in (0,T],\ x_0 \in \mathbb{R}. \end{aligned}$$

Actually, we can derive a differential equation describing the evolution of the mean distribution which represents a bound, namely

$$\begin{cases} \bar{m}_t \leq \bar{m}_0 e^{\rho t} \\ \rho = \alpha + (-\tfrac{2\beta^2}{b+s\bar{m}_t} + \tfrac{\sigma^2}{\gamma^2})(P_{11}(t) + P_{12}). \end{cases}$$

The equation above corresponds to saying that the mean distribution converges exponentially to zero in absence of the stochastic disturbances (the Brownian motion), under the assumption that ρ is strictly negative.

α	β	a	b	θ	q	γ	$\bar{m}_0$
0.1	2	2	1	10	0.1	1	80

TABLE I
SIMULATION PARAMETERS

Input: Set of parameters as in Table I.
Output: Distribution function m_t, mean $\bar{m}_t$ and standard deviation std(m_t).

Initialize. Generate x_0 given $\bar{m}_0$ and std(m_0)
for time $t = 0, 1, \ldots, T-1$ **do**
  if $t > 0$, **then** compute m_t, $\bar{m}_t$, and std(m_t)
  end if
  for player $i = 1, \ldots, n$ **do**
  Compute control $\tilde{u}$ using current $\bar{m}_t$
    compute new state x_{t+1} by executing (1)
  end for
end for
STOP

Fig. 2. Simulation algorithm

VI. Numerical studies

In this section a system consisting of $n = 10^3$ indistinguishable players, with dynamics (1), is considered. Suppose the players seek to minimise cost functionals of the form (2) subject to an adversary disturbance, *i.e.* consider Problem 1. The optimal control and worst-case disturbance are given by (6). However, as it is not possible to obtain closed-form solutions for the PDEs associated to the mean-field system, the approximate solutions given by the control (16) and the disturbance (17) are adopted, where the matrix

$$P = \left[\begin{array}{cc} P_{11}(\bar{m}) & 0 \\ 0 & \frac{q}{4\theta} \end{array} \right],$$

with $P_{11}(\bar{m}) = \sqrt{(\sigma^2 + 2\alpha)^2 + 8(\frac{\beta^2}{b+h\bar{m}} - \frac{\sigma^2}{2\gamma^2})} + \sigma^2 + 2\alpha$, is the positive definite solution to the algebraic Riccati equation (15). The numerical results are obtained using the algorithm in Figure VI for a discretised set of states. The parameter σ determines the influence of the Brownian motion, $\mathcal{B}_t$, and the disturbance ζ_t and simulations have been run for two different values of σ, namely $\sigma_0 = 0$, $\sigma_1 = 0.1$. The selection $\sigma = \sigma_0$ corresponds to the case in which there is no disturbance and the dynamics (1) is deterministic. The simulations have also been run for two different values of s, namely $s_1 = 0.5$ and $s_2 = 1.5$ Recall that large values of s correspond to large penalties when congestion occurs. The remainder of the parameters are as shown in Table I.

Figure 3 shows the time histories of the states of the players with the weights $s = s_1$ (top row) and $s = s_2$ (bottom row) and the paramters $\sigma = \sigma_0$ (left column) and $\sigma = \sigma_1$ (right column). Figure 4 shows the distribution, m_t, of the players states at different times for the four different selections of parameters. The initial and final distributions are indicated by the dashed and solid curves, respectively, whereas the distribution at intermediate times are denoted by the dotted curves. Figure 5 shows the time histories of the mean, $\bar{m}_t$, (left) and the standard deviation (right) for $s = s_1$ (top) and $s = s_2$ (bottom). The solid curves corrspond to $\sigma = \sigma_1$ whereas the dashed lines correspond to $\sigma = \sigma_0$.

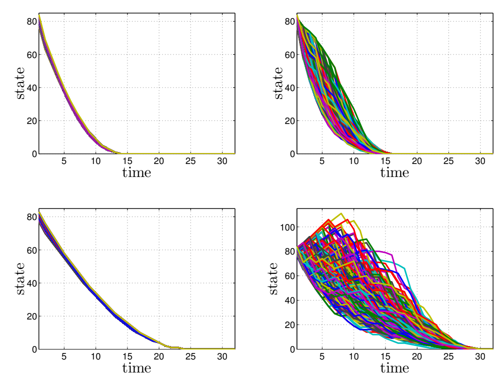

Fig. 3. Time histories of the state of each player. Top row: $s = s_1$, bottom row: $s = s_2$, left column: $\sigma = \sigma_0$, right column: $\sigma = \sigma_1$.

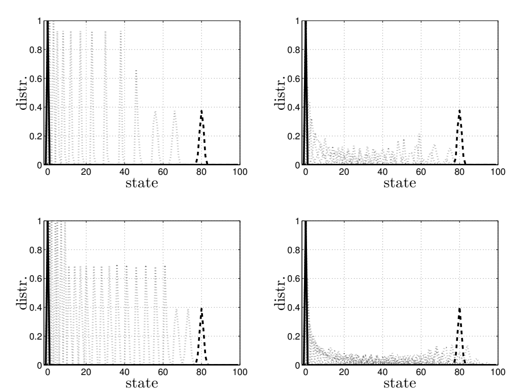

Fig. 4. The initial (dashed line), final (solid line) and intermediate (dotted lines) distribution, m_t, of the states of the players. Top row: $s = s_1$, bottom row: $s = s_2$, left column: $\sigma = \sigma_0$, right column: $\sigma = \sigma_1$.

Note that in all four cases the players successfully drive their states to zero. However, for a given value of the parameter s,the convergence fastest in the absence of noise and disturbances, *i.e.* when $\sigma = \sigma_0$. Figure

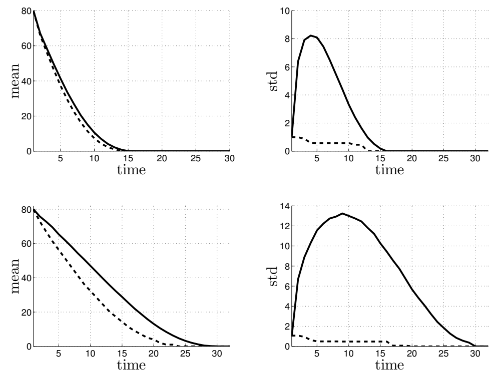

Fig. 5. Time histories of the mean $\bar{m}_t$ (left) and the standard deviation (right) of the states of the players for $s = s_1$ (left), $s = s_2$ (right), $\sigma = \sigma_0$ (dashed line) and $\sigma = \sigma_1$ (solid line).

6 shows the the time histories of the control actions (16) of the players with $s = s_1$ (top row) and $s = s_2$ (bottom row), and $\sigma = \sigma_0$ (left column) and $\sigma = \sigma_1$ (right column). For the case in which $\sigma = \sigma_1$, it is clear that when $s = s_1$ is selected the players put a larger effort at the beginning of the simulation than when $s = s_2$ is selected, and the same is true for $\sigma = \sigma_0$. Since $s_2 > s_1$, this implies that in the former case a larger penalty is incurred when congestion occurs and therefore one would expect the players to stall to avoid this, resulting in the convergence to the zero equilibrium being somewhat slower. The simulations are consistent with this, as for a given value of σ it takes more time for the players to drive their states to zero when the parameter $s = s_2$ is selected in place of $s = s_1$. The simulations show that the control actions (16) solve the robust-mean field problem for the crowd-averse system of players.

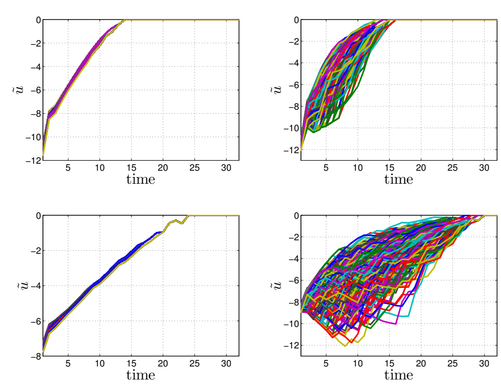

Fig. 6. Time histories of the control actions $\tilde{u}$ of the players. Top row: $s = s_1$, bottom row: $s = s_2$, left column: $\sigma = \sigma_0$, right column: $\sigma = \sigma_1$.

VII. Concluding remarks

We have illustrated robust mean-field games as a paradigm for crowd-averse systems. Future directions include i) the extension of the approximation method to more general cost functionals, ii) the study of the case with "local" mean-field interactions rather than "global" as in the current scenario, and iii) the analysis of crowd-seeking scenarios in contrast to the crowd-averse cases analyzed in this paper.

References

[1] Y. Achdou, F. Camilli, I. Capuzzo Dolcetta, Mean field games: numerical methods for the planning problem, SIAM J. of Control Opt., vol. 50, pp. 77–109, 2012.
[2] F. Bagagiolo and D. Bauso, Mean-field games and dynamic demand management in power grids, Dynamic Games and Applications, in print.
[3] M. Bardi, Explicit solutions of some Linear-Quadratic Mean Field Games, Network and Heterogeneous Media, 7, pp. 243–261, 2012.
[4] D. Bauso, H. Tembine, and T. Başar, Robust Mean Field Games with Application to Production of an Exhaustible Resource, Proc. of 7th IFAC Symposium on Robust Control Design, Aalborg, Danemark, 2012.
[5] D. Bauso, Q. Zhu and T. Başar, Mixed Integer Optimal Compensation: Decompositions and Mean-Field Approximations, Proc. of 2012 American Control Conference, Montreal, Montreal, Canada, 2012.
[6] O. Gueant, J. M. Lasry, and P. L. Lions, Mean-field games and applications, Paris-Princeton Lectures, Springer, pp. 1–66, 2010.
[7] M.Y. Huang, P.E. Caines, and R.P. Malhamé, Large Population Stochastic Dynamic Games: Closed Loop Kean-Vlasov Systems and the Nash Certainty Equivalence Principle, Communications in Information and Systems, vol. 6, no. 3, pp. 221–252 (2006).
[8] M.Y. Huang, P.E. Caines, and R.P. Malhamé, Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria, IEEE Trans. on Automatic Control, vol. 52. no. 9, pp. 1560–1571, 2007.
[9] A. Lachapelle, J. Salomon, and G. Turinici, Computation of Mean Field Equilibria in Economics, Math. Models Meth. Appl. Sci., vol. 20, pp. 1–22, 2010.
[10] J.-M. Lasry and P.-L. Lions, Mean field games, Japanese J. Math., vol. 2, pp. 229–260, 2007.
[11] K. A. Loparo and X. Feng, Stability of stochastic systems. The Control Handbook, CRC Press, pp. 1105-1126, 1996.
[12] Z. Ma, D. S. Callaway, and I. A. Hiskens, Decentralized Charging Control of Large Populations of Plug-in Electric Vehicles, IEEE Transactions on Control System Technology, vol. 21, no.1, pp. 67–78, 2013.
[13] R. Pesenti, and D. Bauso, Mean Field Linear Quadratic Games with Set Up Costs, Dynamic Games and Applications, vol. 3, no. 1, pp. 89–104, March 2013.
[14] M. Sassano, and A. Astolfi, Dynamic Approximate Solutions of the HJ Inequality and of the HJB Equation for Input-Affine Nonlinear Systems, IEEE Trans. on Automatic Control, vol. 57, no. 10, pp. 2490–2503, October 2012.
[15] M. Sassano, and A. Astolfi, Approximate finite-horizon optimal control without PDEs, Systems & Control Letters, vol. 62 pp. 97–103, 2013.
[16] H. Tembine, Q. Zhu, and T. Başar, Risk-sensitive mean-field stochastic differential games, Proc. of 2011 IFAC World Congress, Milan, Italy, August 29 - September 2, 2011.
[17] H. Tembine, J. Y. Le Boudec, R. ElAzouzi, E. Altman, Mean field asymptotic of Markov decision evolutionary games, IEEE Gamenets 2009.