A two-point boundary value formulation of a class of multi-population mean-field games

A. Cristofaro, T. Mylvaganam and D. Bauso

***Abstract*— We consider a multi-agent system consisting of several populations. The interaction between large populations of agents seeking to regulate their state on the basis of the distribution of the neighboring populations is studied. Examples of such interactions can typically be found in social networks and opinion dynamics, where heterogeneous agents or clusters are present and decisions are influenced by individual objectives as well as by global factors. In this paper, such a problem is posed as a multi-population mean-field game, for which solutions depend on two partial differential equations, namely the Hamilton-Jacobi-Bellman equation and the Fokker-Planck-Kolmogorov equation. The case in which the distributions of agents are sums of polynomials and the value functions are quadratic polynomials is considered. It is shown that for this class of problems, which can be considered as approximations of more general problems, a set of ordinary differential equations, with two-point boundary value conditions, can be solved in place of the more complicated partial differential equations characterizing the solution of the multi-population mean-field game.**

I. INTRODUCTION

Due to their wide variety of applications, multi-agent systems are ubiquitous and have, as a consequence, gained interest in recent years. Such systems arise in a wide range of contexts including robotics [1], [2], [3], power systems [4], [5] and optimization [6], [7]. Consequently, different approaches for control design for multi-agent systems are available in the literature. Several of these approaches exploit notions borrowed from game theory. For instance, in [8] game theory and cooperative control are applied to the problem of optimizing energy production in wind farms; in [9], [10], [11] control laws for monitoring a region using multiple unmanned vehicles equipped with sensors are developed through the formulation of a game; in [12], [13], [14], [15] collision avoidance and formation flying for multi-agent systems is achieved through the formulation of a nonzero-sum game. For game theoretic problems with a large number of players, the framework provided by mean-field games, first introduced in [16] and [17], may be of interest.Mean-field games arise in the context of applications with a large number of agents, such as power systems (see, for instance, [18]) and opinion dynamics.

A. Cristofaro is with the School of Science and Technology – Mathematics Division, University of Camerino, 62032 Camerino MC, Italy (Email: andrea.cristofaro@unicam.it).

T. Mylvaganam is with the Department of Aeronautics, Imperial College London, London SW7 2AZ, UK (Email:t.mylvaganam@imperial.ac.uk).

D. Bauso is with the Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield S1 3JD, U.K., and with the Dipartimento dellInnovazione Industriale e Digitale, Università di Palermo, 90128 Palermo, Italy (E-mail: d.bauso@sheffield.ac.uk).

In "classic" mean-field games it is assumed that all players (also referred to as agents) are *indistinguishable* [19]. More recently the theory has been extended to problems in which the agents are not identical. For example, mean-field games with minor and major agents have been considered in [20]. Drawing inspiration from the setting considered in [21], in this paper we consider a system consisting of several *populations* of agents, wherein each agent seeks to minimize (via its control input) a cost function which depends on neighboring populations. Agents belonging to different populations satisfy different (linear) dynamics and seek to optimize different cost functions. Similarly to what has been done in [22], [23] for a class of single-population mean-field games, we demonstrate that for the class of multi-population mean-field games considered herein, the corresponding control strategies rely on the solution of a system of ordinary differential equations forming a two-point boundary value problem.

The remainder of this paper is organised as follows. The problem formulation and terminology is given in Section II before a choice of cost function is introduced and discussed in Section III. The main result of this paper, namely the formulation of the solution of the multi-population mean-field game as a two-point boundary value problem, is provided in Section IV. A stability analysis of the system and numerical simulations are then presented in Sections V and VI, respectively. Finally, some concluding remarks are provided in Section VII.

Notation. We denote with $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space. 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 FORMULATION

We consider a system consisting of N (where N is finite) *populations*, each of which is associated with an index $i = 1, \ldots, N$. We assume that each population consists of infinitely many indistinguishable agents and that interactions between the different populations are limited. In particular the interactions are dictated by a time-invariant communication topology which is modeled by a static directed graph as illustrated for a four-population example in Figure 1. The graph, which we denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, consists of a set $\mathcal{V} = \{V_1, \ldots, V_N\}$ of vertices and a set $\mathcal{E}$ of edges connecting pairs of vertices. Each vertex V_i corresponds

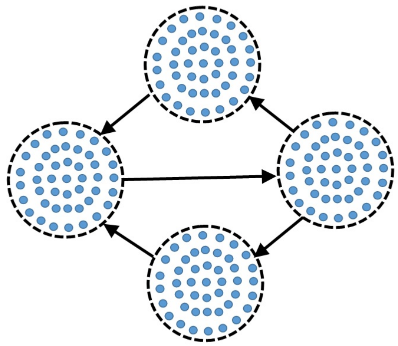

Fig. 1. A system consisting of four populations (denoted by dashed circular lines) of infinitely many agents (represented by solid circular markers). The arrows indicate interactions between populations.

to population i, $i = 1, \dots, N$, and the edges represent interactions between two populations. In particular if $(i, j) \in \mathcal{E}$ if there is a directed edge from population i to population j and population i is then said to be a neighbor of population j. Finally, $\mathcal{N}_i$ denotes the set of all neighbors of population i. Note that while, self loops are omitted in the graph (see Figure 1), each population is said to be a neighbor of itself, that is $(i, i) \in \mathcal{N}_i$, for $i = 1, \dots, N$.

Each agent belonging to population i satisfies the linear dynamics

$$dx_{i,t} = [\alpha_i x_{i,t} + \beta_i u_{i,t} + \sigma_i w_{i,t}]\, dt + \sigma_i x_{i,t} d\mathcal{B}_t\,, \tag{1}$$

where $x_i(t) \in \mathbb{R}$ is the state, $u_i(t) \in \mathbb{R}$ denotes the control input of agents belonging to population i, $w_i(t) \in \mathbb{R}$ denotes a disturbance acting on agents belonging to population i and $\alpha_i \in \mathbb{R}$, $\beta_i \in \mathbb{R}$ and $\sigma_i \in \mathbb{R}$ are scalar parameters, $i = 1, \dots, N$. Note that in (1) and in the remainder of this paper, the subscript t is used to indicate time dependence.

While the dynamics (1) model individual agents, each population is represented by a *probability distribution function*

$$m_i : \mathbb{R} \times [0, +\infty[\to \mathbb{R},\ (x, t) \mapsto m_{i,t}(x)\,,$$

which is such that $\int_{\mathbb{R}} m_{i,t}(x)dx = 1$ for every t, and for all $i = 1, \dots, N$. The individual states $x_{i,t}$, $i = 1, \dots, N$, provide a *microscopic model* of the system, whereas the probability distribution functions $m_{i,t}$, $i = 1, \dots, N$, provide a *macroscopic model* of the system. The distribution $m_{i,t}$ is sometimes referred to as the *mean-field* corresponding to population i.

In addition to the possibly differing dynamics (1), agents belonging to different populations differ in that they strive towards achieving different objectives. The objective of population i, for $i = 1, \dots, N$, is described by the cost function

$$\begin{aligned} J_i(x_i, u_i, m_{\mathcal{N}_i}) = \mathbb{E}\Big(&\Psi_i(x_{i,T}, m_{\mathcal{N}_i,T}, T) \\ &+ \int_0^T \Big[q_i(x_i, m_{\mathcal{N}_i,t}, t) + \frac{r_i}{2}u_{i,t}^2 - \frac{\gamma_i}{2}w_{i,t}^2\Big]\, dt\Big)\,, \end{aligned} \tag{2}$$

where $T > 0$ is the (finite) time horizon,

$$m_{\mathcal{N}_i,t} = \{m_{j,t} | j \in \mathcal{N}_i\}$$

denotes the set of probability density distributions of all neigbours of population i, $q_i(x_{i,t}, m_{\mathcal{N}_i,t}, t)$ is a running cost, $\Psi(x_{i,T}, m_{\mathcal{N}_i,T}, T)$ is a terminal cost, $r_i > 0$ and $\gamma_i > 0$. While each agent belonging to population i seeks to minimise the cost functions (2) via the control input $u_{i,t}$, it is assumed that the external disturbance influencing the dynamics of the agent seeks to maximise (2) via the $w_{i,t}$.

The problem described above, which we will refer to as a *robust multi-population mean-field game* is defined as follows.

Problem 1: (**Robust multi-population mean-field game**) Let $\mathcal{B}$ be a one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $x_{i,0}$ be independent of $\mathcal{B}$ and with density $m_{i,0}(x)$, for $i = 1, \dots, N$. The robust multi-population mean-field game lies in determining $u_{i,t}$ and $w_{i,t}$ solving the minimax problems

$$\inf_{u_i} \sup_{w_i}\ J_i(x_i, u_i, m^*_{\mathcal{N}_i}, w_i)\,, \tag{3}$$

for $i = 1, \dots, N$, subject to the dynamics (1), where $m^*_{\mathcal{N}_i,t}$ denotes the optimal mean field trajectories of the populations in the neighborhood of population i.
In the setting considered herein, a multi-population version of classic mean-field game theory, similarly to that considered in [21], provides convenient tools for designing the control laws $u_{i,t}$.

In what follows, we assume that the probability distribution of each population is described by a polynomial function.

Assumption 1: The probability distribution of each of the populations i, $i = 1, \dots, N$, is described by

$$m_{i,t}(x) = a_{i,0t} + \sum_{j=1}^{n} \frac{1}{j} a_{i,jt} x^j\,, \tag{4}$$

where $n > 0$, $a_{i,jt} \in \mathbb{R}$, for $j = 1, \dots, n$, and where $m_i(x)$ is of compact support. In particular the support is such that $\int_{\mathbb{R}} m_{i,t}(x)dx = 1$ for every t.

Remark 1: The probability distribution (4) can be interpreted as the n-th order Taylor series approximation of a more general probability distribution function.

We consider the case in which the running costs q_i, $i = 1, \dots, N$, are quadratic functions of the state. Namely,

$$q_i(x_i, m_{\mathcal{N}_i}) = c_{i,0t} + \sum_{j=1}^{2} \frac{1}{j} c_{i,jt} x_{i,t}^j\,, \tag{5}$$

where $c_{i,jt}$, for $j = 0, \dots, 2$, are scalar coefficients, which are functions of the probability distributions $m_{\mathcal{N}_i}$, for $i = 1, \dots, N$.[1] Similarly, we consider the case in which the terminal costs Ψ_i, $i = 1, \dots, N$, are quadratic functions of the state, *i.e.*

$$\Psi_i(x_{i,T}, m_{\mathcal{N}_i,T}, T) = \psi_{i,0} + \sum_{j=1}^{2} \frac{1}{j} \psi_{i,j} x_{i,T}^j\,, \tag{6}$$

[1]Note that while the running cost associated with a population i is quadratic in the state x_i, it can also be a function of the distribution of any population $j \in \mathcal{N}_i$, through the selection of the coefficients $c_{i,jt}$, $j = 0, \dots, 2$, as discussed in Section III.

where $\psi_{i,j}$, for $j = 0, 1, 2$ are scalar coefficients which are functions of the probability distributions $m_{\mathcal{N}_i}$, for $i = 1, \ldots, N$.

III. Cost function design

Each agent belonging to a population i seeks to minimise the cost function (2). The behaviour of each population will then depend on the selection of the weights r_i and γ_i as well as the running cost q_i, for $i = 1, \ldots, N$. The behaviour is, in part, dictated by the design of the coefficients $c_{i,jt}$ and $\psi_{i,jt}$, for $i = 1, \ldots, N$ and $j = 0, 1, 2$, which appear in the running cost (5). In this section, a possible design of these coefficients is discussed in some more details.

Consider the quadratic approximation of the density distributions of each population given by

$$m^q_{i,t}(x) = a_{i,0t} + \sum_{j=1}^{2} \frac{1}{j} a_{i,jt} x^j \,,$$

for $i = 1, \ldots, N$. Then the coefficients

$$c_{i,jt} = \sum_{k \in \mathcal{N}_i} a_{k,jt} \,, \tag{7}$$

and

$$\psi_{i,j} = \sum_{k \in \mathcal{N}_i} a_{k,jT} \,, \tag{8}$$

for $j = 0, \ldots, 2$ and for $i = 1, \ldots, N$ yield running costs and terminal costs which incentivize agents to move towards regions in which the population of neighboring agents is low. Thus, the cost function is such that agents are incentivized to be *crowd-averse*.

Note that the running cost with coefficients given by (7) is such that $q_i = \sum_{k \in \mathcal{N}_i} m^q_{k,t}(x_{i,t})$ and the terminal cost with coefficients given by (8) is such that $\Psi_i = \sum_{k \in \mathcal{N}_i} m^q_{k,T}(x_{i,T})$.

Remark 2: Different choices for the running cost coefficients are feasible and potentially interesting. For instance, one might consider the quadratic approximation of the squared difference between the distributions of two neighboring populations, this attempting to model a *similarity seeking* behavior among the populations.

IV. Main results/Two-point boundary value formulation

Let us suppose that $r_i \neq 0$, $\gamma_i \neq 0$, and define the robust Hamiltonian for the i-th population, i.e., the Hamiltonian of the robust mean field game (3) as

$$\begin{aligned}\tilde{H}_i(x_{i,t}, p_t, m_{\mathcal{N}_i,t}) = &\inf_{u_i} \sup_{w_i} \Big\{ q_i(x_{i,t}, m_{\mathcal{N}_i,t}, t) + \frac{r_i u_i^2 - \gamma_i w_i^2}{2} \\ &+ p_t(\alpha_i x_{i,t} + \beta_i u_i + \sigma_i w_i) \Big\}.\end{aligned}$$

Let us begin by computing the supremum part. The function $w_i \longmapsto -\frac{\gamma_i^2 w_i^2}{2} + p_t \sigma_i w_i$ is strictly concave and has a global maximizer given by

$$w^*_{i,t} = \frac{\sigma_i}{\gamma_i} p_t. \tag{9}$$

Similarly, the function $u_i \longmapsto \frac{r_i^2 u_i^2}{2} + p_t \beta_i u_i$ is strictly convex and its global minimum is attained by

$$u^*_{i,t} = -\frac{\beta_i}{r_i} p_t. \tag{10}$$

We introduce $v_{i,t}(x_i)$ as the (upper) value of the robust optimization problem associated with the i-th population under the worst-case disturbance starting from time t at state x_i. Thus, the generic expressions of the worst-case disturbance and optimal control input are given by

$$\begin{cases} w^*_{i,t} = \frac{\sigma_i}{\gamma_i} \partial_{x_i} v_{i,t}, \\ u^*_{i,t} = -\frac{\beta_i}{r_i} \partial_{x_i} v_{i,t}, \end{cases} \tag{11}$$

where $v_{i,t}$ satisfies the Hamilton-Jacobi-Bellman equation

$$\begin{aligned}\partial_t v_{i,t}(x_i) &+ \tilde{H}_i(x_i, \partial_{x_i} v_{i,t}, m_{\mathcal{N}_i,t}) \\ &+ \frac{\sigma^2 x_i^2}{2} \partial^2_{x_i x_i} v_{i,t}(x_i) = 0,\end{aligned} \tag{12}$$

$$v_T(x_i) = \Psi_i(x_{i,T}, m_{\mathcal{N}_i,T}, T), \tag{13}$$

which is coupled with the Fokker-Planck-Kolmogorov equation for the distribution $m_{i,t}$

$$\begin{aligned}\partial_t m_{i,t}(x_i) + \partial_{x_i} \Big(m_{i,t}(x_i) \partial_p \tilde{H}_i(x_i, \partial_{x_i} v_{i,t}, m_{\mathcal{N}_i,t}) \Big) \\ -\frac{1}{2} \sigma^2 \partial^2_{x_i x_i} \left(x_i^2 m_{i,t} \right) = 0,\end{aligned} \tag{14}$$

$$m_{i,0}(x) = a_{i,00} + \sum_{j=1}^{N} \frac{1}{j} a_{i,j0} x^j. \tag{15}$$

Note that the solution to the first PDE is the value function, which contains the population distributions as scheduling parameters. Conversely, the second PDE is defined on variable population distributions and is parametrized in the value functions. The existence of solutions for problem (12)-(15) are guaranteed whenever the following conditions are fulfilled. Let the initial distribution $m_{i,0}$ be absolutely continuous, with a continuous density function having a finite second moment. As the integrand of the cost is convex in the input u_i and concave in the disturbance w_i, one gets a family of convex-concave stage cost functions. The drift dynamics in (1) is linear, and hence Lipschitz continuous because the coefficients $\alpha_i, \beta_i, \sigma_i$ are bounded. We assume that the Fenchel transform of the running cost q_i is Lipschitz in its arguments. Finally, we assume that the function $p \longmapsto \frac{\sigma_i^2}{\gamma_i^2} \|p\|^2 + \tilde{H}_i$ is strictly convex, differentiable and Lipschitz continuous. Note that this last condition is weaker than the typical convexity assumption on the Hamiltonian. Under the above main assumptions, the existence of a solution is established in [16, Theorem 2.6]. Any solution of the above system of equations is referred to as *worst-disturbance* *feedback multi-population mean-field equilibrium.*

Recalling that the running cost (5) is defined by a quadratic polynomial function, by similarity arguments, in the following we focus on the search of quadratic value functions of the form

$$\begin{cases} v_{i,t}(x_{i,t}) &= q_{i,0t} + \sum_{j=1}^{2} \frac{1}{j} q_{i,jt} x_{i,t}^{j}, \text{ in } \mathbb{R} \times [0,T] \\ v_{i,T}(x_{i,t}) &= g(x_{i,T}, m_{i,T}(x_{i,T})) \\ &= q_{i,0T} + \sum_{j=1}^{2} \frac{1}{j} q_{i,jT} x_{i,t}^{j} \\ &= a_{i,0T} + \sum_{j=1}^{2} \frac{1}{j} a_{i,jT} x_{i,t}^{j} . \end{cases} \tag{16}$$

Bearing this in mind, the mean-field system (12)-(15) associated to the robust mean-field game for the i-th population (3) is recast into the system of equations:

$$\begin{cases} \partial_t v_{i,t} + \left[-\frac{\beta_i^2}{2r_i} + \frac{\sigma_i^2}{2\gamma_i} \right] (\partial_{x_i} v_{i,t})^2 \\ +\alpha_i x_{i,t} \partial_{x_i} v_{i,t} + c_{i,0t} + c_{i,1t} x_i, t + \frac{1}{2} c_{i,2t} x_i, t^2 \\ +\frac{1}{2} \sigma_i^2 x^2 \partial^2_{x_i x_i} v_{i,t} = 0, \text{ in } \mathbb{R} \times [0,T[, \\ v_{i,T}(x_{i,t}) = q_{i,0T} + \sum_{j=1}^{2} \frac{1}{j} q_{i,jT} x_{i,t}^{j}, \text{ in } \mathbb{R}, \\ \partial_t m_{i,t} + \sum_{j=1}^{n} a_{i,jt} \Big[(1 + \frac{1}{j}) \\ \qquad \Big(\alpha_i - \frac{\beta_i^2}{r_i} q_{i,2t} + \frac{\sigma_i^2}{2\gamma_i^2} q_{i,2t} \Big) x_{i,t}^{j} \\ + \Big(-\frac{\beta_i^2}{r_i} q_{i,1t} + \frac{\sigma_i^2}{2\gamma_i^2} q_{i,1t} \Big) x_{i,t}^{j-1} \Big] \\ + a_{i,0t} \Big(\alpha_i - \frac{\beta_i^2}{r_i} q_{i,2t} + \frac{\sigma_i^2}{2\gamma_i^2} q_{i,2t} \Big) \\ -\frac{1}{2} \sigma^2 \partial^2_{x_i x_i} \left(x_{i,t}^2 m_{i,t} \right) = 0, \text{ in } \mathbb{R} \times [0,T[, \\ m_{i,0}(x) = a_{i,00} + \sum_{j=1}^{n} \frac{1}{j} a_{i,j0} x^{j} \text{ in } \mathbb{R}. \end{cases} \tag{17}$$

It is worth noticing that interactions between different populations are expressed in the above system through the coefficients $c_{i,jt}$, which interconnect the distributions m_k as k varies in the neighborhood $\mathcal{N}_i$. As a matter of fact, whenever $c_{i,jt},\ j = 0,1,2$, only depend on m_i, each population acts independently of the others, i.e. no interaction is enforced. The problem then reduces to two independent single-population mean-field games similar to those considered in [23].

A common way to tackle this type of problems is iteratively solving the Hamilton-Jacobi-Bellman equation for fixed m_i and by entering the optimal u_i obtained from (11) in the Fokker-Planck-Kolmogorov equation in (17), until a fixed point in v_i and m_i is reached. However this usually requires considerable computational effort and the solution cannot be obtained in closed-form. On the other hand, the following theorem establishes that the mean-field system (17) can be equivalently described by a two-point boundary value problem, which is more readily solved.

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

$$\begin{cases} \dot{q}_{i,0t} + \left[-\frac{\beta_i^2}{2r_i} + \frac{\sigma_i^2}{2\gamma_i} \right] q_{i,1t}^2 + c_{i,0t} = 0\,, \\ \dot{q}_{i,1t} + \left[-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i} \right] q_{i,1t} q_{i,2t} + \alpha_i q_{i,1t} + c_{i,1t} = 0\,, \\ \frac{1}{2} \dot{q}_{i,2t} + \left[-\frac{\beta_i^2}{2r_i} + \frac{\sigma_i^2}{2\gamma_i} \right] q_{i,2t}^2 + \left[\alpha_i + \frac{\sigma_i^2}{2} \right] q_{i,2t} \\ +\frac{1}{2} c_{i,2t} = 0\,, \\ q_{i,jT} = c_{i,jT}\,, \\ \dot{a}_{i,0t} + \Big[\alpha_i + \Big(-\frac{\beta_i^2}{b} + \frac{\sigma_i^2}{2\gamma_i^2} \Big) q_{2t} \Big] a_{i,0t} \\ + \left[-\frac{\beta_i^2}{2r_i} + \frac{\sigma_i^2}{2\gamma_i} \right] q_{i,1t} a_{i,1t} - \sigma_i^2 a_{i,0t} = 0, \\ \dot{a}_{i,1t} + 2 \Big[\alpha_i + \Big(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i} \Big) q_{i,2t} \Big] a_{i,1t} \\ + \left[-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i} \right] q_{i,1t} a_{i,2t} - \frac{6}{2} \sigma_i^2 a_{i,1t} = 0, \\ \frac{1}{j} \dot{a}_{i,jt} + \Big(1 + \frac{1}{j}\Big) \Big[\alpha + \Big(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i} \Big) q_{i,2t} \Big] a_{i,jt} \\ + \left[-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i} \right] q_{i,1t} a_{i,j+1t} \\ -\frac{(j+1)(j+2)}{2j} \sigma_i^2 a_{i,jt} = 0, \qquad j = 2, \ldots, n-1 \\ \frac{1}{n} \dot{a}_{i,nt} + \left(1 + \frac{1}{n}\right) \Big[\alpha_i + \Big(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i} \Big) q_{i,2t} \Big] a_{i,nt} \\ -\frac{(n+2)(n+1)}{2n} \sigma^2 a_{i,nt} = 0 \\ a_{i,j0} \text{ given for all } j = 1, \ldots, n. \end{cases} , \tag{18}$$

for $i = 1, \ldots, N$. The optimal control and worst disturbance are then given by

$$\begin{cases} \tilde{u}_{i,t} = -\frac{\beta_i}{r_i} (q_{i,2t} x_{i,t} + q_{i,1t}), \\ \tilde{w}_{i,t} = \frac{\sigma_i}{\gamma_i} (q_{i,2t} x_{i,t} + q_{i,1t})\,, \end{cases} \tag{19}$$

for $i = 1, \ldots, N$.

Remark 3: Note that the initial conditions $a_{i,j0}$ are given for all j, whereas the final conditions $q_{i,jT} = c_{i,jT}$ for $j = 0,1,2$ are unknown a-priori. Thus (18) is a somewhat atypical two-point boundary value problem, since one of the boundary conditions is unknown a-priori. Similarly to what has been done in [23] the problem is transformed into a standard two-point boundary value problem by performing the change of coordinates $\tilde{q}_{i,jt} = q_{i,jt} - c_{i,jt}$ for $j = 0,1,2$. The final condition is then given by $\tilde{q}_{i,jT} = 0$ and the modified problem is readily solved using numerical methods, such as the shooting method.

V. Stability results

In this section it is shown that the stochastic differential equation describing the closed-loop system has an exponentially and asymptotically stable equilibrium. Substituting the optimal control and the worst-case disturbance (19) into the dynamics for $x_{i,t}$ yields the closed-loop system

$$\begin{aligned} dx_{i,t} &= \left[\alpha_i x_{i,t} + \beta_i u^*_{i,t} + \sigma_i w^*_{i,t}\right] dt + \sigma x_{i,t} d\mathcal{B}_t \\ &= \left[\alpha_i + \left(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i}\right) q_{i,2t}\right] x_{i,t} dt \\ &\quad + \left(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i}\right) q_{i,1t} dt + \sigma_i x_{i,t} d\mathcal{B}_t, \\ &\qquad\qquad t \in (0,T],\ x_0 \in \mathbb{R}\,. \end{aligned}$$

Consider now the following assumption.

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

$$\begin{aligned} -\kappa_i x_{i,t} \quad \geq \quad & \left[\alpha_i + \left(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i}\right) q_{i,2t}\right] x_{i,t} \\ & + \left(-\frac{\beta_i^2}{r_i} + \frac{\sigma_i^2}{\gamma_i}\right) q_{i,1t}\,, \end{aligned} \tag{20}$$

for $i = 1, \ldots, N$.

With Assumption 2 the stability analysis can be performed within the framework of stochastic stability theory [24]. Consider the infinitesimal generator

$$\mathcal{L} = \frac{1}{2}\sigma^2 x_t^2 \frac{d^2}{dx_t^2} - \kappa_i x_t \frac{d}{dx_t}\,, \tag{21}$$

and the Lyapunov function $V(x) = x^2$. The stochastic derivative of $V(x)$ is obtained by applying the infinitesimal generator to $V(x)$. This yields

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

Proposition 1 ([24]): Suppose Assumption 2 holds. 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 1: Let Assumption 2 hold. If $[\sigma_i^2 - 2\kappa_i] < 0$ then there exists $\bar{x}_i$ such that $\lim_{t \to \infty} x_{i,t} = \bar{x}_i$ almost surely and

$$P_{x_0}\Big\{ \sup_{T \leq t < +\infty} (x_{i,t} - \bar{x}_i)^2 \geq \lambda \Big\} \leq \frac{V(x_0 - \bar{x}_i) e^{-\psi T}}{\lambda},$$

for some $\psi > 0$ and for $i = 1, \ldots, N$.

VI. Numerical studies

A numerical example illustrating the theory is presented in this section. We consider two distinct populations, namely a *red* population and a *blue* population, consisting of 3100 and 5400 indistinguishable agents, respectively. We use the subscripts 1 and 2 to denote the red and blue populations, respectively. The parameters of the system and the weights have been selected according to the values reported in Table I. For the sake of simplicity, uniform weights have been considered. The red population is initially distributed along the interval $[\frac{1}{2}, \frac{3}{2}]$, while the blue population along the wider interval $[-\frac{5}{2}, -\frac{1}{2}]$. The initial distributions are indicated by the dashed lines in Figures 3 and 4 and are both second degree polynomials, *i.e.* $n = 2$. A bidirectional communication among the agents is assumed to underlie the dynamics of the system, namely $m_{\mathcal{N}_i} = \{m_{1,t}, m_{2,t}\}$, for $i = 1, 2$. Note that this corresponds to a connected communication graph.

	α_i	β_i	σ_i	r_i	γ_i
red (population 1)	-0.01	0.5	0.05	1	0.4
blue (population 2)	-0.01	0.5	0.05	1	0.4

TABLE I
SIMULATION PARAMETERS

We examine the scenario in which both populations evaluate their best strategy according to the cost (2) with the coefficients $\{c_{1,jt}\}$ and $\{c_{2,jt}\}$ in (5) are given by (7).

The trajectories of the agents belonging to each population are shown in Figure 2. The probability density distributions $m_{1,t}$ (red) and $m_{2,t}$ (blue) at three different time instances are shown in Figures 3 and 4: the initial and final distributions are denoted by the dashed and solid lines, respectively, whereas the distribution at an intermediate time is denoted by the dotted lines. Note that the agents trajectories and distributions indicate a *crowd-averse* behavior as expected. Finally, the solution of the two-point boundary value problem (18) is shown in Figure 5 for completeness.

VII. Conclusion

In this paper we consider a multi-agent system consisting of several *populations* of agent. Agents belonging to a particular population seek to regulate their states on the basis of their state *and* the distributions of neighboring populations. The problem is formulated as a multi-population mean-field game and it is demonstrated that its solution requires solving

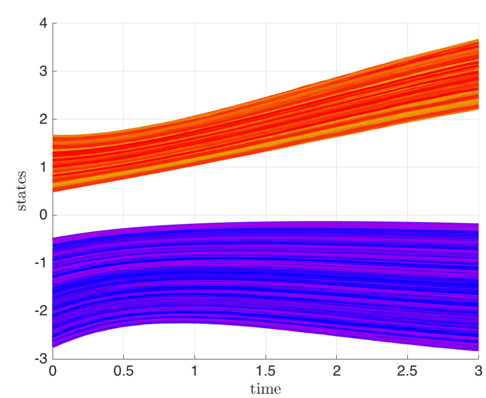

Fig. 2. Time histories of the state of players of red population and blue population, in the case of uniform cost function (7)

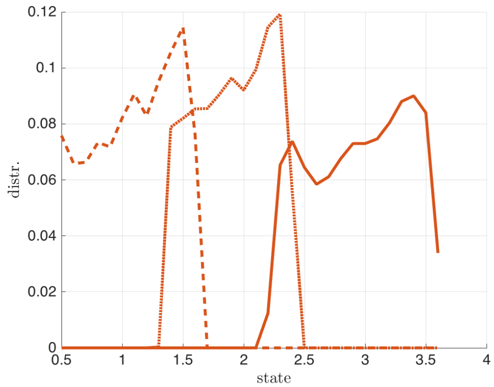

Fig. 3. Distribution of the red population for initial condition (dashed), intermediate condition (dotted) and final condition (solid)

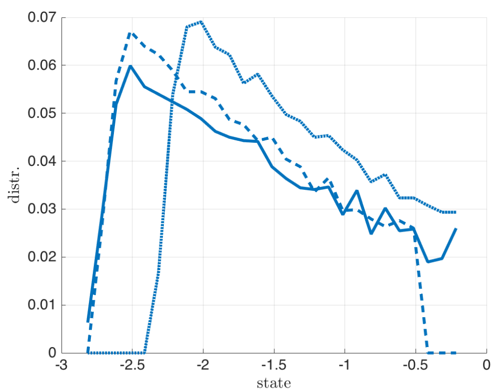

Fig. 4. Distribution of the blue population for initial condition (dashed), intermediate condition (dotted) and final condition (solid)

a two-point boundary value problem in place of the partial differential equations, namely the Hamilton-Bellman-Jacobi equation and the Fokker-Planck-Kolmogorov equation, that typically characterise the solution of a mean-field game. The result is demonstrated on a numerical example involving two populations.

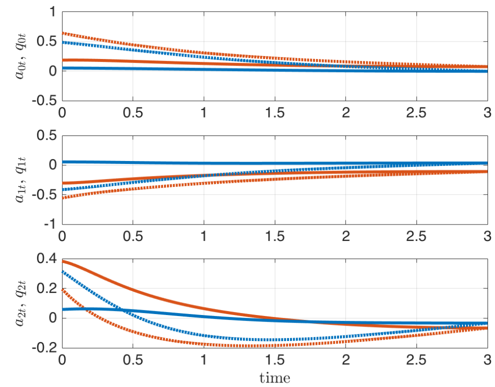

Fig. 5. Solution of the two-point boundary value problem (18). Dotted curves refer to the coefficients of the value function, while solid curves represent the coefficients of the distribution

References

[1] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *IEEE transactions on robotics and automation*, vol. 17, no. 6, pp. 947–951, 2001.

[2] P. Ogren, M. Egerstedt, and X. Hu, "A control lyapunov function approach to multi-agent coordination," in *Proceedings of the 40th IEEE Conference on Decision and Control*, vol. 2, 2001, pp. 1150–1155.

[3] A. Cristofaro, A. Renzaglia, and A. Martinelli, "Distributed information filters for mav cooperative localization," in *Distributed Autonomous Robotic Systems*, 2013, pp. 133–146.

[4] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings, "Theoretical and practical foundations of large-scale agent-based micro-storage in the smart grid," *Journal of Artificial Intelligence Research*, vol. 42, pp. 765–813, 2011.

[5] T. Mylvaganam and A. Astolfi, "Control of microgrids using a differential game theoretic framework," in *54th IEEE Conference on Decision and Control*, 2015.

[6] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[7] R. L. Raffard, C. J. Tomlin, and S. P. Boyd, "Distributed optimization for cooperative agents: Application to formation flight," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 3. IEEE, 2004, pp. 2453–2459.

[8] J. R. Marden, S. D. Ruben, and L. Y. Pao, "A model-free approach to wind farm control using game theoretic methods," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 4, pp. 1207–1214, 2013.

[9] T. Mylvaganam and A. Astolfi, "Approximate optimal monitoring: preliminary results," in *American Control Conference*, 2012.

[10] ——, "Approximate optimal monitoring," in *14th European Control Conference*, 2014.

[11] P. B. Sujit and D. Ghose, "Multiple agent search of an unknown environment using game theoretical models," in *Proceedings of the 2004 American Control Conference*, vol. 6, 2004, pp. 5564–5569.

[12] T. Mylvaganam, M. Sassano, and A. Astolfi, "A differential game approach to multi-agent collision avoidance for wheeled mobile robots," *IEEE Transactions on Automatic Control (under review)*, 2015.

[13] T. Mylvaganam and M. Sassano, "Autonomous collision avoidance for wheeled mobile robots using a differential game approach," *European Journal of Control*, 2017.

[14] D. Gu, "A differential game approach to formation control," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 1, pp. 85–93, 2008.

[15] T. Mylvaganam and A. Astolfi, "A differential game approach to formation control for a team of agents with one leader," in *American Control Conference*, 2015.

[16] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[17] M. Huang, R. P. Malhamé, and P. E. Caines, "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[18] A. De Paola, D. Angeli, and G. Strbac, "Distributed control of micro-storage devices with mean field games," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1119–1127, 2016.

[19] O. Guéant, J.-M. Lasry, and P.-L. Lions, "Mean field games and applications," *Paris-Princeton lectures on mathematical finance 2010*, pp. 205–266, 2011.

[20] M. Nourian and P. E. Caines, "-nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents," *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 3302–3331, 2013.

[21] D. Bauso, "Consensus via multi-population robust mean-field games," *Systems & Control Letters*, vol. 107, pp. 76–83, 2017.

[22] D. Bauso, T. Mylvaganam, and A. Astolfi, "A two-point boundary value formulation of a mean-field crowd-averse game," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7819–7824, 2014.

[23] T. Mylvaganam, D. Bauso, and A. Astolfi, "Mean-field games and two-point boundary value problems," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, 2014, pp. 2722–2727.

[24] K. Loparo and X. Feng, "Stability of stochastic systems," *The Control Handbook*, pp. 1105–1126, 1996.