The University Of Sheffield.

This is a repository copy of *Mean Field Linear Quadratic Games with Set Up Costs*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/89671/

Version: Accepted Version

**Article:**
Bauso, D. and Pesenti, R. (2013) Mean Field Linear Quadratic Games with Set Up Costs. Dynamic Games and Applications, 3 (1). 89 - 104. ISSN 2153-0785

https://doi.org/10.1007/s13235-012-0069-0

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

**Noname manuscript No.**
(will be inserted by the editor)

# Mean field linear quadratic games with set up costs

**Dario Bauso · Raffaele Pesenti**

Received: date / Accepted: date

**Abstract** This paper studies linear quadratic games with set up costs monotonic on the number of active players, namely, players whose action is non-null. Such games arise naturally in joint replenishment inventory systems. Building upon a preliminary analysis of the properties of the best response strategies and Nash equilibria for the given game, the main contribution is the study of the same game under large population. We also analyze the influence of an additional disturbance in the spirit of the literature on $H_\infty$ control. Numerical illustrations are provided.

**Keywords** Mean field games · Linear quadratic differential games · Joint-replenishment

## 1 Introduction

In this paper, we study linear quadratic games with set up costs monotonic on the number of *active players*, namely, players whose action is non-null.

Monotonicity can be used to model situations where imitation of others' behaviors on the part of single players is costly or rewarding (see, e.g., congestion games [12]).

Cost monotonicity arises naturally in multi-retailer inventory application whenever an opportune level of coordination of the retailers' replenishment strategies may lead to individual costs reduction (see, cf., [4]). Consider, for instance, a classical scenario where the transportation cost is shared among

D. Bauso
DICGIM, Università di Palermo, 90128 Palermo, Italy
Tel.: +39-320-4648142
E-mail: dario.bauso@unipa.it

R. Pesenti
Dipartimento di Management, Università di Venezia - Ca' Foscari, Italy
E-mail: pesenti@unive.it

all retailers who reorder at a given time instant, i.e., the active retailers. This is the case when a single truck delivers reordered goods to all active retailers. Evidently, 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 literature on cost-sharing games [18])

As a first contribution, we analyze some properties of the best response strategies. In particular, we show that best response strategies are non-idle in the sense that a player never switches from being inactive to active for fixed behaviors of the other players (fixed set up costs). Non-idleness is used to derive an iterative procedure to compute Nash equilibria.

We then turn to consider large population games and in doing so we link our study to *mean field games* [9,11,16]. **The relevance of considering a continuum of players is in that we can describe any applications where we have a large number, infinite in the limit, of players sharing a service facility.** The theory of mean field games was first formulated by Lasry and Lions [11] and studies the interactions among infinite homogeneous players, i.e., players that show identical behavior in a same situation. This theory has shown flexibility to many applicative fields from engineering [9], economics [8], [5], [6], physics and biology [10]. From a mathematical point of view, the mean field approach constitutes of a system of two PDEs: a Hamilton-Jacobi-Bellman equation and a Fokker-Planck equation coupled in a forward-backward way. Mean field games with linear quadratic costs have been analyzed in the literature [3]. In this context, the presence of an additional set up cost introduces an element of novelty.

As a second contribution, we show that, for the problem at hand, most properties enjoyed by the game with finite players still hold when the number of players tends to infinity. This observation allows us to claim that fixed points exist and that these are associated to mean field equilibria.

A third contribution of the paper is the analysis of the influence of an additional disturbance in the spirit of the literature on $H_\infty$-optimal control [1,16]. This part is not present in the conference version of this paper [13]. In this context, we show that, even in presence of a disturbance, best response strategies are non-idle. Building upon this result we propose a decomposition method to find a mean-field equilibrium.

The paper is organized as follows. In Section 2, we introduce the game. In Section 3 we analyze some properties of best response strategies. In Section 4, we discuss Nash equilibria. In Section 5, we consider the game with large population and illustrate the robust mean field approach. In Section 6, we provide numerical illustrations and conclude in Section 7.

**Notation.** We denote by $P = \{1, 2, \dots, n\}$ a set of $n$ players. We use index $i$ to refer to the generic $i$th player. Likewise, index $-i$ refers to all players other than $i$. We use $\mathbb{R}_+$ to denote the set of non-negative reals. Open and closed intervals between scalars $a$ and $b$ are denoted by $[a, b]$ and $(a, b)$ respectively. We use $[0, T]$ to denote a finite horizon from 0 to $T$. Given a function of time

$\phi(\cdot) : [0, T) \to \mathbb{R}$, we denote by $\phi(t)$ its value at time $t \in [0, T)$. We use $\phi[\xi](\cdot)$ to express the dependence of the function on a given parameter or function $\xi$.

## 2 Game definition

**Let us introduce the game model, which comprises of players, controls and associated measure theoretic spaces and cost functionals.**

Each player $i \in P$ is characterized by a state variable $x_i(\cdot) \in \mathbb{R}$, an initial state $x_i^0 \in \mathbb{R}$, a measurable control $t \mapsto u_i(t)$, taking value, for all $t \in [0, T)$, in the set $\mathbb{R}$. The state variable evolves according to the dynamics

$$\begin{cases} \dot{x}_i(t) = u_i(t), & t \in [0\,T) \\ x_i(0) = x_i^0 \end{cases}. \tag{1}$$

Let us also introduce a measurable opponents' control $t \mapsto u_{-i}(t)$, taking value, for all $t \in [0, T)$, in the set $\mathbb{R}^{n-1}$ and denote the sets for the measurable controls $u$ and $u_{-i}$ by

$$\begin{aligned} U_i &= \left\{ u_i : [0, T) \to \mathbb{R} \middle| u_i \text{ measurable} \right\}, \\ U_{-i} &= \left\{ u_{-i} : [0, T) \to \mathbb{R}^{n-1} \middle| u_{-i} \text{ measurable} \right\}. \end{aligned} \tag{2}$$

**In order to define the cost functional,** let $K$, $\alpha$, and $\beta$ be given positive constants; $\delta : \mathbb{R} \to \{0, 1\}$ be defined as in (3) and $a : \mathbb{R}^{n-1} \to \mathbb{R}_+$ as in (4) where $b$ is a constant greater than 0:

$$\delta(u_i(t)) = \begin{cases} 0 & \text{if } u_i(t) = 0, \\ 1 & \text{otherwise;} \end{cases} \tag{3}$$

$$a(u_{-i}(t)) = b + \frac{1}{n} \left( 1 + \sum_{j \in P \setminus \{i\}} \delta(u_j(t)) \right). \tag{4}$$

**The cost functional includes square penalties on final state, on current state and control as well as a fixed cost if the player is active. We reiterate here that a player is active at time $t$ if its control $u_i(t)$ is non-null.** The $i$th cost function is then

$$J_i(x_i^0, u_i, u_{-i}) = \int_0^T \left( \tfrac{K\delta(u_i(t))}{a(u_{-i}(t))} + x_i(t)^2 + \alpha u_i(t)^2 \right) dt + \beta x_i(T)^2. \tag{5}$$

**Observe that the fixed cost $\frac{K\delta(u_i(t))}{a(u_{-i}(t))}$ is distributed among all the active players and is monotonically decreasing with $a(.)$. The role of this term in the cost functional is to capture so called "crowd seeking" or "imitation" phenomena in that every player benefits from being active when the number of active players increases.**

## 3 Properties of non-dominated strategies

Let the set of the *non-anticipating strategies* for the first player be

$$M = \Big\{ \mu_i = \mu_i[x_i^0, \cdot] : U_{-i} \to U_i \Big| u^a_{-i}(s) = u^b_{-i}(s), \forall s \in [0,t] \Longrightarrow \tag{6}$$
$$\Longrightarrow \mu_i[x_i^0, u^a_{-i}](s) = \mu_i[x_i^0, u^b_{-i}](s), \forall s \in [0,t], \forall u^a_{-i}, u^b_{-i} \in U_{-i}, \forall t \in [0,T) \Big\}.$$

The notion of non-anticipating behavior strategies has a long history [2, 7, 15, 14, 17] in the theory of differential games. Essentially a non-anticipating strategy returns a control/decision at time $t$ as a function only of the history (for instance, the play path) up to time $t$ of the game.

Hereafter, we consider only strategies $\mu_i[x_i^0, u_{-i}]$ such that

1. for each $t \in [0,T)$, $u_i(t) = \mu_i[x_i^0, u_{-i}](t) = 0$ or $sign(\mu_i[x_i^0, u_{-i}](t)) = -sign(x_i(t))$ where $x_i(t)$ is solution of
$$\begin{cases} \dot{x}_i(t) = \mu_i[x_i^0, u_{-i}](t), & t \in [0,T) \\ x_i(0) = x_i^0 \end{cases}, \tag{7}$$
2. $\mu_i[x_i^0, u_{-i}]$ is piece-wise continuous.

**In other words, suppose that each agent can measure $x_i(t)$, then the above property establishes that the sign of the non-anticipating strategy, when different from zero, is opposite to the sign of $x_i(t)$.**

There is no loss of generality in such a choice as, given the player $i$ dynamics and cost, for no reason $i$ would control its state so to increase its state norm.

**After taking $u_i(t) := \mu_i[x_i^0, u_{-i}](t)$ for all $i$, the cost functional in (5) can be rewritten as:**

$$J_i(x_i^0, u_i, u_{-i}) = \int_0^T \left( \frac{K\delta(\mu_i[x_i^0, u_{-i}](t))}{a(u_{-i}(t))} + x_i(t)^2 + \alpha\mu_i[x_i^0, u_{-i}](t)^2 \right) dt + \beta x_i(T)^2 .$$

We say that a strategy $\mu_i[x_i^0, u_{-i}]$ is *non-idle* if, for each interval $[t_1, t_2]$, $0 \le t_1 < t_2 \le T$ in which the set up cost $\frac{K}{a(u_{-i}(t))}$ is non-decreasing, $u_i(t) := \mu_i[x_i^0, u_{-i}](t) > 0$ for all $t_1 \le t \le t_1 + \Delta t$ and $u_i(t) = 0$ for all $t_1 + \Delta t < t \le t_2$, for some $0 \le \Delta t \le t_2 - t_1$. Then, a player $i$ that implements a non-idle strategy, over the considered interval, is either always active or is always inactive or is first active and then inactive, but in no case it remains some time inactive before becoming active. Hereafter, we define *switching time instant*, the time in which a non-idle strategy $u_i(t)$ becomes non-active, i.e., the time $\inf\{t : u_i(t) = 0\}$.

For each interval $[t_1, t_2] \subset [0,T)$, let us denote by $\tilde{U}_{-i}[t_1, t_2]$ the set of the strategies $u_{-i}$ such that the set up cost $\frac{K}{a(u_{-i}(t))}$ does not decrease in $[t_1, t_2]$. Furthermore we say that $u_i^a$ is dominated by $u_i^b$ with respect to $\tilde{U}_{-i}[t_1, t_2]$, if

$$J_i(x_i^0, u_i^a, u_{-i}) - J_i(x_i^0, u_i^b, u_{-i}) > 0, \quad \text{for all } u_{-i} \in \tilde{U}_{-i}[t_1, t_2],$$

where $u_i^l(t) := \mu_i^l[x_i^0, u_{-i}](t)$, the label $l \in \{a, b\}$, and $\mu_i^l[x_i^0, u_{-i}](t)$ is a given strategy for all $u_{-i} \in \tilde{U}_{-i}[t_1, t_2]$.

Hereafter, we say that a strategy is *non-dominated* if it is not dominated with respect to $\tilde{U}_{-i}[t_1, t_2]$, for each interval $[t_1, t_2] \subset [0, T)$ such that the set up cost $\frac{K}{a(u_{-i}(t))}$ does not decrease.

The following two lemmas prove that a non-dominated strategy for a player $i$ is non-idle and that the instantaneous set up cost paid by an active player cannot decrease over time.

**Lemma 1** *Given a player $i$ and an interval $[t_1, t_2] \subset [0, T)$ such that the set up cost $\frac{K}{a(u_{-i}(t))}$ does not decrease, a strategy $u_i$ that is not non-idle is dominated.*

*Proof* Consider two strategies $\mu_i^a[x_i^0, u_{-i}]$ and $\mu_i^b[x_i^0, u_{-i}]$ such that $\mu_i^a[x_i^0, u_{-i}](t) = \mu_i^b[x_i^0, u_{-i}](t)$ for all $t \neq [t_1, t_2]$ and

$$\begin{cases} \mu_i^a[x_i^0, u_{-i}](t) = 0, & t \in [t_1, t_1 + \Delta t) \\ \mu_i^a[x_i^0, u_{-i}](t) \neq 0, & t \in [t_1 + \Delta t, t_2) \end{cases}$$

$$\begin{cases} \mu_i^b[x_i^0, u_{-i}](t) = \mu_i^a[x_i^0, u_{-i}](t + \Delta t), & t \in [t_1, t_2 - \Delta t) \\ \mu_i^b[x_i^0, u_{-i}](t) = 0, & t \in [t_2 - \Delta t, t_2) \end{cases}.$$

To see this, let us denote by $x_i(t_1) = \int_0^{t_1} u_i^a(t)dt + x_i^0 = \int_0^{t_1} u_i^b(t)dt + x_i^0$, and $x_i(t_2) = \int_0^{t_2} u_i^a(t)dt + x_i^0 = \int_0^{t_2} u_i^b(t)dt + x_i^0$. In the following, we prove that

$$J_i(x_i^0, u_i^a, u_{-i}) - J_i(x_i^0, u_i^b, u_{-i}) > 0, \quad \text{for all } u_{-i} \in \tilde{U}_{-i}[t_1, t_2]. \tag{8}$$

Indeed, the costs induced by the two strategies are equal for $0 \leq t \leq t_1$ and $t_2 \leq t \leq T$, as in such interval the two strategies assume the same values and induce the same states for the player.

Then consider the interval $t_1 \leq t \leq t_2$, the cost paid by $u_i^a$ is

$$\int_{t_1}^{t_1+\Delta t} x_i(t_1)^2 dt + \int_{t_1+\Delta t}^{t_2} \frac{K}{a(u_{-i}(t))} dt + \\ + \int_{t_1+\Delta t}^{t_2} \alpha u_i^a(t)^2 dt + \int_{t_1+\Delta t}^{t_2} x_i(t)^2 dt.$$

Differently, the cost paid by $u_i^b$ is

$$\int_{t_1}^{t_2-\Delta t} x_i(t)^2 dt + \int_{t_1}^{t_2-\Delta t} \frac{K}{a(u_{-i}(t))} dt + \\ + \int_{t_1}^{t_2-\Delta t} \alpha u_i^a(t+\Delta t)^2 dt + \int_{t_2-\Delta t}^{t_2} x_i(t_2)^2 dt.$$

Now, note that $\int_{t_1+\Delta t}^{t_2} \frac{K}{a(u_{-i}(t))} dt \geq \int_{t_1}^{t_2-\Delta t} \frac{K}{a(u_{-i}(t))} dt$ since $\frac{K}{a(u_{-i}(t))}$ does not decrease for $t_1 \leq t \leq t_2$. In addition, observe that $\int_{t_1+\Delta t}^{t_2} u_i^a(t)^2 dt =$

$\int_{t_1}^{t_2-\Delta t} u_i^a(t+\Delta t)^2 dt$, and $\int_{t_1+\Delta t}^{t_2} x_i(t)^2 dt = \int_{t_1}^{t_2-\Delta t} x_i(t)^2 dt$, then the inequality (8) holds true, as it becomes

$$J_i(x_i^0, u_i^a, u_{-i}) - J_i(x_i^0, u_i^b, u_{-i})) \geq (x_i(t_1)^2 - x_i(t_2)^2)\Delta t > 0, \quad \text{for all } u_{-i} \in \tilde{U}_{-i}[t_1, t_2].$$

Hence, the lemma is proved. □

**Lemma 2** *If all the players play non-dominated strategies, $\frac{K}{a(u_{-i}(t))}$ does not decrease for all $i \in P$ and all $0 < t < T$.*

*Proof* The value $\frac{K}{a(u_{-i}(t))}$ decreases for some player $i$ and some $0 < t < T$, if there is at least another player $j$ that in $t^j$ switches from being inactive to being active.

Let us first prove the result under the assumption that no more than one player can become active at each time instant. Then, there exists a value $t_1 \geq 0$, a value $\Delta t > 0$, and an interval $0 \leq t_1 < t_1 + \Delta t = t^j < t_2 < T$, such that player $j$ is first inactive and then active even if $\frac{K}{a(u_{-j}(t))}$ remains constant. Then, by Lemma 1, player $j$ cannot be playing a non-dominated strategy.

Given the above argument, for $\frac{K}{a(u_{-i}(t))}$ to decrease for some player $i$, we must assume that a set $S$ of players, with $|S| \geq 2$, coordinates to switch from being inactive to being active at time $t^j > 0$. Even in this case, there exists a value $t_1 \geq 0$, a value $\Delta t$, and an interval $0 \leq t_1 < t_1 + \Delta t = t^j < T$, such that $\frac{K}{a(u_{-s}(t))}$ remains constant, for all $s \in S$. Following the same line of reasoning of Lemma 1, it is immediate to prove that strategies that coordinate the switch at time $t_1$ induce less costs for all the players in $S$ and then they dominate the current strategies (that coordinate the switch at time $t_j$). We can conclude that the strategy that coordinates the switch at time $t_j$ cannot be a non-dominated one. □

As there is no loss of generality in assuming that all the players play non-dominated strategies, hereafter, for all $i \in P$, we can assume that $\frac{K}{a(u_{-i}(t))}$ does not decrease over time. Furthermore, we can also assume that $U_{-i} = \bigcup_{0<t_2<T} \tilde{U}_{-i}[0, t_2]$. Consequently, we can rephrase our previous definition of dominance as follows

We say that $u_i^a$ is dominated by $u_i^b$, if

$$J_i(x_i^0, u_i^a, u_{-i}) - J_i(x_i^0, u_i^b, u_{-i}) > 0, \quad \text{for all } u_{-i} \in U_{-i}.$$

We say that $u_i^a$ is *weakly* dominated by $u_i^b$ when the previous condition holds weakly, that is, $J_i(x_i^0, u_i^a, u_{-i}) - J_i(x_i^0, u_i^b, u_{-i}) \geq 0$ for all $u_{-i} \in U_{-i}$.

We define *switching feedback strategy at* $\tau$ any control $\mathbf{u}_i[\tau]$ that satisfies:

$$\begin{aligned} \mathbf{u}_i[\tau](t) &:= \mu_i[x_i^0, u_{-i}](t) \\ &= \begin{cases} f(t,\tau)x_i(t) & \text{for } 0 \leq t \leq \tau \\ 0 & \text{for } \tau < t \leq T \end{cases}. \end{aligned} \tag{9}$$

We will show later on that the class of strategies above also includes best response strategies. These strategies will be obtained from solving a linear quadratic control problem, which justifies linearity in state $x_i(t)$.

In the hypotheses of the above two lemmas, the following corollary holds.

**Corollary 1** *For all $\tau$ such that $0 \leq \tau \leq T$ and for each player $i$ there exists a unique non-dominated switching feedback strategy $\mathbf{u}_i[\tau]$ as in (9).*

*Proof* For all $\tau$ such that $0 \leq \tau \leq T$, the non-dominated strategy is

$$\mathbf{u}_i[\tau](t) = \begin{cases} \tilde{u}_i(t) & \text{for } 0 \leq t \leq \tau \\ 0 & \text{for } \tau < t \leq T \end{cases},$$

where $t \mapsto \tilde{u}_i(t)$ solves the problem below:

$$\begin{aligned}\tilde{u}_i &:= \arg\min\{\int_0^\tau (\frac{K}{a(u_{-i}(t))} + x_i(t)^2 + \alpha u_i(t)^2)dt \\ &\quad +((T-\tau)+\beta)x_i(\tau)^2\} \\ &= \arg\min\{\int_0^\tau (x_i(t)^2 + \alpha u_i(t)^2)dt \\ &\quad +((T-\tau)+\beta)x_i(\tau)^2\}.\end{aligned}$$

The equality holds as the value of $\frac{K}{a(u_{-i}(t))}$ is independent of $\tilde{u}_i(t)$. In addition, if $x_i^0 > 0$, the second problem, because of the quadratic structure of the costs, presents a unique optimal continuous solution of type $\tilde{u}_i(t) = f(t,\tau)x(t)$. The latter strategy is independent of the fixed cost and is different from zero for $0 \leq t \leq \tau$, as it can be directly verified explicitly solving the optimization problem. In this context note that this problem is a quadratic control problem that can be analytically solved using the maximum principle or a differential Riccati equation. □

Hereafter, for any realization of $u_{-i}$ and therefore $\frac{K}{a(u_{-i}(t))}$, we say that the best response strategy of player $i$ is the switching feedback control at $t_i^*$ defined as:

$$\begin{aligned}\mathbf{u}_i[t_i^*] &:= \arg\min_{\mathbf{u}_i[\tau]:0\leq\tau\leq T}\{J(x_0^i, \mathbf{u}_i[\tau], u_{-i})\} \\ &=: \mu_i^*[x_i^0, u_{-i}].\end{aligned} \tag{10}$$

Note that a strategy solution of (10) always exists and is unique, as it can be verified analytically that $J(x_0^i, \mathbf{u}_i[\tau], u_{-i})$ is a continuous strictly convex function of $\tau$.

The next lemma relates the switching times of two different players.

**Lemma 3** *Given two players $i$ and $j$, such that $x_i^0 \geq x_j^0 > 0$, if $\mathbf{u}_j[t_j^*]$ is a best response strategy for player $j$, then all the strategies of player $i$ $\mathbf{u}_i[\tau]$ where $\tau < t_j^*$ are dominated.*

*Proof* The statement of this lemma can be directly verified by explicitly determining the values of $J(x_i^0, \mathbf{u}_i[\tau], u_{-i})$ and $J(x_j^0, \mathbf{u}_j[\tau], u_{-i})$ and observing that $J(x_i^0, \mathbf{u}_i[\tau], u_{-i})$ decreases for $\tau \in [0, t_j^*]$ as long as $J(x_j^0, \mathbf{u}_j[\tau], u_{-j})$ decreases in the same interval. The latter is true as $\mathbf{u}_j[t_j^*]$ is the best response strategy for player $j$. □

The above lemma can be rephrased by saying that according to their best responses if player $j$ is active then player $i$ is active too.

The next theorem states under which condition a player active at time $t = 0$ becomes inactive in a following time instant. Specifically, it points out the dependence of the switching time instant of a non-dominated strategy on the value of the fixed cost.

**Theorem 1** *According to a non-dominated strategy, player $i$ is active as long as the instantaneous set up cost satisfies the following condition*

$$\frac{K\alpha}{a(u_{-i}(t))} \leq (((T-t)+\beta)x_i(t))^2. \tag{11}$$

*When the above condition is satisfied, a non-dominated strategy is bounded as in (12), where $\gamma := -((T-t)+\beta)x_i(t)$ and $\Delta := (((T-t)+\beta)x_i(t))^2 - K\alpha/a(u_{-i}(t))$:*

$$\frac{\gamma - \sqrt{\Delta}}{\alpha} \leq u_i(t) \leq \frac{\gamma + \sqrt{\Delta}}{\alpha}. \tag{12}$$

*Proof* We analyze under which circumstances player $i$, active at time $t$, remains so for a further time interval $\Delta t > 0$. Then, let us look at interval $[t, t+\Delta t]$ and consider a *non-null strategy*, where $u_i(t) > 0$, and a *null strategy*, with $u_i(t) = 0$, for $t \in [t, t+\Delta t]$. Let us compare the cost to go from $t$ to $T$ induced by such strategies. The cost to go of the null strategy is

$$\int_t^{t+\Delta t} x_i(t)^2 d\tau + \int_{t+\Delta t}^T x_i(t)^2 d\tau + \beta x_i(T)^2.$$

Similarly, the cost to go of the non-null strategy is the one displayed below, with $\Delta x_i = \int_t^{t+\Delta t} u_i(\tau)d\tau$:

$$\int_t^{t+\Delta t} \left( \frac{K}{a(u_{-i}(\tau))} + \alpha u_i(\tau)^2 + x_i(\tau)^2 \right) d\tau$$
$$+ \int_{t+\Delta t}^T (x_i(t) + \Delta x_i)^2(t) d\tau + \beta(x_i(t) + \Delta x_i)^2.$$

Then, we compute the difference of the two costs for $\Delta t \to 0$, to obtain

$$\left( \frac{K}{a(u_{-i}(t))} + \alpha u_i(t)^2 \right) dt + 2(T-t)x_i(t)dx_i + 2\beta x_i(t)dx_i.$$

Since $dx_i = u_i(t)dt$, after dividing by $dt$ the latter can be rewritten as

$$\frac{K}{a(u_{-i}(t))} + \alpha u_i(t)^2 + 2(T-t)x_i(t)u_i(t) + 2\beta x_i(t)u_i(t).$$

Hence, the non-null strategy provides a lower cost than the null strategy, and therefore we would rather have $u_i(t) > 0$ in $t$, if and only if the above difference is non-positive, that is if $\alpha u_i^2(t) + 2((T-t)+\beta)x_i(t)u_i(t) + \frac{K}{a(u_{-i}(t))} \leq 0$. In turn, this last inequality holds if and only if conditions (11) and (12) are satisfied. □

An immediate consequence of the above theorem is that a player is certainly never active if $K\alpha > (b+1-1/n)((T+\beta)x_i^0)^2$.

Lemma 3 also implies that if all the players $j \neq i$ play their best responses, and using $\mathbf{u}_{-i}[t^*_{-i}]$ to denote their set of best response strategies in compact form, then it holds

$$\frac{K\alpha}{a(\mathbf{u}_{-i}[t^*_{-i}](t_i^*))} = (((T-t_i^*)+\beta)x_i(t_i^*))^2. \tag{13}$$

## 4 Nash equilibria

In this section, we show how to determine a set of Nash equilibria strategies for players in $P$ under the assumption that $0 < x_1^0 \leq x_2^0 \leq \ldots \leq x_n^0$. To this end, we heavily exploit Lemma 3 to determine the best response of the players.

Preliminarily, for each player $i$ let us define $\hat{K}_i := \frac{K}{b+\frac{n-i+1}{n}}$ and consider the following ***auxiliary optimal control problem***, independent of $u_{-i}$,

$$\min \hat{J}_i(x_i^0, u_i, m) = \int_0^T [\hat{K}_i\delta(u_i(t)) + x_i(t)^2 + \alpha u_i(t)^2]dt + \beta x_i(T)^2. \tag{14}$$

Applying the same line of reasoning used in the previous sections, we can affirm that the optimal solution of (14) is a switching strategy with switching time instant $\hat{t}_i$ such that $\hat{t}_i = 0$, if $\hat{K}_i\alpha > ((T+\beta)x_i^0)^2$, $\hat{t}_i = T$, if $\hat{K}_i\alpha < ((T+\beta)x_i^0)^2$, otherwise $0 \leq \hat{t}_i \leq T$ is the solution of

$$\hat{K}_i\alpha = (((T-\hat{t}_i)+\beta)x_i(\hat{t}_i))^2. \tag{15}$$

where $x_i(\hat{t}_i)$ is the trajectory of $i$ when the switching strategy $\mathbf{u}_i[\hat{t}_i]$ is implemented. Note that $\mathbf{u}_i[\hat{t}_i]$ is also the best response strategy for player $i$ if $\frac{K}{a(u_{-i}(\hat{t}_i))} = \hat{K}_i$, that is, if at the switching time instant the only active players are the ones with state greater than or equal to $x_i(\hat{t}_i)$, or, that is the same, as the trajectories of best strategies cannot intersect, the only active players are the ones with initial state greater than or equal to $x_i^0$. In other words, $\hat{t}_i$ is the last time instant in which it is convenient for player $i$ to remain active even if there are only other $n-i$ active players.

Lemma 3 implies that if all the players play their best responses, then strategy $\mathbf{u}_1[t_1^*]$ for player 1 must satisfy:

$$\frac{K}{a(\mathbf{u}_{-1}[t^*_{-1}](t))} = \begin{cases} \frac{K}{b+1} =: \hat{K}_1 & \text{if } 0 \leq t \leq t_1^* \\ 0 & \text{if } t_1^* < t \leq T \end{cases}.$$

From the latter condition, and invoking conditions (13)-(15), we can infer that $t_1^* = \hat{t}_1$ and also that player 1 has a unique non-dominated strategy $\mathbf{u}_1[t_1^*] = \mathbf{u}_1[\hat{t}_1]$.

Let us now consider the generic player $i > 1$. It holds

$$t_i^* = \max\{t_{i-1}^*, \hat{t}_i\}. \tag{16}$$

Indeed, Lemma 3 implies that player $i$ must be active at least as long as player $i-1$ is active, hence $t_i^* \geq t_{i-1}^*$. Lemma 3 also implies that if $t_i^* > t_{i-1}^*$, then in $t_i^*$ the only active players are the ones with state greater than or equal to $x_i(t_i^*)$. These two observations imply that $t_i^*$ is either equal to $t_{i-1}^*$ or equal to $\hat{t}_i$, that is that player $i$ can consider only two strategies $\mathbf{u}_i[t_{i-1}^*]$ or $\mathbf{u}_i[\hat{t}_i]$. Finally, note that player $i$ chooses $\mathbf{u}_i[\hat{t}_i]$, if $\hat{t}_i > t_{i-1}^*$ because it is convenient for player $i$ to remain active even if only other $n-i$ players are active after $t_{i-1}^*$.

## 5 Large number of players

Let us now reformulate our game from a mean field perspective. To this end, let $m_t(\cdot) : \mathbb{R}^+ \to [0,1]$ be the distribution of the players' states at time $t$. Given this, for all $x \in \mathbb{R}^+$, the differential term $dm_t(x) = m_t(x)dx$ accounts for the percentage of players whose state $x(t) \in [x, x+dx]$. **For sake of simplicity, suppose that $m_t(x)$ is continuous and differentiable in the variable $x$. In the following we revise the steps that lead to a Nash equilibrium described in the previous section. To this end, we need to introduce the auxiliary optimal control problem (equivalent to (14)). In this context, consider the following notation.** For each player $i$, given his current state $x_i(t)$, we define the function $\tilde{a}_i(m_t) := b + \int_{x_i(t)}^{+\infty} dm_t(x)$. Essentially, $\tilde{a}_i(m_t)$ depends on the percentage of players with states greater than $x_i(t)$. **The considerations over the state trajectories that precede Lemma 3 imply the following time-invariance condition $\int_{x_i(t)}^{+\infty} dm_t(x) = \int_{x_i^0}^{+\infty} dm_0(x)$.** Then, we can rewrite $\hat{K}_i$ as a function of the distribution at time $t$:

$$\hat{K}_i(m_t) = \frac{K}{\tilde{a}_i(m_t)}. \tag{17}$$

With the above definition in mind, the objective function of the auxiliary control problem (14) takes on the form

$$\hat{J}_i(x_i^0, u_i, m) = \int_0^T [\hat{K}_i(m_t)\delta(u_i(t)) + x_i(t)^2 + \alpha u_i(t)^2]dt + \beta x_i(T)^2.$$

If we denote by $v$ the value function, then the mean-field problem appears as:

$$v(x_i^0, m) = \inf_{u_i} \hat{J}_i(x_i^0, u_i, m) \tag{18}$$

$$\dot{x}_i(t) = u_i(t). \tag{19}$$

**Again, following the same line of reasoning used in the previous sections, we can affirm the the optimal solution of the above problem is a switching strategy with switching time instant $\hat{t}_i$. We use such a value in Subsection 5.1 to determine player $i$ best strategy in presence of a large number of players. In the same context, in Subsection 5.2, we describe the evolution over time of the players' state cumulative distribution $Q(y,t) := \int_y^{\infty} dm_t(x)$.**

### 5.1 Generic player $i$ best strategy

Let us first consider the generic player $i$ best strategy. The recursive equation (16) allows player $i$ to determine the switching time instant $t_i^*$ of its best strategy $\mathbf{u}_i[t_i^*]$ and hence to characterize the strategy itself. Unfortunately, equation (16) is of no practical use in presence of a large number of players as it would force player $i$ to wait for the decision of all the players from 1 to $i-1$ before being able to compute $t_i^*$. For this reason, player $i$ may decide to play an approximatively optimal strategy $\mathbf{u}_i[\tilde{t}_i^*]$ based on an estimate $\tilde{t}_i^*$ of $t_i^*$. In particular, we observe that we may rewrite equation (16) as

$$t_1^* = \max\{\hat{t}_i, \max_{j<i}\{\hat{t}_j\}\}.$$

Then, for any subset $S \subseteq \{1, 2, \dots, i-1\}$, the value

$$\tilde{t}_i^* = \max\{\hat{t}_i, \max_{j\in S}\{\hat{t}_j\}\} \leq t_i^*$$

is an estimate, and in particular a lower bound, of the switching time instant $t_i^*$. Needless to say that the $\tilde{t}_i^*$ becomes a better and better estimate of $t_i^*$, and hence $\mathbf{u}_i[\tilde{t}_i^*]$ a better and better approximation of the best strategy $\mathbf{u}_i[t_i^*]$, as the subset $S$ includes more and more elements of $\{1, 2, \dots, i-1\}$.

The above kind of approximate strategy requires that player $i$ communicates with the players in $S$ to acquire the values of $\hat{t}_j$. Player $i$ can play a different approximate strategy that just needs the observation of the behavior of player $i-1$ as described in the following.

Player $i$ remains active as long as $i-1$ is active. Then, at the switching time instant $t_{i-1}$ of $i-1$, player $i$ decides whether it is convenient to remain active or not and for how long. If all the players use such an approximate strategy, this approximation identifies the best strategy from the switching time instant of $i-1$ on. Indeed, from such time instant player $i$ can determine its best strategy based on the number of active players: all the players from 1 to $i-1$ are not active any more, viceversa, all the players from $i+1$ to $n$ remain active at least as long as $i$ is active. Unfortunately, player $i$ cannot play its best strategy until the switching time instant of $i-1$ as it cannot a priori know its value. As the optimal choice would be a strategy of type $\tilde{u}_i(t) = f(t)x(t)$, player $i$ can approximate such a strategy, as an example fixing the value of $f(t)$ to a constant. In such a context a reasonable choice could be $f(t) \approx -\frac{\beta}{2\alpha}$ as $f(\tilde{t}) = -\frac{T+\beta}{2\alpha}$ for $0 \leq \tilde{t} \leq T$. Then, the approximate strategy for player $i$ turns out to be

$$\begin{aligned} &\mu_i[x_i^0, u_{-i}](t) = \\ &= \begin{cases} -\frac{\beta}{2\alpha}x & \text{for } 0 \leq t \leq t_{i-1} \\ \mu_i^*[x_i(t_{i-1}), u_{-i}](t - t_{i-1}) & \text{for } t_{i-1} < t \leq T \end{cases}, \end{aligned}$$

where $\mu_i^*[x_i(t_{i-1}), (\cdot)]$ is the best response strategy with initial state $x_i(t_{i-1})$.

### 5.2 Evolution of the cumulative distribution

We now study how $Q(y,t) = \int_y^\infty dm_t(x)$ evolves over time. Specifically, as the trajectories of players with different initial states do not cross, it must satisfy the transport equation

$$\frac{\partial}{\partial t}Q(y,t) = -u(y,t)\frac{\partial}{\partial y}Q(y,t), \tag{20}$$

where $u(y,t)$ is the control applied at time $t$ by a player with state $x(t) = y$.

As the best strategy of a player depends only on its initial state, we observe that, for each initial state $x_0$ and time instant $t$ we can write $x(t) - x_0 = \int_0^t \tilde{u}(\tau)d\tau$, where $\tilde{u}$ is the best strategy of a player with initial state $x_0$. Then the solution of (20) is

$$Q(y,t) = Q(y - \int_0^t \tilde{u}(\tau)d\tau, 0),$$

as it can be directly verified computing the partial derivatives of $Q(y,t)$ and exploiting the fact that $\tilde{u}(t) = u(x(t),t)$.

The above results generalize to all the cases in which players choose strategies that depend only on the initial states. We also observe that the more the time to go $T-t$ gets closer to 0 the higher must be the state of a player for being convenient for the player to be active. Formally, there exists an increasing function $\lambda : [0,T] \to \mathbb{R}$ such that

$$u(y,t) = \begin{cases} 0 & \text{for } y \le \lambda(t) \\ f(t)y & \text{for } y > \lambda(t) \end{cases}.$$

Hence, we can rewrite $Q(y,t)$ as

$$Q(y,t) = \begin{cases} Q(y - \int_0^t \tilde{u}(\tau)d\tau, 0) & \text{for } 0 \le t \le \lambda^{-1}(y) \\ Q(y, \lambda^{-1}(y)) & \text{for } \lambda^{-1}(y) < t \le T \end{cases}.$$

### 5.3 Robust mean-field formulation

In this section, we analyze the influence of an additive disturbance on the mean-field dynamics.

Each player $i \in P$ is characterized by a state variable $x_i(\cdot) \in \mathbb{R}$, an initial state $x_i^0 \in \mathbb{R}$, a measurable control $t \mapsto u_i(t)$ and disturbance $t \mapsto z_i(t)$, taking value, for all $t \in [0,T)$, in the set $\mathbb{R}$. The state variable evolves according to the dynamics

$$\begin{cases} \dot{x}_i(t) = u_i(t) + \sigma z_i(t), & t \in [0\,T) \\ x_i(0) = x_i^0 \end{cases}. \tag{21}$$

Here $\sigma$ is a positive scalar weighting the influence of the disturbance $z_i(t)$ on $x_i(t)$. In the spirit of the literature on $H_\infty$-optimal control [1], we consider the following objective function to minimize with respect to $u_i$ and maximize with respect to $z_i$

$$J_i(x_i^0, u_i, m, z_i) = \int_0^T [\hat{K}_i(m_t)\delta(u_i(t)) + x_i(t)^2 + \alpha u_i(t)^2]dt - \zeta^2 \int_0^T z_i(t)^2 dt + \beta x_i(T)^2.$$

Note that the above objective function differs from (14) in the additional term $\zeta^2 \int_0^T z_i(t)^2 dt$, which accounts for the disturbance energy.

Denote by $v$ the value function, then the robust mean-field problem appears as:

$$v(x_i^0, m) = \inf_{u_i} \sup_{z_i} J_i(x_i^0, u_i, m, z_i) \quad (22)$$

$$\dot{x}_i(t) = u_i(t) + \sigma z_i(t). \quad (23)$$

The value of $z_i$ that maximizes the above cost is called *worst-case disturbance*. Also, if the value function is upper bounded by zero then we can conclude that the ratio between the original cost (5) obtained from implementing the $H_\infty$ optimal control (the value of $u_i$ that minimizes the above cost) and the disturbance energy is upper bounded by $\zeta^2$.

In accordance with the results illustrated in the preceding sections, the next result elaborates on the importance of non-idle strategies.

**Lemma 4** *Given the robust mean field formulation (22)-(23), a strategy that is not non-idle is dominated.*

*Proof* Consider the systems $\dot{x}_i(\tau) = u_i(\tau) + \sigma z_i(\tau)$ with

$$\begin{array}{l} u_i(\tau) = 0 \text{ for } t \le \tau \le t + \Delta t \\ u_i(\tau) \neq 0 \text{ for } t + \Delta t < \tau \le t + 2\Delta t \end{array} \quad (24)$$

and $\dot{y}_i(\tau) = u_i(\tau) + \sigma z_i(\tau)$ with

$$\begin{array}{l} u_i(\tau) \neq 0 \text{ for } t \le \tau \le t + \Delta t \\ u_i(\tau) = 0 \text{ for } t + \Delta t < \tau \le t + 2\Delta t \end{array} \quad (25)$$

such that $x_i(t) = y_i(t)$ and $\hat{K}_i(m_t)$ is not increasing over time.

Consider the following two problems

$$\hat{v}(x_i(t)) := \lim_{\Delta t \to 0} \min_{u_i} \sup_{z_i} \left\{ \int_t^{t+\Delta t} (x_i^2(t) - \zeta^2 z_i^2(t))dt + \right.$$

$$\left. + \int_{t+\Delta t}^{t+2\Delta t} (\hat{K}_i(m_t) + x_i^2(t) + \alpha u_i^2(t) - \zeta^2 z_i^2(t))dt + \phi x_i^2(t + 2\Delta t) - \phi x_i^2(t) \right\} =$$

$$= \hat{K}_i(m_t) + \alpha u_i^{*2} - 2\zeta^2 z_i^{*2} + 2x_i(t)(x_i(t) + 2\phi(u_i + 2z_i^*\sigma))$$

where $u_i^* = -\frac{\phi}{2\alpha}x_i(t)$ and $z_i^* = -\frac{\phi\sigma}{2\zeta^2}x_i(t)$ and

$$\begin{aligned}\tilde{v}(x_i(t)) := \lim_{\Delta t\to 0}\min_{u_i}\sup_{z_i}\Bigg\{\int_t^{t+\Delta t}(\hat{K}_i(m_t)+y_i^2(t)+\alpha u_i^2(t)-\zeta^2 z_i^2(t))dt+ \\ +\int_{t+\Delta t}^{t+2\Delta t}(y_i^2(t)-\zeta^2 z_i^2(t))dt+\phi y_i^2(t+2\Delta t)-\phi y_i^2(t)\Bigg\} = \\ = \hat{K}_i(m_t)+\alpha u_i^{*2}-2\zeta^2 z_i^{*2}+2y_i(t)(x(t)+2\phi(u_i+2z_i^*\sigma))\end{aligned}$$

where $u_i^* = -\frac{\phi}{2\alpha}y_i(t)$ and $z_i^* = -\frac{\phi\sigma}{2\zeta^2}y_i(t)$ .

Now, if $\hat{K}_i(m_t)$ remains constant between $t$ and $t+2\Delta t$, consider the difference

$$\frac{1}{T^2}\hat{v}(x_i(t)) - \tilde{v}(x_i(t)) = -2u_i^* x_i(t) > 0$$

otherwise, if $\hat{K}_i(m_t)$ changes value in $t$, consider the difference

$$\hat{v}(x_i(t)) - \tilde{v}(x_i(t)) = K_i(m_{t^-}(\cdot)) - \hat{K}_i(m_{t^+}(\cdot)) > 0.$$

The above results imply that for sufficiently small $\Delta t$ strategy (24) is (at least weakly) dominated by the non-idle strategy (25). Hence, a strategy that is not non-idle is always at least weakly dominated by a non-idle strategy and the lemma is proved. □

The importance of the above result is in that we can decompose the problem into two subproblems. The first subproblem describes the system behavior in the interval $[0,\tau]$ during which the player is supposed to be active, whereas the second subproblem captures the behavior in the interval $[\tau, T]$ where the player is supposed to be inactive. It is worth noting that $\hat{K}_i(m_t)$ enters only in the computation of $\tau$. Once we have computed $\tau$, we can decompose the problem and get rid of $\hat{K}_i(m_t)$. Actually, It turns out that $\hat{K}_i(m_t)$ no longer appears in the rest of the paper. More formally, subproblem 1 reads as:

**(Subproblem 1)** $\inf_{u_i}\sup_{z_i}\int_0^\tau[x_i(t)^2+\alpha u_i(t)^2]dt - \zeta^2\int_0^\tau z_i^2(t)dt + \hat{\Phi}(x_i(\tau))$

$$\dot{x}_i(t) = u_i(t) + \sigma z_i(t),$$

where $\hat{\Phi}(x_i(\tau))$ is the best quadratic approximation of the optimal cost-to-go and the way to obtain this approximation is yet to be introduced.

Conversely, Subproblem 2 can be formalized as:

**(Subproblem 2)** $\Phi(x_i(\tau)) := \sup_{z_i}\int_\tau^T x_i(t)^2 dt - \zeta^2\int_\tau^T z_i^2(t)dt + \beta x_i(T)^2.$

$$\dot{x}_i(t) = \sigma z_i(t).$$

We are in a position to state the main results of this section. The first result provides a solution to Subproblem 2 and establishes a relation between the worst-case disturbance and the current state.

**Theorem 2** *Solution to Subproblem 2 is the following worst-case disturbance:*

$$z_i^*(t) = \frac{\sigma}{2\zeta^2} p(t) x_i(t) \tag{26}$$

*where $p(t)$ satisfies the differential algebraic Riccati equation:*

$$\dot{p}(t) = -\frac{\sigma^2}{2\zeta^2} p(t)^2 - 2, \quad p(T) = 2\beta. \tag{27}$$

*Proof* The main idea of the proof is to exploit the quadratic structure of the cost and the related Riccati theory. Actually the optimization problem is convex and quadratic on the the disturbance. Then, the worst-case disturbance is a function of the state $x(t)$ through $p(t)$ as established in (26). It remains to notice that (27) is the differential Riccati equation specialized to the model under study. □

The second result provides a solution to Subproblem 1 in terms of the optimal control and worst-case disturbance. Suppose $\hat{\Phi}(x_i(\tau)) := \phi x_i(\tau)^2$ for a given scalar $\phi > 0$ then we can establish the following result.

**Theorem 3** *Solution to Subproblem 1 is the following pair of optimal control and worst-case disturbance:*

$$u_i^*(t) = -\frac{1}{2\alpha} p(t) x_i(t) \tag{28}$$

$$z_i^*(t) = \frac{\sigma}{2\zeta^2} p(t) x_i(t) \tag{29}$$

*where $p(t)$ satisfies the differential algebraic Riccati equation:*

$$\dot{p}(t) = \Big[\frac{1}{2\alpha} - \frac{\sigma^2}{2\zeta^2}\Big] p(t)^2 - 2, \quad p(\tau) = 2\phi. \tag{30}$$

*Proof* Observe that the optimization problem has a quadratic structure of the cost and as such we can refer to the related Riccati theory. Then, both the optimal control and the worst-case disturbance are functions of the state $x_i(t)$ through $p(t)$ as established in (28)-(29). It remains to notice that (30) is the differential Riccati equation specialized to the model under study. □

Note that the distribution depends upon the optimal control and disturbance following the *advection equation*

$$\partial_t m_t(x) + div\,(m_t(x)(u^*(x) + \sigma z^*)) = 0, \quad \forall\, x \in \mathbb{R}^+,$$

where *div* stands for the divergence operator.

**In this section we have supposed that $m_t(x)$ is differentiable in $x$. Should this not be true we need to resort to so-called *weak solutions* of the above PDE as defined in [5], Definition 1.**

| $n$ | $T$ | $\alpha$ | $\beta$ | $K$ | $x_i(0)$ | $\tau$ |
|---|---|---|---|---|---|---|
| 200 | 20 | 20 | 1 | 1600 | $[0, 150]$ | $1, 1.5, \ldots, T$ |

**Table 1** Simulations data.

## 6 Numerical illustrations

In this section we provide numerical illustrations for a large number of players evolving according to system (21) and with simulations data as reported in Table 1.

In particular, the number of players is $n = 200$ and the horizon is $T = 20$.

The parameters appearing in the cost (5) are set as follows: $\alpha = 20$, $\beta = 1$, and $K = 1600$. Initial states $x_i(0)$ for all $i$ are uniformly distributed over the interval $[0, 150]$. We also discretize the set of possible switching times and so $\tau \in \{1, 1.5, \ldots, T\}$.

The algorithm used to numerically illustrate the players' behavior accepts the simulations data as input and returns the best response strategies $\mathbf{u}_i[t_i^*]$ as in (10) and the associated state distribution $dm(x, t)$.

The algorithm is designed as follows. First, we initialize the state by using the Matlab in-built functions *rand* to generate a realization of the random variable $x(0)$ and *sort* to reorder the agents for increasing states.

For every possible value of the switching time $\tau \in \{1, 1.5, \ldots, T\}$, and for all players $i = 1, \ldots, n$, we compute the optimal (we say optimal as for fixed $\tau$ the strategy $\mathbf{u}_i[\tau]$ is independent of the other players' behaviors) strategy $\mathbf{u}_i[\tau]$ as in (9).

To do this, we solve the following differential Riccati equation in the scalar variable $p(t)$ $t \in [0, \tau]$:

$$\dot{p}(t) = \frac{1}{2\alpha} p(t)^2 - 2, \quad p(\tau) = 2(T - \tau) + \beta.$$

The solution of the above ordinary differential equation with boundary value on final time is obtained using the Matlab in-built function *ode45* with step size 0.1. Function $f(t, \tau)$ appearing in (9) is then derived by setting $f(t, \tau) = -\frac{1}{2\alpha} p(t)$. As a result we have $\mathbf{u}_i[\tau](t) = -\frac{1}{2\alpha} p(t) x_i(t)$ for all $t \in [0, \tau]$.

For every player $i = 1, \ldots, n$, we then extract by brute force comparison, the strategy $\mathbf{u}_i[t_i^*]$ as in (10). Hence, we simulate the state evolution with $\mathbf{u}_i[t_i^*]$ and illustrate the results in Fig. 1. One can observe that for most of the players, especially those with a higher initial state, the switching time $t_i^*$ is around 15. Players usually stop before reaching zero as expected in consequence of the presence of a fixed cost $K$ in the cost function. A player with a state relatively close to zero at a time $t \approx T$ ($t$ is approaching the end of the horizon $T$) will be inactive to avoid paying the fixed cost.

**Algorithm**

**Input:** Simulations data
**Output:** best response strategies $\mathbf{u}_i[t_i^*]$ (10) and associated state distribution $dm(x,t)$.

1 : **Initialize** state $x(0) \leftarrow rand[0, 150]$,
2 : **for** $\tau = 1, 1.5, \ldots, T$ **do**
3 :   **for** player $i = 1, \ldots, n$ **do**
4 :     compute $\mathbf{u}_i[\tau]$ (9) and associated cost,
5 :   **end for**
6 : **end for**
7 : **for** player $i = 1, \ldots, n$ **do**
8 :   extract $\mathbf{u}_i[t_i^*]$ as in (10);
    simulate state evolution with $\mathbf{u}_i[t_i^*]$;
    compute distribution $dm(x,t)$.
9 : **end for**

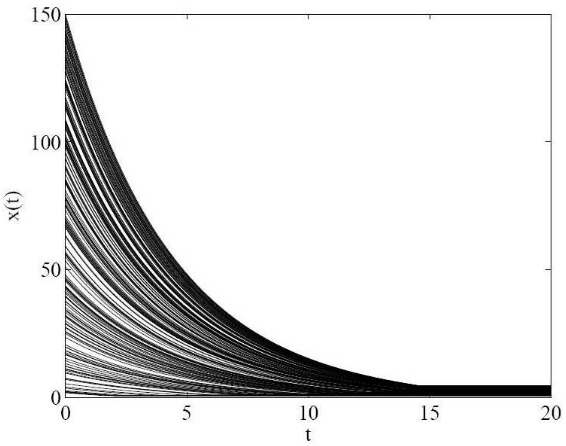


**Fig. 1** Time plot of state $x(t)$ with best response strategies $\mathbf{u}_i[t_i^*]$ as in (10).

## 7 Conclusions

Inspired by joint replenishment inventory systems, we have introduced linear quadratic games with set up costs monotonic on the number of active players, namely, players whose action is non-null. We have first analyzed the properties of the best response strategies and Nash equilibria for the given game. The obtained results are extended to the same game under large population with or without an additive disturbance.

## References

1. Başar, T., Bernhard, P.: $H^\infty$-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach. Birkhäuser (1991)

2. Bagagiolo, F., Bauso, D.: Objective function design for robust optimality of linear control under state-constraints and uncertainty. ESAIM: Control, Optimisation and Calculus of Variations **17**, 155–177 (2011)
3. Bardi, M.: Explicit solutions of some linear-quadratic mean field games. submitted (2012). Available at http://hal.inria.fr/docs/00/66/44/42/PDF/Bardi.pdf
4. Bauso, D., Giarrè, L., Pesenti, R.: Consensus in noncooperative dynamic games: A multiretailer inventory application. Transactions on Automatic Control **53**(4), 998–1003 (2008)
5. Bauso, D., Tembine, H., Başar, T.: Robust mean field games with application to production of an exhaustible resource. In: Proceedings of 7th IFAC Symposium on Robust Control Design. Aalborg, Danemark (2012)
6. Bauso, D., Zhu, Q., Başar, T.: Mixed integer optimal compensation: Decompositions and mean-field approximations. In: Proceedings of 2012 American Control Conference, Montreal. Montreal, Canada (2012)
7. Elliot, N.J., Kalton, N.: The existence of value in differential games of pursuit and evasion. Journal of Differential Equations **12**, 504–523 (1972)
8. Gueant, O., Lasry, J.M., Lions, P.L.: Mean field games and applications. Paris-Princeton Lectures (2010)
9. Huang, M., Caines, P., Malhamé, R.: Population cost-coupled lqg problems with non-uniform agents: Individual-mass behaviour and decentralized $\epsilon$ - nash equilibria. Transactions on Automatic Control **52**(9), 1560–1571 (2007)
10. Lachapelle, A., Salomon, J., Turinici, G.: Computation of mean field equilibria in economics. Mathematical Models and Methods in Applied Science **20**, 1–22 (2010)
11. Lasry, J.M., Lions, P.L.: Mean field games. Japanese Journal of Mathematics **2** (2007)
12. Milchtaich, I.: Congestion games with player-specific payoff functions. Games & Economic Behavior **13**, 111–124 (1996)
13. Pesenti, R., Bauso, D.: Mean field linear quadratic games with set up costs. In: Proceedings of NetGCoop 2011. Paris, France (2011)
14. Roxin, E.: The axiomatic approach in differential games. Journal of Optimization Theory and Applications **3**, 153–163 (1969)
15. Soulaimani, A.S., Quincampoix, M., Sorin, S.: Approchability theory, discriminating domain and differential games. SIAM Journal of Control and Optimization **48**(4), 2461–2479 (2009)
16. Tembine, H., Zhu, Q., Basar, T.: Risk-sensitive mean field stochastic games. In: Proceedings of IFAC World Congress. Milan, Italy (2011)
17. Varaiya, P.: The existence of solution to a differential game. SIAM Journal of Control and Optimization **5**, 153–162 (1967)
18. Watts, A.: Uniqueness of equilibrium in cost sharing games. Journal of Mathematical Economics **37**, 47–70 (2002)