Robust Dynamic Programming in N Players Uncertain Differential Games

Manuel JIMÉNEZ-LIZÁRRAGA[1,*], Sara V. RODRÍGUEZ-SÁNCHEZ[2],
Naín DE LA CRUZ[1], César Emilio VILLARREAL[2]

[1] *Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Nuevo León (UANL),
San Nicolás de los Garza, Nuevo León, México*

[2] *Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León (UANL),
San Nicolás de los Garza, Nuevo León, México*

*e-mail: manuel.jimenezlzr@uanl.edu.mx, sara.rodriguezsn@uanl.edu.mx,
nca200881@gmail.com, cesar.villarrealrd@uanl.edu.mx*

Abstract. In this paper we consider a non-cooperative N players differential game affected by deterministic uncertainties. Sufficient conditions for the existence of a robust feedback Nash equilibrium are presented in a set of min-max forms of Hamilton–Jacobi–Bellman equations. Such conditions are then used to find the robust Nash controls for a linear affine quadratic game affected by a square integrable uncertainty, which is seen as a malicious fictitious player trying to maximize the cost function of each player. The approach allows us to find robust strategies in the solution of a group of coupled Riccati differential equation. The finite, as well as infinite, time horizon cases are solved for this last game. As an illustration of the approach, the problem of the coordination of a two-echelon supply chain with seasonal uncertain fluctuations in demand is developed.

Key words: differential games, LQ games, dynamic supply chain.

1. Introduction

Differential games stand as a suitable framework for modelling strategic interaction between different agents (known as players), where each of them is looking for the minimization or, equivalently, the maximization of his individual criterion (Engwerda, 2005; Başar and Olsder, 1999). In such a multi-player scenario, none of the players is allowed to maximize his profits or objectives at the expense of the rest of the players. Therefore, the solution of the game is given in a form of *"equilibrium of forces"*.

Among different types of solutions, the so called Nash equilibrium is the most extensively used in the game theory literature. In this solution none of the players can improve their criteria by unilaterally deviating from their Nash strategy; therefore, no player has an incentive to change his decision. When the full state information is available to all the players to realize their decision strategy in each point of time, this is called a *feedback*

---

*Corresponding author.

Nash equilibrium (Engwerda, 2005; Başar and Olsder, 1999; Friedman, 1971). In order to find such feedback strategies, the optimal control tools are applied, specifically an equivalent N players form of the Hamilton–Jacobi–Bellman equation is required to be solved for each of the players. In the case of the non-cooperative Nash equilibrium solution framework, each player deals with a single criterion optimization problem (the standard optimal control problem), with the actions of the remaining players taking fixed equilibrium values.

Although the notion of robustness is such an important feature in the control theory, there are not many studies of dynamic games that are affected by some sort of uncertainties or disturbances. Some recent developments on this topic can be mentioned. Jiménez-Lizárraga and Poznyak (2007) presented a notion of open loop Nash equilibrium (OLNE) where the parameters of the game are within a finite set and the solution is given in terms of the worst-case scenario, that is, the result of the application of certain control input (in terms of the cost function value) is associated with the worst or least favourable value of the unknown parameter. The article of Jank and Kun (2002) shows also an OLNE and derives conditions for the existence and uniqueness of a worst case Nash equilibrium (WCNE); however, in this case they considered that the uncertainty belongs to a Hilbert functional space and enters adding up into the time derivative of the state variables. A similar problem is considered in a quite recent work (Engwerda, 2017), where the author shows that the WCNE can be derived by finding an ONLE of an associated differential game with $2N$ initial state constraints, the author derives necessary and sufficient conditions for the solution of the finite time problem. The work of Jungers *et al.* (2008) deals with a game with polytopic uncertainties that reformulated the problem as a nonconvex coupling between semi-definite programming to find the Nash type controls. Other related approaches include: using the Nash strategy to design robust controls for linear systems (Chen and Zhou, 2001). Another way to deal with uncertainties is to view them as an exogenous input (a fictitious player) (Chen *et al.*, 1997). In the work of van den Broek *et al.* (2003) the definition of equilibria is extended to deal with two cases: a soft-constrained formulation whose basis is given by Jank and Kun (2002), where the fictitious player is introduced in the criteria via a weighting matrix.

In this work, inspired in the works of Jank and Kun (2002) and Engwerda (2005, 2017), we analyse a deterministic N-player non-zero sum Differential Game case, considering finite, as well as infinite, time horizon in the performance index and a L^2 perturbation which is considered as a fictitious player trying to maximize the cost of each i-th player.

Assuming the player has access to the full state information, we are interested in finding a type of robust feedback Nash strategies, that guarantee a robust equilibrium when the players consider the worst case of the perturbation with respect to their own point of view. To that end, a set of robust form of the HJB equations are introduced; each of these equations compute not only the minimum of the i-th player control; but the maximum or worst case uncertainty from his point of view; resulting in a min-max form of the known HJB equations for a N players game. To the best of the authors' knowledge, using such a robust HJB equation has not been considered before to find a robust feedback Nash equilibrium in linear quadratic deterministic games, which stand as an important case to study. To summarize, the contributions of this work are as follows:

1. Presentation of the general conditions of robust worst case feedback Nash equilibrium by means of a robust form of the HJB equation for N players non-zero sum games.
2. Based on such a formulation, it gives the solution for the finite time horizon for the linear affine quadratic uncertain game.
3. It gives the solution of the infinite horizon for the linear affine dynamics.
4. It illustrates the result solving a problem of coordination of a two-echelon supply chain with seasonal uncertain fluctuations in demand. Such a case has not been treated before.

The development of this paper is as follows. In Section 2 we state, formally, the general problem of a differential game and the conditions for the Robust Nash Equilibrium to exist. Then, in Section 3 we define the dynamics of the problem analysed and the type of functional cost we have to minimize for a finite time horizon problem, we also state a theorem based on dynamic programming to find the robust controls for each player. In Section 4 we analyse the case of infinite time horizon. Finally, Section 5 follows with a numerical example. The purpose of this last section is to show how to apply the formulas obtained in Sections 3–4 and then compare our results against a finite time differential game which does not consider perturbation in the solution of the problem, which is the common problem treated, but the system itself is affected by some sort of perturbation.

2. Problem Statement

In this section we exploit the principle of dynamic programming in order to find all the robust feedback Nash equilibrium strategies for each player of a Non-zero sum uncertain differential game. We begin by presenting the general sufficient conditions for such a robust equilibrium to exist. Towards that end, consider the following N-person uncertain differential game with initial pair $(s, y) \in [0, T] \times \mathbb{R}^{n \times 1}$ described by the following initial value problem

$$\dot{x}(t) = f\left(x(t), \left(u_1(t), u_2(t), \ldots, u_N(t)\right), w(t), t\right), \quad x(s) = y,$$
$$\text{a.e. } t \in [s, T], \quad T < +\infty, \tag{1}$$

where $x(t) \in \mathbb{R}^{n \times 1}$ is the state column vector of the game and $u_i(t) \in \mathbb{R}^{l_i \times 1}$ is the control strategy at time t for each player i that may run over a given control region $U_i \subset \mathbb{R}^{l_i \times 1}$, i represents the number of player for $i \in \{1, \ldots, N\}$, $u_{\hat{i}}$ is the vector of strategies for the rest of the players, $\hat{i}$ is the *counter-coalition* of players counteracting the player with index i and $w(t) \in \mathbb{R}^{q \times 1}$ is a finite unknown disturbance in the sense that $\int_0^T \|w(t)\|^2 dt < +\infty$, that is, w is square integrable or, stated another way, $w \in L^2[0, T]$. The cost function as individual aim performance

$$J_i(s, y, u_i, u_{\hat{i}}, w) := \int_s^T g_i\left(x(t), u_i(t), u_{\hat{i}}(t), w(t), t\right) dt + h_i\left(x(T)\right), \tag{2}$$

which contains the integral term as well as a terminal state is given in the standard Bolza form.

Throughout the article we shall use the next notations:

- ^AB is the set of functions from the set A to the set B.
- A^t is the transpose of the matrix A.
- $I_{N,i} := \{k \in \mathbb{N} : 1 \leqslant k \leqslant N \text{ and } k \neq i\}$.
- $\mathscr{U}^i_{\text{adm}}[s_0, s_1] := \{u_i \in {}^{[s_0,s_1]}U_i : u_i \text{ is measurable}\}$.
- $\mathscr{U}^i_{\text{adm}} := \mathscr{U}^i_{\text{adm}}[0, T]$ is the set of all admissible control strategies.
- $\mathscr{U}^{\hat{i}}_{\text{adm}} := \mathscr{U}^1_{\text{adm}} \times \cdots \times \mathscr{U}^{i-1}_{\text{adm}} \times \mathscr{U}^{i+1}_{\text{adm}} \times \cdots \times \mathscr{U}^N_{\text{adm}}$.
- $\mathscr{U}_{\text{adm}} := \times_{i=1}^{N} \mathscr{U}^i_{\text{adm}}$.
- If $u \in \mathscr{U}_{\text{adm}}$, for $t \in [0, T]$, $u(t) := (u_1(t), u_2(t), \ldots, u_N(t))$.
- $D_i f$ denotes the partial derivative of f with respect to the i-th component.
- $\mathbb{1}_A$ denotes the indicator function of a set A.

Hypothesis 1. *The control region U_i is a subset of $\mathbb{R}^{l_i \times 1}$. The maps f, g_i and h_i are such that for all $(u_i, u_{\hat{i}}, w) \in \mathscr{U}^i_{\text{adm}} \times \mathscr{U}^{\hat{i}}_{\text{adm}} \times L^2[0, T]$, equation (1) admits an a.e. unique solution and the function J_i given in (2) is well defined; in general we assume the conditions given by Yong and Zhou (1999, p. 159).*

REMARK 1. We assume that the integrand g_i given in Equation (2) is positive definite, then the cost function J_i could not take negative values.

2.1. *Robust Feedback Nash Equilibrium*

Next, we introduce the worst case uncertainty from the point of view of the i-th player according to the complete set of controls u_j, with $j \in \{1, \ldots, N\}$ (Jank and Kun, 2002; Engwerda, 2017):

$$J_i(s, y, u_i, u_{\hat{i}}, w^*_{i,u_i,u_{\hat{i}}}) := \max_{w \in L^2[s,T]} J_i(s, y, u_i, u_{\hat{i}}, w). \tag{3}$$

In this paper we want to extend the robust Nash equilibrium notion, previously introduced by Jank and Kun (2002) for an open loop information structure, to a full state feedback information for an N players game.

DEFINITION 1. The control strategies $u^{\text{rn}}_1, u^{\text{rn}}_2, \ldots, u^{\text{rn}}_N$, are said to be *robust feedback Nash equilibrium*, where $(u^{\text{rn}}_i,)^N_{i=1} \in \mathscr{U}_{\text{adm}}$, if for any vector of admissible strategies

$$(u_i, u_{\hat{i}}) \in \mathscr{U}^i_{\text{adm}} \times \mathscr{U}^{\hat{i}}_{\text{adm}}, \quad \text{for } i \in \{1, \ldots, N\},$$

assuming the existence of the corresponding maximizing uncertainty function $w^*_{i,u_i,u^{\text{rn}}_{\hat{i}}} \in L^2[0, T]$ from the point of view of the i-th player. Then, we have the next set of inequalities:

$$J_i(s, y, u^{\text{rn}}_i, u^{\text{rn}}_{\hat{i}}, w^*_{i,u^{\text{rn}}_i,u^{\text{rn}}_{\hat{i}}}) \leqslant J_i(s, y, u_i, u^{\text{rn}}_{\hat{i}}, w^*_{i,u_i,u^{\text{rn}}_{\hat{i}}}). \tag{4}$$

In those conditions, we say also that $(u_1^{\text{rn}}, u_2^{\text{rn}}, \ldots, u_N^{\text{rn}})$ is a *vector of robust feedback Nash strategies for the whole set of players.*

Hypothesis 2. *There is a unique vector of robust feedback Nash strategies for the whole set of players.*

Now in order to find the robust feedback Nash equilibrium control strategies for the problem given by (2) subject to (1), we consider the following definition.

DEFINITION 2. Consider the *N*-tuples of strategies $(u_1, u_2, \ldots, u_N)$ and the *robust value function* for the *i*-th player as:

$$
\begin{aligned}
V_i(s, y) &:= \min_{u_i \in U_i} J_i\left(s, y, u_i, u_{\hat{i}}^{\text{rn}}, w_{i,u_i,u_{\hat{i}}^{\text{rn}}}^*\right), \quad \text{for } i \in \{1, 2, \ldots, N\}, \\
V_i\left(T, x(T)\right) &:= h_i\left(x(T)\right),
\end{aligned}
\tag{5}
$$

for any particular initial pair $(s, y) \in [0, T) \times \mathbb{R}^{n \times 1}$. The function V_i is also called the *robust Bellman function.*

REMARK 2. Notice that the minimization operation over u_i considers that the rest of the players are fixed in their Robust strategies (4) and each $w_{i,u_i,u_{\hat{i}}}^*$ satisfies (3).

2.2. Robust Dynamic Programming Equation

Let us explore the Bellman principle of optimality (Poznyak, 2008) for the robust value function V_i associated with the min-max posed problem for the *i*-th player, considering the rest of the participants as well as the signal function w fixed.

For $u_i \in \mathscr{U}_{\text{adm}}^i$, let us take $v_i = \mathbb{1}_{[s,\hat{s})} u_i + \mathbb{1}_{[\hat{s},T)} u_i^{\text{rn}}$ and note that $v_i \in \mathscr{U}_{\text{adm}}^i$. Using the Bellman principle of optimality for the functional $J_i(s, y, v_i, u_{\hat{i}}^{\text{rn}}, \cdot)$, where J_i is given in equation (2), and using also equation (5) given in Definition 2 we have:

$$
\begin{aligned}
V_i(s, y) &\leqslant J_i\left(s, y, v_i, u_{\hat{i}}^{\text{rn}}, w_{i,v_i,u_{\hat{i}}^{\text{rn}}}^*\right) = \max_{w \in L^2[s,T]} J_i\left(s, y, v_i, u_{\hat{i}}^{\text{rn}}, w\right) \\
&= \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\left(x(t), u_i(t), u_{\hat{i}}^{\text{rn}}(t), w(t), t\right) \mathrm{d}t \right. \\
&\qquad \left. + \int_{\hat{s}}^T g_i\left(x(t), u_i^{\text{rn}}(t), u_{\hat{i}}^{\text{rn}}(t), w(t), t\right) \mathrm{d}t + h_i\left(x(T)\right) \right\} \\
&= \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\left(x(t), u_i(t), u_{\hat{i}}^{\text{rn}}(t), w(t), t\right) \mathrm{d}t + V_i\left(\hat{s}, x(\hat{s})\right) \right\}, \quad (6)
\end{aligned}
$$

where the control strategies $u_{\hat{i}}^{\text{rn}}$ are robust Nash controls defined in (4) and $x(\hat{s})$ is such that x fulfills (1) when $u_j = u_j^{\text{rn}}$ for $j \neq i$ and $w = w_{i,u_i,u_{\hat{i}}^{\text{rn}}}^*$ that is described in the

Definition 1. Hence, taking the minimum in the right part of (6) over u_i, the inequality yields to

$$rlV_i(s,y) \leqslant \min_{u_i \in \mathscr{U}_{\mathrm{adm}}^i[s,T]} \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t \right.$$
$$\left. + V_i\big(\hat{s}, x(\hat{s})\big) \right\}. \tag{7}$$

On the other hand, beside for any $\delta > 0$, there is a control $u_{i,\delta} \in \mathscr{U}_{\mathrm{adm}}^i$, with the property:

$$V_i(s,y) + \delta \geqslant \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x_\delta(t), u_{i,\delta}(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t + V_i\big(\hat{s}, x_\delta(\hat{s})\big) \right\}, \tag{8}$$

where x_δ is the solution of (1) under the application of the control $u_{i,\delta}$ keeping the rest of the players fixed. Indeed, if there is a $\delta > 0$ such that for any $u_i \in \mathscr{U}_{\mathrm{adm}}^i$ we have

$$V_i(s,y) + \delta < \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t + V_i\big(\hat{s}, x(\hat{s})\big) \right\},$$

then, taking $u_i = u_i^{\mathrm{rn}}$ and using the Bellman principle of optimality, we would obtain

$$V_i(s,y) + \delta < \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i^{\mathrm{rn}}(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t + V_i\big(\hat{s}, x(\hat{s})\big) \right\}$$
$$\leqslant \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i^{\mathrm{rn}}(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t \right.$$
$$\left. + \int_{\hat{s}}^T g_i\big(x(t), u_i^{\mathrm{rn}}(t), u_{\hat{i}}^{\mathrm{rn}}(t), w_{i,u_{\hat{i}}^{\mathrm{rn}}, u_{\hat{i}}^{\mathrm{rn}}}^*(t), t\big) \mathrm{d}t + h_i\big(x(T)\big) \right\}$$
$$= V_i(s,y),$$

arriving to a contradiction. So, from the inequality (8) we get

$$V_i(s,y) + \delta \geqslant \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x_\delta(t), u_{i,\delta}(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t + V_i\big(\hat{s}, x_\delta(\hat{s})\big) \right\}$$
$$\geqslant \min_{u_i \in \mathscr{U}_{\mathrm{adm}}^i[s,T]} \max_{w \in L^2[s,T]} \left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i(t), u_{\hat{i}}^{\mathrm{rn}}(t), w(t), t\big) \mathrm{d}t \right.$$
$$\left. + V_i\big(\hat{s}, x(\hat{s})\big) \right\}. \tag{9}$$

Now, as in the inequality (9), the value of δ is positive, but arbitrary, we have

$$
\begin{aligned}
V_i(s, y) \geqslant \min_{u_i \in \mathscr{U}_{\text{adm}}^i[s,T]} \max_{w \in L^2[s,T]} &\left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i(t), u_i^{\text{rn}}(t), w(t), t\big) \mathrm{d}t \right. \\
&\left. + V_i\big(\hat{s}, x(\hat{s})\big) \right\}
\end{aligned}
\tag{10}
$$

(Fattorini, 1999; Poznyak, 2008).

From the inequalities (7) and (10), we have arrived to the next theorem that is a robust form of the dynamic programming equation, for the problem in consideration.

Theorem 1. *Let the basic assumption of Section 2 hold, then for any initial pair* $(s, y) \in [0, T) \times \mathbb{R}^{n \times 1}$, *the following relationship holds*:

$$
\begin{aligned}
V_i(s, y) = \min_{u_i \in \mathscr{U}_{\text{adm}}^i[s,T]} \max_{w \in L^2[s,T]} &\left\{ \int_s^{\hat{s}} g_i\big(x(t), u_i(t), u_i^{\text{rn}}(t), w(t), t\big) \mathrm{d}t \right. \\
&\left. + V_i\big(\hat{s}, x(\hat{s})\big) \right\},
\end{aligned}
\tag{11}
$$

for all $\hat{s} \in [s, T]$.

The development of the principle of optimality to equation (11), leads immediately to the following result:

Theorem 2. *Let's consider the uncertain affine N-players differential game given by (1)–(2), where T is finite and the full state information is known. In this case the vector of control strategies* $(u_i^{\text{rn}}, u_i^{\text{rn}})$ *provides a robust feedback equilibrium if there exists a continuously differentiable function* $V_i : [0, T] \times \mathbb{R}^{n \times 1} \to \mathbb{R}$ *satisfying the following partial differential equation*:

$$
\begin{aligned}
-\mathrm{D}_1 V_i\big(t, x(t)\big) = \min_{u_i \in \mathscr{U}_{\text{adm}}^i} \max_{w \in L^2[s,T]} &\left\{ \mathrm{D}_2 V_i\big(t, x(t)\big)^{\mathrm{t}} f\big(x(t), \hat{u}_i(t), w(t), t\big) \right. \\
&\left. + g_i\big(x(t), u_i(t), u_i^{\text{rn}}(t), w(t), t\big) \right\}; \\
V_i\big(T, x(T)\big) = h_i\big(x(T)\big), \quad &\text{for } i \in \{1, \dots, N\},
\end{aligned}
\tag{12}
$$

where $\hat{u}_i(t) = (u_1^{\text{rn}}(t), \dots, u_{i-1}^{\text{rn}}(t), u_i(t), u_{i+1}^{\text{rn}}(t), \dots, u_N^{\text{rn}}(t))$, *and the corresponding min-max cost for each player is*

$$
J_i^* = V_i(s, y).
\tag{13}
$$

REMARK 3. The partial differential equation (12) of Theorem 2 is called the *robust Hamilton–Jacobi–Bellman (RHJB) equation*. In previous important works dealing with

the design of robust H_∞ controllers using a dynamic game approach (Başar and Bernhard, 2008; Aliyu, 2011) the min-max version of the value function was already found. It was also found that actually when all the players are fixed in their robust Nash controls the game became a zero-sum game, played out between the i-th player and the uncertainty. Equation (12) is an extension to the N players non-zero sum game; however, to the best of our knowledge, this case has not been introduced yet.

3. Finite Time Horizon N Players Linear Affine Quadratic Differential Game

Once the general conditions for the existence of a robust feedback Nash equilibrium in an uncertain differential game are established, we turn now to the special case of the linear affine quadratic differential games (LAQDG). In this section we consider the case where the time horizon is finite, that is, $T < +\infty$. The game is played by N participants which are trying to minimize certain loss inflicted by a disturbance, besides, the functional cost of the game is restricted by the corresponding differential equation. Therefore, in this section we assume that:

$$f\big(x(t), u(t), w(t), t\big) := A(t)x(t) + \sum_{j=1}^{N} B_j(t)u_j(t) + E(t)\, w(t) + c(t),$$

$$x(t) \in \mathbb{R}^{n \times 1}, \qquad x(0) = x_0, \qquad u_j(t) \in \mathbb{R}^{l_j \times 1}, \quad 0 \leqslant t \leqslant T < +\infty \tag{14}$$

the functions for the cost of each player are given by the following quadratic functions:

$$g_i\big(x(t), u_i(t), u_{\hat{\imath}}(t), w(t), t\big) := x(t)^{\mathrm{t}} Q_i(t) x(t)$$

$$+ \sum_{j=1}^{N} u_j(t)^{\mathrm{t}} R_{i,j}(t) u_j(t) - w(t)^{\mathrm{t}} W_i(t) w(t), \tag{15}$$

$$h_i\big(x(T)\big) := x(T)^{\mathrm{t}} Q_{i\mathrm{f}} x(T),$$

where j represents the number of the player, $A(t) \in \mathbb{R}^{n \times n}$, and $B_j(t) \in \mathbb{R}^{n \times l_j}$, for $j \in \{1, \ldots, N\}$, are the known system and controls matrices; $x(t)$ is the state vector of the game and u_j is the control strategy for the j-th player; $c(t) \in \mathbb{R}^{n \times 1}$ is an exogenous and known signal. In this case w is the same as in (1), that is, a finite disturbance entering the system through the matrix $E(t) \in \mathbb{R}^{n \times q}$. The performance index for each i-th player is given again in standard Bolza form, the strategy for the player i is u_i while $u_{\hat{\imath}}$ are the strategies of the rest of the players. The term $w(t)^{\mathrm{t}} W_i(t) w(t)$ is the unknown uncertainty, which is trying to maximize the cost J_i from the point of view of the i-th player. The cost matrices are assumed to satisfy: $Q_i(t) = Q_i(t)^{\mathrm{t}} \geqslant \mathbf{0}$, $Q_{i\mathrm{f}} = Q_{i\mathrm{f}}^{\mathrm{t}} \geqslant \mathbf{0}$ and $W_i(t) = W_i(t)^{\mathrm{t}} > \mathbf{0}$ (symmetric and semipositive/positive definite matrices); $R_{i,i}(t) = R_{i,i}(t)^{\mathrm{t}} > \mathbf{0}$ and $R_{i,j}(t) = R_{i,j}(t)^{\mathrm{t}} \geqslant \mathbf{0}$, where inequalities mean inequalities component by component. Assume also that the players have access to the full state information pattern, that is, they

measure $x(t)$, for all $t \in [0, T]$. All the involved squared matrices are assumed to be non-singular.

For the linear affine dynamics given in (14), equation (12) can be rewritten as follows:

$$
\begin{aligned}
-D_1 V_i (t, x(t)) = \min_{u_i \in \mathscr{U}_{\text{adm}}^i} \max_{w \in L^2[0,T]} & \left\{ D_2 V_i (t, x(t))^{\text{t}} \left(A(t)x(t) + B_i(t)u_i(t) \right. \right. \\
& \left. + \sum_{j \in I_{N,i}} B_j(t)u_j^{\text{rn}}(t) + E(t)w(t) + c(t) \right) \\
& + x(t)^{\text{t}} Q_i(t)x(t) + u_i(t)^{\text{t}} R_{i,i}(t)u_i(t) \\
& \left. + \sum_{j \in I_{N,i}} u_j^{\text{rn}}(t)^{\text{t}} R_{i,j}(t)u_j^{\text{rn}}(t) - w(t)^{\text{t}} W_i(t)w(t) \right\}
\end{aligned}
\tag{16}
$$

with terminal condition as $V_i(T, x(T)) = x(T)^{\text{t}} Q_{if} x(T)$. With this condition and if the assumptions mentioned above are satisfied, the Robust Feedback Nash Equilibrium can be directly obtained as

$$
\begin{aligned}
u_i^{\text{rn}} = \arg \min_{u_i \in \mathscr{U}_{\text{adm}}^i} & \left\{ D_2 V_i (\cdot, x(\cdot))^{\text{t}} \left(Ax + B_i u_i + \sum_{j \in I_{N,i}} B_j u_j + E w_{i,u_i,u_{\hat{i}}^{\text{rn}}}^{*} + c \right) \right. \\
& \left. + x^{\text{t}} Q_i x + u_i^{\text{t}} R_{i,i} u_i + \sum_{j \in I_{N,i}} u_j^{\text{rnt}} R_{i,j} u_j^{\text{rn}} - w_{i,u_i,u_{\hat{i}}^{\text{rn}}}^{*\,\text{t}} W_i w_{i,u_i,u_{\hat{i}}^{\text{rn}}}^{*} \right\},
\end{aligned}
\tag{17}
$$

and worst case uncertainty from the point of view of the i-th player is obtained as

$$
\begin{aligned}
w_{i,u_i,u_{\hat{i}}}^{*} = \arg \max_{w \in L^2[0,T]} & \left\{ D_2 V_i (\cdot, x(\cdot))^{\text{t}} \left(Ax + \sum_{j=1}^{N} B_j u_j + E w + c \right) \right. \\
& \left. + x^{\text{t}} Q_i x + \sum_{j=1}^{N} u_j^{\text{t}} R_{i,j} u_j - w^{\text{t}} W_i w \right\}.
\end{aligned}
\tag{18}
$$

REMARK 4. Notice that the value of $w_{i,u_i,u_{\hat{i}}}^{*}$ given in (18) does not depend of $(u_i, u_{\hat{i}})$. So, in this particular case, we shall denote such value just by w_i^{*}.

Theorem 3. *The robust feedback Nash strategies for the uncertain LQ affine game (14)–(15), has the next linear form:*

$$
u_i^{\text{rn}} = -R_{i,i}^{-1} B_i^{\text{t}} (P_i x + m_i)
\tag{19}
$$

and the worst case uncertainty from the point of view of the i-th player is:

$$
w_i^{*} = W_i^{-1} E_i^{\text{t}} (P_i x + m_i),
\tag{20}
$$

where the set of N Riccati type coupled equations P_i satisfy the following boundary value problem:

$$-\dot{P}_i = \tilde{A}^t P_i + P_i \tilde{A} + Q - P_i S_i^t P_i + P_i M_i P_i + \sum_{j \in I_{N,i}} P_j S_{j,i} P_j;$$

$$P_i(T) = Q_{if},$$

$$\tilde{A} := A - \sum_{j=1}^{N} S_j P_j, \tag{21}$$

where $S_i := B_i R_{i,i}^{-1} B_i^t$, $S_{j,i} := B_j R_{j,j}^{-1} R_{i,j} R_{j,j}^{-1} B_j^t$, $M_i := E W_i^{-1} E^t$, for $i, j \in \{1, \ldots, N\}$; and the m_i are the "shifting vectors" governed by the following coupled linear differential equations:

$$-\dot{m}_i = A^t m_i - P_i S_i m_i - \sum_{j \in I_{N,i}} P_i S_j m_j - \sum_{j \in I_{N,i}} P_j S_j m_i$$

$$+ P_i M_i m_i + P_i c + \sum_{j \in I_{N,i}} P_j S_{j,i} m_j; \quad m_i(T) = 0, \tag{22}$$

and the value of the robust Nash cost is:

$$J_i^* = x(0)^t P_i x(0) + 2 m_i(0)^t x(0) + \int_0^T \left(m_i^t S_i m_i - 2 \sum_{j \in I_{N,i}} m_i^t S_j m_j + m_i^t M_i m_i \right.$$

$$\left. + \sum_{j \in I_{N,i}} m_j^t P_j S_{j,i} P_j m_j + 2 m_i^t c \right), \tag{23}$$

where J_i^* is the optimum value of (2).

The proof of this theorem is presented in Appendix A.

4. Infinite Time Horizon Case

In this section we consider the same linear affine quadratic game when the time horizon is infinite. As the case analysed in the last section, the players are trying to minimize certain loss inflicted by a disturbance, besides, the functional cost of the game is restricted by a differential equation which considers an affine term. In this type of game the functional cost is given by:

$$J_i(t, x, u_i, u_{\hat{\imath}}, w) = \int_0^{+\infty} \left(x^t Q_i x + \sum_{j=1}^{N} u_j^t R_{j,i} u_j - w^t W_i w \right), \tag{24}$$

and the constraint has the following form:

$$\dot{x}(t) = Ax(t) + \sum_{j=1}^{N} B_j u_j(t) + Ew(t) + c(t). \tag{25}$$

The involved matrices are constant with corresponding dimension and involved matrices in (24), satisfy equivalent restriction of the finite time counterpart. Following Engwerda (2005), we assume that $c \in L^2_{\text{exp, loc}}$, that is, locally square integrable and converging to zero exponentially. In this case, the system of algebraic Riccati equations takes on the form:

$$\tilde{A}^t P_i + P_i \tilde{A} + Q - P_i S_i^t P_i + P_i M_i P_i + \sum_{j \in I_{N,i}} P_j S_{j,i} P_j = 0, \tag{26}$$

where $\tilde{A} := A - \sum_{j=1}^{N} S_j P_j$.

To find the solution to this problem the completion of the square method is developed in Poznyak (2008) and the following theorem is stated.

Theorem 4. *For the differential game problem given by the equations* (24)–(25), *if the algebraic Riccati equations* (26) *possess symmetric stabilizing solutions* P_i, *then the infinite time horizon Robust Nash Equilibrium strategies are given by*

$$u_i(t) = -R_{i,i}^{-1} B_i^t (P_i x(t) + m_i(t)), \tag{27}$$

and the worst case will be given by

$$w_i(t) = T_i^{-1} E^t (P_i x(t) + m_i(t)), \tag{28}$$

where each m_i *fulfills the equation*

$$m_i(t) = \int_t^{+\infty} \left(e^{-(A_i - \sum_{j \neq i}^{N} (P_j S_{j,i} - P_i S_j))(t-s)} P_i c(s) \right) ds, \tag{29}$$

and

$$A_i = A^t - \sum_{j=1}^{N} P_j S_j + \sum_{j=1}^{N} P_j S_{j,i} + P_i M_i.$$

Moreover, the optimal value J_i^* *is given by*

$$J_i^* = x(0)^t P_i x(0) + 2m_i(0)^t x(0) + n_i(0) \tag{30}$$

and the closed loop states equation has the form

$$\dot{x}(t) = \left(A - \sum_{j=1}^{N} B_j R_{j,j}^{-1} B_j^t P_j \right) x(t) - \sum_{j=1}^{N} B_j R_{j,j}^{-1} B_j^t m_j(t) + Ew(t) + c(t). \tag{31}$$

The proof of this theorem is found in Appendix A.

5. Numerical Example: A Differential Game Model for a Vertical Marketing System with Demand Fluctuation and Seasonal Prices

Consider a noncooperative game in a two-echelon supply chain established between two chain agents (Dockner *et al.*, 2000; Jørgensen, 1986); a single supplier (called the manufacturer) and a single distributor (called the retailer). The manufacturer is in charge of selling a product type to a single retailer over a period of time T at the price $p_1(t)$. The retailer is in charge of distributing and marketing that product, at a price $p_2(t) = p_1(t) + r_2(t)$, where $r_2(t)$ represents the profit margin gained by the retailer at time t per each unit sold. In this case, let us set $r_2 = 0.2p_1$.

The dynamic of the game is established by both players searching for a Nash equilibrium in their coordination contract, and furthermore facing some source of uncertainties. For this particular case assume that the retailer deals with a demand that evolves exogenously over time, with the quantity sold per time unit, d, depending not only on price p_1, but also on the time t elapsed, $d = d(p, t)$. The exogenous change in demand presented here is due to seasonal fluctuations. Under such enviroment, the profit equations for each players are J_1 and J_2 with the following quadratic structure:

$$J_1 = F_{1f}x_1^2(26) + \int_0^{26} \left(\frac{c_1(t)}{2}u_1^2(t) + \frac{h_1}{2}x_1^2(t) \right) dt, \tag{32}$$

$$J_2 = F_{2f}x_2^2(26) + \int_0^{26} \left(-p_2w^2(t) + \left(\frac{p_1 + c_2(t)}{2} \right) u_2^2(t) + \frac{h_2}{2}x_2^2(t) \right) dt, \tag{33}$$

subject to the following dynamic

$$\dot{x}_1 = u_1 - u_2,$$
$$\dot{x}_2 = u_2 - d - ew, \tag{34}$$

where J_1 indicates the operating cost faced by the manufacturer given by the holding cost and the production cost, plus a small penalization of the inventories at the final time of the horizon. On the other hand, J_2 indicates the operating cost incurred in by the retailer obtained by the holding cost, the production cost (including the price paid to the manufacturer for the products), and the perturbation signal w seen as a malicious fictitious player, and a small penalization of the inventories at the final time of the horizon. This game involves the dynamic changes of the inventory for each player (x_1, x_2), with the production rate (u_1, u_2) as decision variables. Moreover, the retailer's dynamic faces an uncertain demand represented by two terms, the deterministic demand d plus the uncertain factor ew.
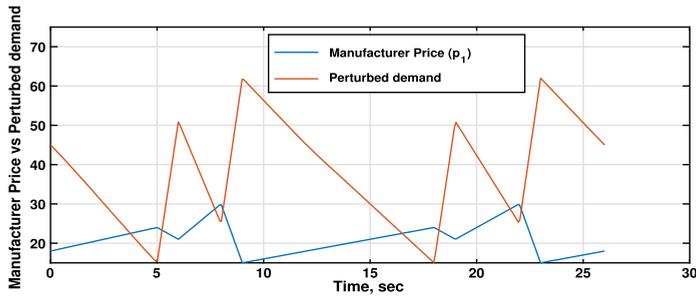
Fig. 1. Price vs Perturbed demand.

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 \\ -d \end{pmatrix},$$

$$F_{1f} = \begin{pmatrix} 8 & 0 \\ 0 & 0 \end{pmatrix}, \quad F_{2f} = \begin{pmatrix} 0 & 0 \\ 0 & 8 \end{pmatrix}, \quad E = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad R_{11} = \frac{c_1}{2},$$

$$R_{22} = \frac{p_1 + c_2}{2}, \quad Q_1 = \begin{pmatrix} \frac{h_1}{2} & 0 \\ 0 & 0 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{h_2}{2} \end{pmatrix},$$

where

$$c_1 = 0.85p_1, \quad p_2 = 2p_1, \quad c_2 = 2, \quad h_1 = 15, \quad h_2 = 10.5,$$

$$p_1(t) = \begin{cases} \frac{1}{5}(6t + 90) & \text{for } 0 \leqslant t \leqslant 5, \\ -3t + 39 & \text{for } 5 < t \leqslant 6, \\ \frac{9}{2}t - 6 & \text{for } 6 < t \leqslant 8, \\ -15t + 150 & \text{for } 8 < t \leqslant 9, \\ \frac{20}{3}t - 45 & \text{for } 9 < t \leqslant 12, \end{cases} \tag{35}$$

$$d(t) = \begin{cases} -5p_1(t) + 135 & \text{for } 0 \leqslant t < 5, \\ -12p_1(t) + 303 & \text{for } 5 < t \leqslant 6, \\ -\frac{26}{9}p_1(t) + \frac{1005}{9} & \text{for } 6 < t \leqslant 8, \\ -\frac{37}{15}p_1(t) + \frac{1485}{15} & \text{for } 8 < t \leqslant 9, \\ -5p_1(t) + 135 & \text{for } 9 < t \leqslant 12. \end{cases} \tag{36}$$

According to the game equations (1) and (2), $N = 2$. We used the Matlab software to solve numerically backward in time (21), thus obtaining the corresponding robust Nash equilibrium strategies for each player. The results of such numerical solution are shown in figures (Figs. 2, 3). In Fig. 1 are depicted the perturbed demand and the manufacturer price.
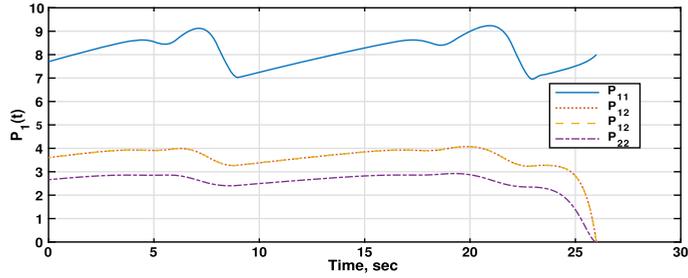
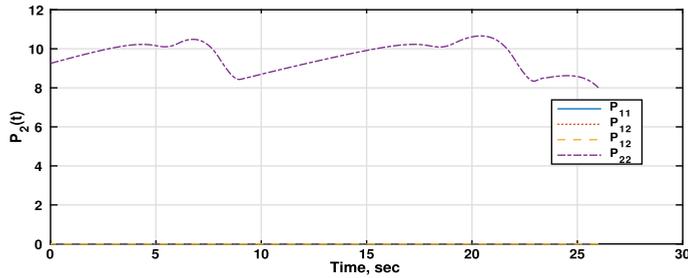Fig. 2. Riccati differential equation player 1 (manufacturer).

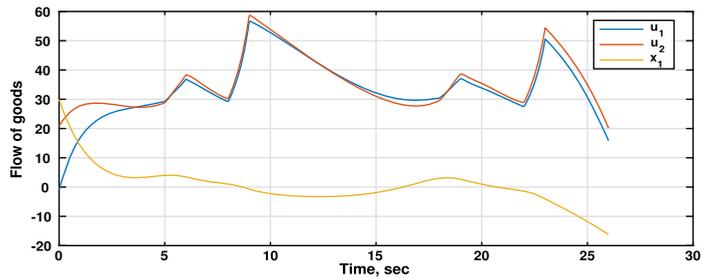Fig. 3. Riccati differential equation player 2 (retailer).

Fig. 4. Comparison between manufacturer produced units (u_1), units demanded by the retailer (u_2), and units left in the manufacturer stock x_1.

Figure 4 shows the behaviour of decision variables from each player, and the state equation from the manufacturer (u_1 production manufacturer's rate, u_2 purchasing retailer's rate, and manufacturer's inventory). Through this figure we can compare the different outputs, for instance, we observe that products left in the manufacturer's stock basically are close to zero. In fact, this figure shows the advantages of better coordination between the different chain agents, in order to reduce bullwip effect. Since, the manufacturer and the retailer share information about customer demand, the produced goods from the manufacturer and the purchased goods from the retailer are deriving in similar behaviour.
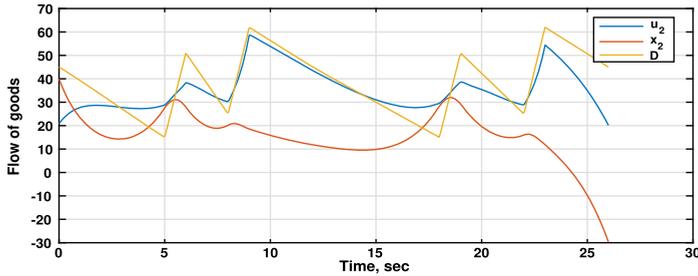
Fig. 5. Comparison between retailer bought units (u_2), units demanded by the final consumer (D), and units left in the retailer stock (x_2).

Also, since there are no restrictions for the states of a given stage in the chain we can see that, at times, we are going to have negative values for this variable. For example, between $t = 8$ and $t = 16$, the units left in stock get a negative value, this only means that the manufacturer has backlogged units. However, we can appreciate that the amount of this backlogged units is minimum. Also for the closing of the season, between $t = 20$ and $t = 25$ it is better for the manufacturer only to have backlogged units. Once we get a Nash equilibrium, any deviation from the output policies would result in a loss for the manufacturer or the retailer.

On the other hand, Fig. 5 shows the behaviour of the retailer's dynamics through the time horizon. We can appreciate that the strategy followed by the retailer differs from the manufacturer in that the retailer uses inventory to face demand uncertainties. The retailer is considering the worst case of any perturbation on demand, but stock units are kept up to the minimum. The decisions at the end of the planning horizon are perturbed by the finite time horizon condition. For that reason the planning horizon was extended to two years in order to avoid such perturbations in the first year.

6. Conclusions

We found the Nash equilibrium control function of an N players differential game affected by an L^2 uncertainty function for a linear quadratic affine performance function in two cases:

1. When we have a finite time horizon. In this case we assume the matrices involved in the performance function are time dependent.
2. When we have an infinite time horizon. In this case we assume the matrices involved in the performance function are constant with respect to the time, and there are only temporal dependence of the uncertainty function, the state function, and an affine term of the constraint given by the linear differential equation.

Both problems are solved in this work using different methods.

A. Appendix

Proof. Proof of Theorem 3. To start the proof we find u_i^{rn} by means of (17) and the fact that

$$\mathrm{D}_2 V_i\big(t, x(t)\big)\big(B_i(t)\big) + 2u_i(t)^{\mathrm{t}} R_{i,i}(t) = 0$$
$$\Rightarrow u_i(t) = -\frac{1}{2} R_{i,i}^{-1}(t) B_i(t)^{\mathrm{t}} \big(\mathrm{D}_2 V_i\big(t, x(t)\big)\big)^{\mathrm{t}}, \tag{37}$$

and we find w_i^*, given in Remark 4 by means of (18) and the fact that

$$\mathrm{D}_2 V_i\big(t, x(t)\big)\big(E(t)\big) + 2w(t)^{\mathrm{t}} W(t) = 0$$
$$\Rightarrow w(t) = \frac{1}{2} W_i^{-1}(t) E(t)^{\mathrm{t}} \big(\mathrm{D}_2 V_i(t, x)\big)^{\mathrm{t}}. \tag{38}$$

Substituting (37) and (38) back in (16) we get

$$
\begin{aligned}
-\mathrm{D}_1 V_i(\cdot, x) = \mathrm{D}_2 V_i(\cdot, x)\bigg(& Ax - \frac{1}{2} R_{i,i}^{-1} B_i^{\mathrm{t}} \big(\mathrm{D}_2 V_i(\cdot, x)\big)^{\mathrm{t}} \\
& - \frac{1}{2} \sum_{j \in I_{N,j}} B_j R_{j,j}^{-1} B_j^{\mathrm{t}} \big(\mathrm{D}_2 V_j(\cdot, x)\big)^{\mathrm{t}} \\
& - E W_i^{-1} E^{\mathrm{t}} \big(\mathrm{D}_2 V_i(\cdot, x)\big)^{\mathrm{t}} + c\bigg) \\
& + x^{\mathrm{t}} Q_i x + \frac{1}{4} \sum_{j=1}^{N} \mathrm{D}_2 V_j(\cdot, x) B_j R_{j,j}^{-1} R_{j,i} R_{j,j}^{-1} B_j^{\mathrm{t}} \big(\mathrm{D}_2 V_j(\cdot, x)\big)^{\mathrm{t}} \\
& + \frac{1}{2} \big(\mathrm{D}_2 V_i(\cdot, x)\big) E W_i^{-1} E^{\mathrm{t}} \big(\mathrm{D}_2 V_i(\cdot, x)\big)^{\mathrm{t}}.
\end{aligned}
\tag{39}
$$

Now we have to solve the first order partial differential equation (39); to do that we propose the solution

$$V_i\big(t, x(t)\big) = x(t)^{\mathrm{t}} P_i(t) x(t) + 2m_i(t)^{\mathrm{t}} x(t) + n_i(t). \tag{40}$$

So

$$\mathrm{D}_2 V_i\big(t, x(t)\big) = 2x(t)^{\mathrm{t}} P_i(t) + 2m_i(t)^{\mathrm{t}} \quad \text{and} \tag{41}$$
$$\mathrm{D}_1 V_i\big(t, x(t)\big) = x(t)^{\mathrm{t}} \dot{P}_i(t) x(t) + 2\dot{m}_i(t)^{\mathrm{t}} x(t) + \dot{n}_i(t). \tag{42}$$

Substituting (41) and (42) into (39), expanding, and grouping terms of the form $x^{\mathrm{t}}(Y)x$ and $x^{\mathrm{t}}(Z)$, yields

$$
x^{\mathrm{t}}\left(\dot{P}_i + \left(A - \sum_{j=1}^{N} B_j R_{j,j}^{-1} B_j^{\mathrm{t}} P_j\right)^{\mathrm{t}} P_i + P_i\left(A - \sum_{j=1}^{N} B_j R_{j,j}^{-1} B_j^{\mathrm{t}} P_j\right) + Q\right.
$$

$$
\left. - P_i B_i R_{i,i}^{-1} B_i^{\mathrm{t}} P_i + P_i E W_i^{-1} E^{\mathrm{t}} P_i + \sum_{j=1}^{N} P_j B_j R_{j,j}^{-1} R_{j,i} R_{j,j}^{-1} B_j^{\mathrm{t}} P_j\right) x
$$

$$
+ x^{\mathrm{t}}\left(\dot{m}_i + A^{\mathrm{t}} m_i - P_i B_i R_{i,i}^{-1} B_i^{\mathrm{t}} m_i - \sum_{j=1}^{N} P_i B_j R_{j,j}^{-1} B_j^{\mathrm{t}} m_j - \sum_{j=1}^{N} P_j B_j R_{j,j}^{-1} B_j^{\mathrm{t}} m_i\right.
$$

$$
\left. + P_i E W_i^{-1} E^{\mathrm{t}} m_i + P_i c + \sum_{j=1}^{N} P_j B_j R_{j,j}^{-1} R_{j,i} R_{j,j}^{-1} B_j^{\mathrm{t}} m_j\right)
$$

$$
+ \left(\dot{n}_i + m_i^{\mathrm{t}} B_i R_{i,i}^{-1} B_i^{\mathrm{t}} m_i - 2\sum_{j=1}^{N} m_i^{\mathrm{t}} B_j R_{j,j}^{-1} B_j^{\mathrm{t}} m_j + m_i^{\mathrm{t}} E W_i^{-1} E^{\mathrm{t}} m_i\right.
$$

$$
\left. + \sum_{j=1}^{N} m_j^{\mathrm{t}} P_j B_j R_{j,j}^{-1} R_{j,i} R_{j,j}^{-1} B_j^{\mathrm{t}} P_j m_j + 2 m_i^{\mathrm{t}} c\right) = 0. \tag{43}
$$

Last equation is only satisfied when (21) and (22) have solution; for the terminal conditions we have

$$
V_i(T, x) = x(T)^{\mathrm{t}} P_i(T) x(T) + 2 m_i(T)^{\mathrm{t}} x(T) + n_i(T) = x(T)^{\mathrm{t}} Q_{i\mathrm{f}} x(T),
$$

implying the terminal conditions of (21) and (22). The value of the optimal functional cost is

$$
J_i^* = x(0)^{\mathrm{t}} P_i(0) x(0) + 2 m_i(0)^{\mathrm{t}} x + n_i(0), \tag{44}
$$

so the theorem is proven. $\square$

Proof. Proof of Theorem 4. To develop the proof of this theorem let us suppose there exists an "energetic function" $V_i : [0, +\infty) \times \mathbb{R}^{n \times 1} \to \mathbb{R}$ with the form

$$
V_i(t, x) = x_i^{\mathrm{t}} P_i x_i + 2 m_i(t)^{\mathrm{t}} x_i + n_i(t),
$$

where m_i is defined in (29) and n_i is given by

$$
n_i(t) = \int_t^\infty \left(m_i^{\mathrm{t}} S_i\, m_i - 2\sum_{j=1}^{N} m_i^{\mathrm{t}} S_j\, m_j + m_i^{\mathrm{t}} M_i m_i + \sum_{j=1}^{N} m_j^{\mathrm{t}} P_j\, S_{j,i}\, P_j\, m_j + 2 m_i^{\mathrm{t}} c\right). \tag{45}
$$

Let $\hat{V}_i : [0, +\infty) \to \mathbb{R}$ the function given by

$$\hat{V}_i(t) := V_i\big(t, x(t)\big).$$

We start by taking the derivative of $\hat{V}_i$, obtaining

$$\hat{V}'_i = 2x^{\mathrm{t}} P_i \dot{x} + 2\dot{m}^{\mathrm{t}}_i x + 2m^{\mathrm{t}}_i \dot{x} + \dot{n}_i. \tag{46}$$

Now the fundamental calculus theorem tells us

$$\hat{V}_i(T) - \hat{V}_i(0) = \int_0^T \hat{V}'_i(t)\mathrm{d}t. \tag{47}$$

Substituting the value of $\dot{x}(t)$ given in (31) into (46) and so into (47) we get

$$
\begin{aligned}
&\hat{V}_i(T) - \hat{V}_i(0) \\
&= \int_0^T \Bigg(2x(t)^{\mathrm{t}} P_i \Bigg(Ax(t) + \sum_{j=1}^N B_j u_j(t) + Ew(t) + c(t) \Bigg) + 2\dot{m}_i(t)^{\mathrm{t}} x(t) \\
&\quad + 2m_i(t)^{\mathrm{t}} \Bigg(Ax(t) + \sum_{j=1}^N B_j u_j(t) + Ew(t) + c(t) \Bigg) - \dot{n}_i(t) \Bigg) \mathrm{d}t.
\end{aligned} \tag{48}
$$

Now we start adding and subtracting $x^{\mathrm{t}} Q_i x$, $u^{\mathrm{t}}_i R_{i,i} u_i$, $w^{\mathrm{t}} W_i w$, so we can express (48) as

$$
\begin{aligned}
&\hat{V}_i(T) - \hat{V}_i(0) \\
&= \int_0^T \Bigg(\Bigg(2x(t)^{\mathrm{t}} P_i Ax(t) + x^{\mathrm{t}} Q_i x + u_i(t)^{\mathrm{t}} R_{i,i} u_i(t) + 2x(t)^{\mathrm{t}} P_i \sum_{j=1}^N B_j u_j(t) \\
&\quad + 2x(t)^{\mathrm{t}} P_i c(t) + 2\dot{m}_i(t)^{\mathrm{t}} x(t) + 2m_i(t)^{\mathrm{t}} Ax(t) \\
&\quad + 2m_i(t)^{\mathrm{t}} \sum_{j=1}^N B_j u_j(t) + 2m_i(t)^{\mathrm{t}} c(t) + \dot{n}_i(t) \Bigg) \\
&\quad + \big(2x(t)^{\mathrm{t}} P_i Ew(t) + 2m_i(t)^{\mathrm{t}} Ew(t) - w(t)^{\mathrm{t}} W_i w(t) \big) \Bigg) \mathrm{d}t \\
&\quad - \int_0^T \big(x(t)^{\mathrm{t}} Q_i x(t) + u_i(t)^{\mathrm{t}} R_{i,i} u_i(t) - w(t)^{\mathrm{t}} W_i w(t) \big) \mathrm{d}t.
\end{aligned} \tag{49}
$$

The second brace can be expressed as a difference of squares as follows:

$$
\begin{aligned}
2\left(x^{\mathrm{t}} P_i E + m_i^{\mathrm{t}} E\right) w &- w^{\mathrm{t}} W_i w \\
&= 2\left(x^{\mathrm{t}} P_i E W^{-\frac{1}{2}} + m_i^{\mathrm{t}} E W^{-\frac{1}{2}}\right) W_i^{\frac{1}{2}} w - \left\| w^{\mathrm{t}} W_i^{\frac{1}{2}} \right\|^2 \\
&= -\left\| x^{\mathrm{t}} P_i E W^{-\frac{1}{2}} + m_i^{\mathrm{t}} E W^{-\frac{1}{2}} - w^{\mathrm{t}} W_i^{\frac{1}{2}} \right\|^2 \\
&\quad + \left\| x^{\mathrm{t}} P_i E W^{-\frac{1}{2}} + m_i^{\mathrm{t}} E W^{-\frac{1}{2}} \right\|^2 .
\end{aligned}
\tag{50}
$$

Now, dealing with the first brace after the equal sign in the right side of (49), we do as follows:

$$
\begin{aligned}
2\left(x^{\mathrm{t}} P_i B_i + m_i^{\mathrm{t}} B_i\right) u_i &+ u_i^{\mathrm{t}} R_{i,i} u_i \\
&= 2\left(x^{\mathrm{t}} P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i \mathrm{t} B_i R_{i,i}^{-\frac{1}{2}}\right) R_{i,i}^{\frac{1}{2}} u_i + \left\| u_i^{\mathrm{t}} R_{i,i}^{\frac{1}{2}} \right\|^2 \\
&= \left\| x^{\mathrm{t}} P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i^{\mathrm{t}} B_i R_{i,i}^{-\frac{1}{2}} + u_i^{\mathrm{t}} R_{i,i}^{\frac{1}{2}} \right\|^2 \\
&\quad - \left\| x^{\mathrm{t}} P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i^{\mathrm{t}} B_i R_{i,i}^{-\frac{1}{2}} \right\|^2 .
\end{aligned}
\tag{51}
$$

Inserting (50) and (51) into (49), defining $S_i := B_i R_{i,i}^{-1} B_i^{\mathrm{t}}$ and $M_i := E W_i^{-1} E^{\mathrm{t}}$, we get

$$
\begin{aligned}
\hat{V}_i(T) - \hat{V}_i(0) = \int_0^T &\Bigg(\Big(x^{\mathrm{t}}\left(P_i A + A^{\mathrm{t}} P_i + Q_i - P_i S_i P_i + P_i M_i m_i\right) x \\
&+ 2x^{\mathrm{t}}\left(\dot{m}_i + A^{\mathrm{t}} m_i - P_i S_i m_i + P_i M_i m_i + P_i c\right) \\
&+ \left(\dot{n} + m_i^{\mathrm{t}} M_i m_i - m_i^{\mathrm{t}} S_i m_i + 2m_i^{\mathrm{t}} c\right) \\
&+ 2x^{\mathrm{t}} P_i \sum_{j \in I_{N,i}} B_j u_j + 2m_i^{\mathrm{t}} \sum_{j \in I_{N,i}} B_j u_j \Big) \\
&+ \left\| x^{\mathrm{t}} P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i^{\mathrm{t}} B_i R_{i,i}^{-\frac{1}{2}} + u_i^{\mathrm{t}} R_{i,i}^{\frac{1}{2}} \right\|^2 \\
&- \left\| x^{\mathrm{t}} P_i E W^{-\frac{1}{2}} + m_i^{\mathrm{t}} E W^{-\frac{1}{2}} - w^{\mathrm{t}} W^{\frac{1}{2}} \right\|^2 \Bigg) \\
&- \int_0^T \left(x^{\mathrm{t}} Q_i x + u_i^{\mathrm{t}} R_{i,i} u_i - w^{\mathrm{t}} W_i w \right) .
\end{aligned}
\tag{52}
$$

In (52) since u_j is minimizing at the same time u_i does, then we know that the form of the control for u_j is

$$
u_j = -R_{j,j}^{-1} B^{\mathrm{t}}\left(P_j x + m_j\right).
\tag{53}
$$

Substitution of (53) into (52) gives

$$
\begin{aligned}
\hat{V}_i(T) - \hat{V}_i(0) = \int_0^T & \left(\left(x^t \left(P_i \left(A - \sum_{j \in I_{N,i}} S_j P_j \right) + \left(A - \sum_{j \in I_{N,i}} S_j P_j \right)^t P_i \right. \right. \right. \\
& + Q_i - P_i S_i P_i + P_i M_i m_i + \sum_{j \in I_{N,i}} P_j S_{j,i} P_j \bigg) x \\
& + 2x^t \left(\dot{m}_i + A^t m_i - P_i S_i m_i + P_i M_i m_i \right. \\
& + \sum_{j \in I_{N,i}} P_j S_{j,i} m_j - \sum_{j \in I_{N,i}} P_j S_j m_i - \sum_{j \in I_{N,i}} P_i S_j m_j + P_i c \bigg) \\
& + \left(\dot{n} + m_i^t M_i m_i - m_i^t S_i m_i \right. \\
& + \sum_{j \in I_{N,i}} m_j^t S_{j,i} m_j - 2 m_i^t \sum_{j \in I_{N,i}} S_j m_j + 2 m_i^t c \bigg) \\
& + \left\| x^t P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i^t B_i R_{i,i}^{-\frac{1}{2}} + u_i^t R_{i,i}^{\frac{1}{2}} \right\|^2 \bigg) \\
& - \left\| x^t P_i E W_i^{-\frac{1}{2}} + m_i^t E W_i^{-\frac{1}{2}} - w^t W_i^{\frac{1}{2}} \right\|^2 \\
& - \left\| x^t \sum_{j \in I_{N,i}} P_j B_j R_{j,j}^{-1} R_{j,i}^{\frac{1}{2}} + \sum_{j \in I_{N,i}} m_j^t B_j R_{j,j}^{-1} R_{j,i}^{\frac{1}{2}} + u_j^t R_{j,i}^{\frac{1}{2}} \right\|^2 \bigg) \\
& - \int_0^T \left(x^t Q_i x + u_i^t R_{i,i} u_i + u_j^t R_{j,i} u_j - w^t W_i w \right). \quad (54)
\end{aligned}
$$

According to (54) we have to find the solution to the N equation. This solution is found by solving simultaneously the differential equations system

$$
\begin{pmatrix} \dot{m}_1 \\ \dot{m}_2 \\ \vdots \\ \dot{m}_N \end{pmatrix} = \begin{pmatrix} A_1 & P_2 S_{21} - P_1 S_2 & \cdots & P_N S_{N1} - P_1 S_N \\ P_1 S_{12} - P_2 S_1 & A_2 & \cdots & P_N S_{N2} - P_2 S_N \\ \vdots & \vdots & \ddots & \vdots \\ P_1 S_{1N} - P_N S_1 & P_2 S_{2N} - P_N S_2 & \cdots & A_N \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{pmatrix} \\
+ \begin{pmatrix} P_1 c \\ P_2 c \\ \vdots \\ P_N c \end{pmatrix},
$$

where the result for m_i are the N equation stated in (29) of Theorem 4.

From the assumptions of equations (24)–(30), this is P_i, m_i, n_i are solutions to the equations (26), (29), (45) respectively, and substituting these equations into (54) we find that

$$
\begin{aligned}
\hat{V}_i(T) - \hat{V}_i(0) = \int_0^T &\left(\left(\left\| x^{\mathrm{t}} P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i^{\mathrm{t}} B_i R_{i,i}^{-\frac{1}{2}} + u_i^{\mathrm{t}} R_{i,i}^{\frac{1}{2}} \right\|^2 \right) \right. \\
&- \left\| x^{\mathrm{t}} P_i E W_i^{-\frac{1}{2}} + m_i^{\mathrm{t}} E W_i^{-\frac{1}{2}} - w^{\mathrm{t}} W_i^{\frac{1}{2}} \right\|^2 \\
&\left. - \left\| x^{\mathrm{t}} \sum_{j \in I_{N,i}} P_j B_j R_{j,j}^{-1} R_{j,i}^{\frac{1}{2}} + \sum_{j \in I_{N,i}} m_j^{\mathrm{t}} B_j R_{j,j}^{-1} R_{j,i}^{\frac{1}{2}} + u_j^{\mathrm{t}} R_{j,i}^{\frac{1}{2}} \right\|^2 \right) \\
&- \int_0^T \left(x^{\mathrm{t}} Q_i x + u_i^{\mathrm{t}} R_{i,i} u_i + u_j^{\mathrm{t}} R_{j,i} u_j - w^{\mathrm{t}} W_i w \right).
\end{aligned}
$$

Now, since m_i converges exponentially to zero, because the matrix $A_i - \sum_{j=1}^N (P_j S_{j,i} - P_i S_j)$ is stable, then n_i also converges, because it depends on the m_i terms and c, which is a locally square integrable function that converges to zero exponentially. Also since $x(T) \to 0$ as $T \to +\infty$, then $\lim_{T \to +\infty} \hat{V}_i(T) = 0$, obtaining

$$
\begin{aligned}
-\hat{V}_i(0) = \int_0^{+\infty} &\left(\left\| x^T P_i B_i R_{i,i}^{-\frac{1}{2}} + m_i^{\mathrm{t}} B_i R_{i,i}^{-\frac{1}{2}} + u_i^{\mathrm{t}} R_{i,i}^{\frac{1}{2}} \right\|^2 \right. \\
&- \left\| x^T P_i E W_i^{-\frac{1}{2}} + m_i^T E W_i^{-\frac{1}{2}} - w^T W_i^{\frac{1}{2}} \right\|^2 \\
&\left. - \left\| x^{\mathrm{t}} \sum_{j \in I_{N,i}} P_j B_j R_{j,j}^{-1} R_{j,i}^{\frac{1}{2}} + \sum_{j \in I_{N,i}} m_j^{\mathrm{t}} B_j R_{j,j}^{-1} R_{j,i}^{\frac{1}{2}} + u_j^{\mathrm{t}} R_{j,i}^{\frac{1}{2}} \right\|^2 \right) - J_i,
\end{aligned}
$$

(55)

where

$$
J_i = \int_0^{+\infty} \left(x^{\mathrm{t}} Q_i x + u_i^{\mathrm{t}} R_{i,i} u_i + u_j^{\mathrm{t}} R_{j,i} u_j - w^{\mathrm{t}} W_i w \right).
$$

This means that the control that minimizes J_i has the form of equation (27) and the term that maximizes J_i has the form of equation (28). In that case, when we substitute u_i and w_i in (55) we find that the optimal control has the value given in (30). $\qquad\square$

Acknowledgements

References

Aliyu, M.D.S. (2011). *Nonlinear $\mathcal{H}_\infty$ Control, Hamiltonian sYSTEMS and Hamiltonian–Jacobi Equations*. CRC Press, Boca Raton Fl.

Başar, T., Bernhard, P. (2008). *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd ed. Birkhäuser.

Başar, T., Olsder, G.J. (1999). *Dynamic Noncooperative Game Theory* 2nd ed. *Classics in Applied Mathematics*, *Vol.* 23. SIAM.

Chen, H., Scherer, C.W., Allgöwer, F. (1997). A game theoretic approach to nonlinear robust receding horizon control of constrained systems. In: *Proceedings of the American Control Conference*, Vol. 5. American Automatic Control Council, pp. 3073–3077.

Chen, X., Zhou, K. (2001). Multiobjective $\mathcal{H}_2/\mathcal{H}_\infty$ control design. *SIAM Journal on Control and Optimization*, 40(2), 628–660.

Dockner, E.J., Jørgensen, S., Long, N.V., Sorger, G. (2000). *Differential Games in Economics and Management Science*. Cambridge University Press.

Engwerda, J. (2005). *LQ Dynamic Optimization and Differential Games*. Wiley.

Engwerda, J.C. (2017). Robust open-loop Nash equilibria in the noncooperative LQ game revisited. *Optimal Control Applications & Methods*, 38(5), 795–813.

Fattorini, H.O. (1999). *Infinite-Dimensional Optimization and Control Theory. Encyclopedia of Mathematics and Its Applications*, Vol. 62. Cambridge University Press.

Friedman, A. (1971). *Differential Games*. Pure and Applied Mathematics, Vol. XXV. Wiley.

Jank, G., Kun, G. (2002). Optimal control of disturbed linear-quadratic differential games. *European Journal of Control*, 8(2), 152–162.

Jiménez-Lizárraga, M., Poznyak, A. (2007). Robust Nash equilibrium in multi-model LQ differential games: analysis and extraproximal numerical procedure. *Optimal Control Applications and Methods*, 28(2), 117–141.

Jørgensen, S. (1986). Optimal production, purchasing and pricing: a differential game approach. *European Journal of Operational Research*, 24(1), 64–76.

Jungers, M., Castelan, E.B., de Pieri, E.R., Abou-Kandil, H. (2008). Bounded Nash type controls for uncertain linear systems. *Automatica. A Journal of IFAC*, 44(7), 1874–1879.

Poznyak, A.S. (2008). *Advanced Mathematical Tools for Automatic Control Engineers, Vol. 1: Deterministic Techniques*. Elsevier.

van den Broek, W.A., Engwerda, J.C., Schumacher, J.M. (2003). Robust equilibria in indefinite linear-quadratic differential games. *Journal of Optimization Theory and Applications*, 119(3), 565–595.

Yong, J., Zhou, X.Y. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations. Applications of Mathematics, Stochastic Modelling and Applied Probability*, Vol. 43. Springer.

M. Jimenez-Lizarraga received his BSc in electrical engineering from Culiacan Institute of Technology, and the MSc and PhD degrees in automatic control from CINVESTAV-IPN, in Mexico, in 1996, 2000, and 2006, respectively. He was a postdoctoral fellow at the ECE Department of the Ohio State University in 2009 and a visiting scholar at Heuristique el Diagnostic des Systemes Complexes at Universite de Technologie de Compiegne in 2015. He is a member of the Mexican Academy of Sciences and is currently with the Faculty of Physical and Mathematical Sciences of the UANL. His research interests include differential games, robust, optimal and sliding mode control and applications such as UAVs guidance and control.

S.V. Rodriguez-Sanchez is an associate professor at the UANL. She received her BSc degree in industrial engineering from the Universidad Autonoma del Estado de Hidalgo, Mexico, in 2002 and her PhD degree in engineering from the Universidad de Lleida, Spain, in 2010. Her current research interests include the application of operation research and analytics in agroindustry and forestry.

N. de la Cruz received his master in industrial physics engineering in 2017 from the Faculty of Physical and Mathematical Sciences of the Autonomous University of Nuevo Leon. His reseach interest include dynamic games and applied mathematics.

C.E. Villarreal received his doctoral degree in mathematics from the Centro de Investigacion y de Estudios Avanzados del Instituto Politecnico Nacional (CINVESTAV-IPN), Mexico, in 1998. He is a professor at the Universidad Autonoma de Nuevo Leon (UANL), Mexico. His research interests include stochastic processes, probability and ergodic theory.