© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

On Robust One-Leader Multi-Followers Linear Quadratic Dynamic Games

Mohsen Saffar[1], Hamed Kebriaei[1] and Luigi Iannelli[2]

***Abstract*— In this paper, the theory of robust min-max control is extended to hierarchical and multi-player dynamic games for linear quadratic discrete time systems. The proposed game model consists of a leader and many followers, while the performance of all players is affected by disturbance. The followers play a Nash game with each other and each of them also plays a zero-sum game with the disturbance. In the higher level of the game, the leader plays a Stackelberg game with the followers and at the same time, plays a zero-sum game with disturbance. The Stackelberg-Nash-saddle point equilibrium of the game is derived using dynamic programming approach and some conditions for existence and uniqueness of the solution are given. Finally an illustrative example is given in the simulation results.**

I. Introduction

Hierarchical decision making has been applied widely in many engineering applications like smart grids, manufacturing and wireless communication networks [1]–[6]. For example, the problem of demand side management in smart grids is modeled as a hierarchical game between a distribution company and users [1], [2]. In manufacturing systems, the problem of inventory control is modeled as a hierarchical decision making problem in which, a manufacturer and retailers constitute a two level supply-chain, where the manufacturer is the leader and retailers are the followers [3], [4]. Also in bandwidth allocation problems in wireless communication networks, the interaction between the service provider and the users is modeled as a two layer decision making problem [5], [6].

The problem of multi-player hierarchical decision making becomes more complicated when players decide repeatedly, in a dynamic environment. Dynamic games have been studied thoughtfully in the literature of control theory [7]. The Nash and Stackelberg equilibrium points are the well-known solutions of a dynamic game, where the players decide simultaneously or as a leader-follower, respectively [8], [9]. Recently, dynamic games have been applied to hierarchical decision making in fields like pricing and spectrum sharing in communication networks [10]. However, still there is lack of a general formulation of dynamic games for hierarchical decision making problem. Since the hierarchical decision making is commonly used as an effective approach for tackling decision making in large-scale environments, the uncertainty is a highly influential input to the system, which affects the performance of the players. Therefore, it is desirable to model the uncertainties and consider them in the optimal decision making of players.

To tackle the uncertainties introduced by disturbances in the environment, the min-max control problem can be looked at as a game theoretic perspective of robust H-infinite control strategy in a dynamic environment [11]. In this framework, the problem is modeled as a zero-sum dynamic game, in which, the uncertainty is considered as a virtual player who wants to maximize the cost function while decision maker wants to minimize it [11]. Some applications of dynamic min-max control can be found in security problems [12], [13]. Recently, the min-max control strategy has been also applied for control over network's losses under TCP-like protocol, where the wireless network system is affected by link failures and packets drop [14], [15]. In [16] the min-max game is also applied for decision making in a two player sequential game. However, to the best of the authors' knowledge, the robust decision making has not been explored in multi-player and dynamic hierarchical games, yet.

This paper addresses the general formulation of dynamic discrete time hierarchical linear quadratic games, considering the uncertainties as a bounded exogenous disturbance to the system. In the proposed dynamic game, the leader acts first and play a Stackelberg game with the followers. In the second level the followers play an n-player non-cooperative game and decide simultaneously, after the leader. All the players, including the leader and the followers, also play a zero sum game with the disturbance. Then, the sub-game perfect equilibrium points of the game are derived by using a dynamic programming approach. It is shown that under certain conditions the proposed dynamic hierarchical game admits a Stackelberg-Nash-saddle point equilibrium which can be interpreted as the robust optimal strategy of the players. It is shown that the robust optimal closed loop strategies of the leader and the followers are in fact the linear feedback strategies and the corresponding Riccati equations are derived. Some conditions for the existence and uniqueness of the equilibrium point of the dynamic hierarchical game are also given.

This paper is organized as follows: The problem formulation, the closed loop robust optimal strategies for the players and the conditions for the existence and uniqueness of the equilibrium points are given in this Section II. Section III includes the simulation results and finally the paper is concluded in Section IV.

[1]Mohsen Saffar and Hamed Kebriaei are with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Iran, email: {mohsen.saffar, kebriaei}@ut.ac.ir

[2]Luigi Iannelli is with the Department of Engineering, University of Sannio in Benevento, Italy, e-mail: luigi.iannelli@unisannio.it

II. Equilibrium Analysis of the Hierarchical Dynamic Game

Consider a dynamic system that has one leader and n followers, all together playing over a states space evolving according to the following difference equation

$$x_{k+1} = Ax_k + B^{n+1}v_k + \sum_{i=1}^{n} B^i u_k^i + Dw_k, \quad (1)$$

where $x_k \in \mathbb{R}^q$ is the value of states, $v_k \in \mathbb{R}^m$ is the leader's decision, $u_k^i \in \mathbb{R}^m$, $i = 1, 2, \ldots, n$ is the i-th follower's decision and $w_k \in \mathbb{R}^p$ is the disturbance, and k points to the index of stage.

Assume that the players decide rationally and all the players are aware of this fact and know all players' parameters and the state dynamics. Moreover, in order to implement a closed-form control strategy, it is assumed that the state information is available to all players. The leader's decision is based on followers' rationality and both leader and followers are trying to set their control strategy in order to minimize their cost function over a finite number of stages, say N. Those cost functions are in quadratic form, as inspired by the robust game theoretic approach [11]. They can be written as

$$J_k^i(x_k, u_\bullet^i, w_\bullet) = \sum_{m=k}^{N-1} \left(x_m^\top Q^i x_m + r^i \|u_m^i\|^2 - \gamma^2 \|w_m\|^2 \right) + x_N^\top Q_N^i x_N, \quad i = 1, 2, \ldots, n \quad (2)$$

$$J_k^{n+1}(x_k, v_\bullet, w_\bullet) = \sum_{m=k}^{N-1} \left(x_m^\top Q^{n+1} x_m + r^{n+1} \|v_m\|^2 - \gamma^2 \|w_m\|^2 \right) + x_N^\top Q_N^{n+1} x_N, \quad (3)$$

where Q^i, Q^{n+1} are all positive semidefinite matrices (such property is denoted by $\geqslant 0$), the terminal weights are positive definite matrices ($Q_N^i, Q_N^{n+1} > 0$), γ and $r^i, r^{n+1} > 0$ are positive scalars, and the sequences of signals of interest are denoted by the subscript $\bullet$, i.e., in (2) $u_\bullet^i = \{u_k^i, \ldots, u_{N-1}^i, u_N^i\}$. Equations (2), (3) describe the followers' and the leader's cost functions, respectively, and weight the decision sequences $u_\bullet^i$ ($v_\bullet$ for the leader) and $w_\bullet$, together with the the state sequence $x_\bullet$ obtained from the initial state x_k through the game dynamics (1). All of the players including the leader and followers play a zero-sum game against the disturbance w, by considering it as a further virtual player who wants to maximize their own cost functions while other players want to minimize them. Here it is assumed for simplicity (but the analysis can be straightforwardly extended) that cost functions have the same weight γ for the disturbance effect. Note that, when considering a single cost function, such weight assumes the role of the disturbance attenuation parameter related to the upper bound of the H-infinite norm of a suitable transfer function [11].

***Theorem** 1:* Consider the proposed hierarchical game including one leader and n-followers with objective functions defined in (2), (3) and the dynamic system (1) affected by disturbance. If there exists a unique Stackelberg-Nash-Saddle point equilibrium including the strategy of all the players', then the following positive definite conditions need to be satisfied

$$\gamma^2 I - D^\top Z_{k+1}^i D > 0, \quad i = 1, 2, \ldots, n+1 \quad (4a)$$
$$(B^i)^\top Z_{k+1}^i B^i + r^i I > 0, \quad i = 1, 2, \ldots, n \quad (4b)$$
$$(B_k')^\top Z_{k+1}^{n+1} B_k' + r^{n+1} I > 0, \quad (4c)$$

where the recursive difference Riccati equations of Z_k^i $\forall k \leqslant N$ are given by

$$Z_k^i = Q^i + (H_k^i)^\top Z_{k+1}^i H_k^i + r^i {L_k^i}^\top L_k^i - \gamma^2 (L_k^{-i})^\top L_k^{-i} \quad i = 1, 2, \ldots, n \quad (5)$$

$$Z_k^{n+1} = Q^{n+1} + (H_k^{n+1})^\top Z_{k+1}^{n+1} H_k^{n+1} + r^{n+1} (L_k^{n+1})^\top L_k^{n+1} - \gamma^2 (L_k^{-(n+1)})^\top L_k^{-(n+1)}, \quad (6)$$

with the terminal conditions on Z_k^i as

$$Z_N^i = Q_N^i > 0 \quad i = 1, 2, \ldots, n \quad (7a)$$
$$Z_N^{n+1} = Q_N^{n+1} > 0, \quad (7b)$$

and

$$H_k^i = A_k' + B_k' L_k^{n+1} + D L_k^{-i} \quad i = 1, 2, \ldots, n \quad (8)$$
$$H_k^{n+1} = A_k' + B_k' L_k^{n+1} + D L_k^{-(n+1)} \quad (9)$$
$$L_k^i = F_k^i (A + B^{n+1} L_k^{n+1}) \quad i = 1, 2, \ldots, n \quad (10)$$
$$L_k^{-i} = F_k^{-i} (A + B^{n+1} L_k^{n+1}) \quad i = 1, 2, \ldots, n \quad (11)$$
$$M_k = \sum_{i=1}^{n} B^i F_k^i \quad (12)$$
$$A_k' = A + M_k A \quad (13)$$
$$B_k' = B^{n+1} + M_k B^{n+1}, \quad (14)$$

with the Stackelberg-Nash-Saddle point being linear with respect to the state and obtained through the gains $L_k^\bullet$ as below:

$$\left({u_k}^{i*}, {w_k}^{i*}, {v_k}^*, {w_k}^{n+1*} \right) = \left(L_k^i, L_k^{-i}, L_k^{n+1}, L_k^{-(n+1)} \right) x_k, \quad (15)$$

where $i = 1, 2, \ldots, n$ refer to the followers, $n+1$ points to the leader, and the indices $-1, -2, \ldots, -n$ indicate the different disturbance worst case feedback strategy for different followers and the index $i = -(n+1)$ refers to the disturbance worst case feedback strategy for the leader.

Proof: Based on the bottom-up principle we want to solve the hierarchal game from down to top. In lower level, the sub-game perfect equilibrium point of the game is obtained by finding the Nash-saddle point equilibrium of the game among the followers and also between each follower and the disturbance. Consider the simultaneous decision making of the n-followers according to cost functions given in (2). First, we want to find the optimal reaction functions of the followers to the strategy of the leader, by taking into account the state and system dynamics information (it is assumed they know it) and the worst case effect of disturbance on their own cost function. That means each follower could get different saddle point equilibria with respect to the disturbance. Thus we will denote by w_k^i the disturbance variable into the i-th follower. Assume that the system is in the k-th stage and all the cost-to-go functions in the next stages have been already optimized. The goal is to calculate the k-th stage saddle point equilibrium for all the followers. Then we have the following value function to be optimized:

$$V_k^i(x_k) = \min_{u_k^i} \max_{w_k^i} J_k^i(x_k, u_k^i, w_k^i), \tag{16}$$

and, by using the dynamic programming approach, the cost function will be:

$$\begin{aligned} J_k^i(x_k, u_k^i) =& x_{k+1}^\top Z_{k+1}^i x_{k+1} + x_k^\top Q^i x_k + r^i \|u_k^i\|^2 - \gamma^2 \|w_k^i\|^2 \\ =&(Ax_k + B^{n+1}v_k + \sum_{j=1}^{n} B^j u_k^j + Dw_k^i)^\top Z_{k+1}^i \\ &\cdot (Ax_k + B^{n+1}v_k + \sum_{j=1}^{n} B^j u_k^j + Dw_k^i) \\ &+ x_k^\top Q^i x_k + r^i \|u_k^i\|^2 - \gamma^2 \|w_k^i\|^2. \end{aligned} \tag{17}$$

In order to calculate optimal strategies of the followers and the worst case disturbance (i.e., $\left(u_k^{i\,*}, w_k^{i\,*}\right)$), the first derivatives of the cost-to-go function with respect to u_k^i and w_k^i are set equal to zero as follows:

$$\frac{\partial J_k^i}{\partial u_k^i} = 0 \rightarrow (B^i)^\top Z_{k+1}^i \sum_{j=1}^{n} B^j u_k^j + r^i u_k^i + (B^i)^\top Z_{k+1}^i D w_k^i$$
$$= -(B^i)^\top Z_{k+1}^i (Ax_k + B^{n+1}v_k) \tag{18a}$$
$$\frac{\partial J_k^i}{\partial w_k^i} = 0 \rightarrow D^\top Z_{k+1}^i \sum_{j=1}^{n} B^j u_k^j + (D^\top Z_{k+1}^i D - \gamma^2 I) w_k^i$$
$$= -D^\top Z_{k+1}^i (Ax_k + B^{n+1}v_k). \tag{18b}$$

If we denote by $\mathcal{B}_k \in \mathbb{R}^{mn\times mn}$, $\mathcal{D}_k \in \mathbb{R}^{pn\times pn}$ and $\mathcal{A}_k \in \mathbb{R}^{pn\times pn}$ the block matrices of the matrix α_k whose expression is reported in (19) such that

$$\alpha_k = \begin{pmatrix} \mathcal{B}_k & \mathrm{diag}(\mathcal{A}_k^\top) \\ \mathcal{A}_k & \mathcal{D}_k \end{pmatrix}, \tag{20}$$

then conditions (18) can be rewritten as:

$$\mathcal{B}_k \mathcal{U}_k^* + \mathrm{diag}(\mathcal{A}_k^\top)\mathcal{W}_k^* = \mathcal{E}_k (Ax_k + B^{n+1}v_k) \tag{21a}$$
$$\mathcal{A}_k \mathcal{U}_k^* + \mathcal{D}_k \mathcal{W}_k^* = \mathcal{F}_k (Ax_k + B^{n+1}v_k), \tag{21b}$$

with $\mathcal{E}_k$ and $\mathcal{F}_k$ suitable matrices and $\mathcal{U}_k^*$ and $\mathcal{W}_k^*$ column vectors obtained by stacking the Nash-saddle-point values $u_k^{i\,*}$ and $w_k^{i\,*}$. As it is shown, there are $2n$ matrix equations that should be solved simultaneously, in order to obtain the Nash-saddle-point equilibrium strategy of all followers in k-th stage of the game. In particular, the first equation involves m unknown components of u_k^i and the second equation p components of w_k^i. System (21) has $(m+p)n$ linear equations with $(m+p)n$ unknowns.

If the Nash-Saddle-point equilibrium exists and it is unique, that has to be linear with respect to the right-hand side of (21) and, in particular, the optimal strategies can be expressed as follow:

$$u_k^{i\,*} = F_k^i (Ax_k + B^{n+1}v_k) \qquad i = 1, 2, \ldots, n \tag{22a}$$
$$w_k^{i\,*} = F_k^{-i} (Ax_k + B^{n+1}v_k), \quad i = 1, 2, \ldots, n, \tag{22b}$$

where F_k^i and F_k^{-i} are matrices related to the inverse of α_k. The value $w_k^{i\,*}$ corresponds to the worst case that could be caused by disturbance on the i-th follower's cost function. It is clear from (22) that the optimal strategies of the followers are dependent on the strategy of the leader. Once the leader made its decision these strategies can be evaluated. In what follows, we will show that leader's strategy is also linear with respect to the state, thus giving an expression

$$v_k^* = L_k^{n+1} x_k, \tag{23}$$

that, used by followers in (22), give the expressions of L_k^i and L_k^{-i} in (10) and (11).

Before looking at the leader's optimal strategy, we have to show what the expression of L_k^{-i} is. Consider the condition (18b) by replacing all the players optimal strategies $u_k^{i\,*}$ and v_k^*:

$$\begin{aligned} (D^\top Z_{k+1}^i D - \gamma^2 I) w_k^{i\,*} =& -D^\top Z_{k+1}^i (A + B^{n+1} L_k^{n+1}) x_k \\ &- D^\top Z_{k+1}^\top \sum_{j=1}^{n} B^j F_k^j (A + B^{n+1} L_k^{n+1}) x_k. \end{aligned} \tag{24}$$

By using (13) and (14) we can write

$$\begin{aligned} w_k^{i\,*} =& (\gamma^2 I - D^\top Z_{k+1}^i D)^{-1} D^\top Z_{k+1}^i \left(A_k' + B_k' L_k^{n+1}\right) x_k \\ =& L_k^{-i} x_k. \end{aligned} \tag{25}$$

Let us focus now on calculating Stackelberg-saddle point equilibrium of the game in higher level which is a Stackelberg game between leader and followers and also a zero sum game between leader and disturbance. By considering that the leader knows the reaction functions of followers and use those functions in its optimal decision making, the followers' reaction functions are put into the states equation and then, the state space equation is changed to the following form:

$$\begin{aligned} x_{k+1} =& Ax_k + B^{n+1}v_k + \sum_{i=1}^{n} (B^i F_k^i (Ax_k + B^{n+1}v_k)) \\ &+ Dw_k \\ =& (A + M_k A)x_k + (B^{n+1} + M_k B^{n+1})v_k + Dw_k \\ =& A_k' x_k + B_k' v_k + Dw_k. \end{aligned} \tag{26}$$

Note that the optimal value of disturbance which was calculated in (22b) is not replaced in (26). The reason is that we also want to calculate the worst case condition for the

$$
\alpha_k = \left(\begin{array}{cccc|cccc}
(B^1)^\top Z^1_{k+1} B^1 + r^1 I & (B^1)^\top Z^1_{k+1} B^2 & \cdots & (B^1)^\top Z^1_{k+1} B^n & (B^1)^\top Z^1_{k+1} D & 0 & \cdots & 0 \\
(B^2)^\top Z^2_{k+1} B^1 & (B^2)^\top Z^2_{k+1} B^2 + r^2 I & \cdots & (B^2)^\top Z^2_{k+1} B^n & 0 & (B^2)^\top Z^2_{k+1} D & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
(B^n)^\top Z^n_{k+1} B^1 & (B^n)^\top Z^n_{k+1} B^2 & \cdots & (B^n)^\top Z^n_{k+1} B^n + r^n I & 0 & 0 & \cdots & (B^n)^\top Z^n_{k+1} D \\
\hline
D^\top Z^1_{k+1} B^1 & D^\top Z^1_{k+1} B^2 & \cdots & D^\top Z^1_{k+1} B^n & D^\top Z^1_{k+1} D - \gamma^2 I & 0 & \cdots & 0 \\
D^\top Z^2_{k+1} B^1 & D^\top Z^2_{k+1} B^2 & \cdots & D^\top Z^2_{k+1} B^n & 0 & D^\top Z^2_{k+1} D - \gamma^2 I & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
D^\top Z^n_{k+1} B^1 & D^\top Z^n_{k+1} B^2 & \cdots & D^\top Z^n_{k+1} B^n & 0 & 0 & \cdots & D^\top Z^n_{k+1} D - \gamma^2 I
\end{array}\right) \tag{19}
$$

leader. The leader's cost-function at k-stage is:

$$
\begin{aligned}
J_k^{n+1}(x_k, v_k) =& x_{k+1}^\top Z_{k+1}^{n+1} x_{k+1} + x_k^\top Q^{n+1} x_{k+1} \\
&+ r^{n+1}\|v_k\|^2 - \gamma^2 \|w_k\|^2 \\
=& (A'_k x_k + B'_k v_k + D w_w)^\top Z_{k+1}^{n+1} \\
&\cdot (A'_k x_k + B'_k v_k + D w_k) + x_k^\top Q^{n+1} x_k \\
&+ r^{n+1}\|v_k\|^2 - \gamma^2 \|w_k\|^2.
\end{aligned} \tag{27}
$$

Taking the derivatives with respect to the control strategy and disturbance and putting them equal to zero we have

$$
\begin{aligned}
\frac{\partial J_k^{n+1}}{\partial v_k} = 0 \rightarrow & (B'_k)^\top Z_{k+1}^{n+1} B'_k v_k + r^{n+1} v_k \\
&+ (B'_k)^\top Z_{k+1}^{n+1} D w_k^{n+1} = -(B'_k)^\top Z_{k+1}^{n+1} A'_k x_k
\end{aligned} \tag{28a}
$$

$$
\begin{aligned}
\frac{\partial J_k^{n+1}}{\partial w_k^{n+1}} = 0 \rightarrow & D^\top Z_{k+1}^{n+1} B'_k v_k + (D^\top Z_{k+1}^{n+1} D - \gamma^2) w_k^{n+1} \\
&= -D^\top Z_{k+1}^{n+1} A'_k x_k.
\end{aligned} \tag{28b}
$$

By following the same approach as above for the followers, the optimal values $(v_k^*,\ w_k^{n+1^*})$ of the leader strategy will be

$$
v_k^* = L_k^{n+1} x_k \tag{29a}
$$

$$
w_k^{n+1^*} = L_k^{-(n+1)} x_k. \tag{29b}
$$

and, in particular, the gain $L_k^{-(n+1)}$ is given by

$$
L_k^{-(n+1)} = \left(\gamma^2 I - D^\top Z_{k+1}^{n+1} D\right)^{-1} D^\top Z_{k+1}^{n+1} (A'_k + B'_k L_k^{n+1}). \tag{30}
$$

Once the optimal strategy of the leader, followers and disturbance is achieved, one can compute the Stackelberg-Nash-saddle point equilibrium of the game given in (22); it comes out that the cost functions must be convex-concave, with respect to the players (follower or leader) and the disturbance in every stage, respectively:

$$
\frac{\partial^2 J_k^i}{\partial {u_k^i}^2} = (B^i)^\top Z_{k+1}^i B^i + r^i I > 0 \tag{31a}
$$

$$
\frac{\partial^2 J_k^i}{\partial {w_k^i}^2} = D^\top Z_{k+1}^i D - \gamma^2 I < 0 \tag{31b}
$$

$$
\frac{\partial^2 J_k^{n+1}}{\partial {v_k}^2} = (B'_k)^\top Z_{k+1}^{n+1} B'_k + r^{n+1} I > 0 \tag{31c}
$$

$$
\frac{\partial^2 J_k^{n+1}}{\partial (w_k^{n+1})^2} = D^\top Z_{k+1}^{n+1} D - \gamma^2 I < 0. \tag{31d}
$$

Now, in order to complete the proof, it is needed to derive Riccati equations. To this aim, the obtained control strategies are replaced into the state equation and then into the cost-to-go functions, as well. The Riccati equation of the leader is as follows:

$$
\begin{aligned}
V_k^{n+1}(x_k) =& x_k^\top (A'_k + B'_k L_k^{n+1} + D L_k^{-(n+1)})^\top Z_{k+1}^{n+1} \\
&\cdot (A'_k + B'_k L_k^{n+1} + D L_k^{-(n+1)}) x_k + x_k^\top Q^{n+1} x_k \\
&+ r^{n+1}\|L_k^{n+1} x_k\|^2 - \gamma^2 \|L_k^{-(n+1)} x_k\|^2 \\
=& x_k^\top ((H_k^{n+1})^\top Z_{k+1}^{n+1} H_k^{n+1} + r^{n+1} (L_k^{n+1})^\top L_k^{n+1} \\
&+ Q^{n+1} - \gamma^2 (L_k^{-(n+1)})^\top L_k^{-(n+1)}) x_k \\
=& x_k^\top Z_k^{n+1} x_k.
\end{aligned} \tag{32}
$$

According to the same manipulations, the followers' Riccati equations is obtained as follows:

$$
\begin{aligned}
V_k^i(x_k) =& x_k^\top (A'_k + B'_k L_k^{n+1} + D L_k^{-i})^\top Z_{k+1}^i \\
&\cdot (A'_k + B'_k L_k^{n+1} + D L_k^{-i}) x_k + x_k^\top Q^i x_k \\
&+ r^i x_k^\top (L_k^i)^\top L_k^i x_k - \gamma^2 x_k^\top (L_k^{-i})^\top L_k^{-i} x_k \\
=& x_k^\top (Q^i + (H_k^i)^\top Z_{k+1}^i H_k^i + r^i (L_k^i)^\top L_k^i \\
&- \gamma^2 (L_k^{-i})^\top L_k^{-i}) x_k = x_k^\top Z_k^i x_k.
\end{aligned} \tag{33}
$$

■

***Theorem** 2:* If conditions (4) of Theorem 1 are satisfied and, moreover, the following condition holds $\forall k, i$

$$
\begin{aligned}
r^i > \sum_{j \neq i} &\|(B^i)^\top Z_{k+1}^i B^j \\
&+ (B^i)^\top Z_{k+1}^i D (\gamma^2 I - D^\top Z_{k+1}^i D)^{-1} D^\top Z_{k+1}^i B^j\|
\end{aligned} \tag{34}
$$

then there exists a unique Stackelberg-Nash-Saddle point equilibrium including the strategy of all players.

Proof: Let us rewrite here the conditions related to the optimal strategies of followers:

$$
\mathcal{B}_k \mathcal{U}_k^* + \mathrm{diag}(\mathcal{A}_k^\top) \mathcal{W}_k^* = \mathcal{E}_k (A x_k + B^{n+1} v_k) \tag{35a}
$$

$$
\mathcal{A}_k \mathcal{U}_k^* + \mathcal{D}_k \mathcal{W}_k^* = \mathcal{F}_k (A x_k + B^{n+1} v_k). \tag{35b}
$$

If condition (4a) is satisfied, then $\mathcal{D}_k$ is invertible and

$$
\mathcal{W}_k^* = \mathcal{D}_k^{-1} \left(-\mathcal{A}_k \mathcal{U}_k^* + \mathcal{F}_k (A x_k + B^{n+1} v_k)\right), \tag{36a}
$$

$$
\begin{aligned}
&\left(\mathcal{B}_k - \mathrm{diag}(\mathcal{A}_k^\top) \mathcal{D}_k^{-1} \mathcal{A}_k\right) \mathcal{U}_k^* \\
&= \left(\mathcal{E}_k - \mathrm{diag}(\mathcal{A}_k^\top) \mathcal{D}_k^{-1} \mathcal{F}_k\right) (A x_k + B^{n+1} v_k).
\end{aligned} \tag{36b}
$$

It is not difficult to show that the matrix $\mathcal{M}_k = \mathcal{B}_k - \operatorname{diag} \mathcal{A}_k^\top \mathcal{D}_k^{-1} \mathcal{A}_k$ can be partitioned into the following blocks:

$$(\mathcal{M}_k)_{ij} = (B^i)^\top Z^i_{k+1} D(\gamma^2 - D^\top Z^i_{k+1} D)^{-1} D^\top Z^i_{k+1} B^j + (B^i)^\top Z^i_{k+1} B^j, \quad i,j = 1,\ldots,n,\ i \neq j \tag{37a}$$

$$(\mathcal{M}_k)_{ii} = (B^i)^\top Z^i_{k+1} D(\gamma^2 - D^\top Z^i_{k+1} D)^{-1} D^\top Z^i_{k+1} B^i + (B^i)^\top Z^i_{k+1} B^i + r^i I, \quad i = 1,\ldots,n. \tag{37b}$$

The blocks $(\mathcal{M}_k)_{ii}$ on the diagonal are symmetric and, due to conditions (4a) and (4b), they are also positive definite matrices since $Z^i_k > 0$ as shown in Theorem 1 and Lemma **??**. Their singular values σ coincide with their eigenvalues λ and

$$\begin{aligned}\|(\mathcal{M}_k)^{-1}_{ii}\|^{-1} &= \sigma_{\min}((\mathcal{M}_k)_{ii}) = \lambda_{\min}((\mathcal{M}_k)_{ii}) \\ &= \lambda_{\min}((\mathcal{M}_k)_{ii} - r^i I) + r^i > 0.\end{aligned} \tag{38}$$

Let us now consider the matrix $(\mathcal{M}_k)_{ii} - r^i I$ rewritten as $(B^i)^\top X^i B^i$ with $X^i \in \mathbb{R}^{q\times q}$ and $M^i \in \mathbb{R}^{q\times m}$. In the case $m > q$ (more inputs variables than state dimension for the i-th follower), B^i is not full column rank and, thus, $\lambda_{\min}((\mathcal{M}_k)_{ii} - r^i I) = 0$. Instead, when $q \geqslant m$, we can apply [17, Theorem 3.2] for getting

$$\lambda_{\min}\left((B^i)^\top X^i B^i\right) \geqslant \lambda_{\min}(X^i) \cdot \lambda_{\min}((B^i)^\top B^i). \tag{39}$$

Thus

$$\|(\mathcal{M}_k)^{-1}_{ii}\|^{-1} \geqslant r^i + \lambda_{\min}(X^i) \cdot \lambda_{\min}((B^i)^\top B^i) \tag{40}$$

If conditions

$$r^i > \sum_{j\neq i} \|(B^i)^\top Z^i_{k+1} B^j + (B^i)^\top Z^i_{k+1} D(\gamma^2 I - D^\top Z^i_{k+1} D)^{-1} D^\top Z^i_{k+1} B^j\| \tag{41}$$

hold, then by (40) it comes out that matrix $\mathcal{M}_k$ is block strictly diagonally dominant and, thus, nonsingular [18]. Its invertibility implies the existence and uniqueness of the Nash equilibrium point.

$$\mathcal{U}^*_k = \mathcal{M}_k^{-1}\left(\mathcal{E}_k - \operatorname{diag}(\mathcal{A}_k^\top)\mathcal{D}_k^{-1}\mathcal{F}_k\right)(Ax_k + B^{n+1}v_k) \tag{42a}$$

$$\mathcal{W}^*_k = \mathcal{D}_k^{-1}\left(\mathcal{F}_k - \mathcal{A}_k \mathcal{M}_k^{-1}\left(\mathcal{E}_k - \operatorname{diag}(\mathcal{A}_k^\top)\mathcal{D}_k^{-1}\mathcal{F}_k\right)\right) \cdot (Ax_k + B^{n+1}v_k) \tag{42b}$$

■

III. Simulation

In this section, an illustrative example of proposed hierarchical dynamic system is presented. The number of state variables and the followers are considered equal to two. The state space model and the output of the system are

$$x_0 = \begin{pmatrix} 5 & 3 \end{pmatrix}^\top$$

$$x_{k+1} = \begin{pmatrix} 2 & -3 \\ 8 & 3 \end{pmatrix} x_k + \begin{pmatrix} 1.9 \\ 0.55 \end{pmatrix} v_k + \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} u^1_k + \begin{pmatrix} 0.2 \\ 0.41 \end{pmatrix} u^2_k + \begin{pmatrix} -0.8 \\ 1 \end{pmatrix} w_k$$

$$y_k = \begin{pmatrix} 1 & 1 \end{pmatrix} x_k$$

The leader's control coefficient is set bigger than the followers, since the leader's decision is more effective decision in some practical cases. The disturbance is considered as a uniform noise, bounded in $[-0.5, 0.5]$. The value of other parameters are given in Table I.

TABLE I
System Parameters

Parameter	Leader	First Follower	Second Follower
r^i	0.5	7	1
Q^i_N	$4 \cdot I_2$	$3.25 \cdot I_2$	$2.5 \cdot I_2$
Q^i	$5 \cdot I_2$	$2 \cdot I_2$	$5 \cdot I_2$

Before simulating the discrete time dynamical system, the values of control feedback signals are calculated as proposed in Section II. After that, the control strategies are substituted in the state difference equation and the behavior of the system is evaluated. Figure 1 shows the behavior of the output signal for optimal values of the control strategy and for random disturbances. As it is shown, although the system is stable, as the attenuation parameter is starting to decrease the system goes toward instability. This is an expected result since as it was shown in the paper this is a necessary condition for existence and uniqueness of the saddle point equilibrium of the game. The optimal decisions in each iteration of all players is given in Figure 2.

Note that the disturbance is considered having its actual value. Therefore, after the stabilization, the control strategies are still affected by a zero mean bounded disturbance.

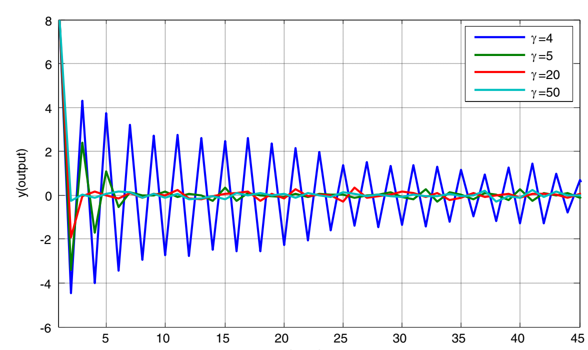

Fig. 1. The output of the system among different attenuation parameter

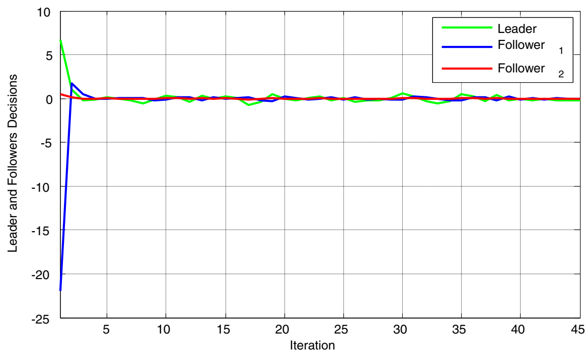

Fig. 2. Players' decisions

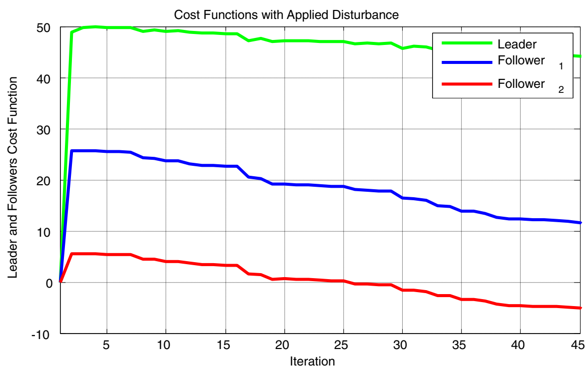

Fig. 3. Players' cost function with the actual disturbances

Figure 3 shows both increasing and decreasing behavior of the cost functions while it is expected to see the accumulating property of the cost. The reason is intuitive and comes from the last part of cost functions in (2)-(3), namely the square of the disturbance. As the states converge to their final value, the state and the decisions' signals will be in a narrow band around zero. But the last term has still an additive negative value that makes the cost function decreasing after some iterations.

IV. Conclusion

In this paper, the H-infinite min-max problem has been extended to a hierarchical multi player robust game, in which the discrete time state space model and the cost function of the players is affected by disturbance.

In the lower level of the hierarchical game, the Nash-saddle equilibrium point has been derived by intersecting the followers' best responses together with disturbance's worst case strategies for each follower.

In the higher level, the leader plays a Stackelberg game with followers and also participates in a min-max game with the disturbance similar to the followers. Once achieved the leader's optimal decision, the Stackelberg-Nash-saddle equilibrium point was calculated. The optimal feedback strategies of the players and the conditions for the existence and uniqueness of the solution have been given.

Current research work is aimed to investigate properties of the Riccati equation solutions and some stability analysis.

References

[1] A. Mohsenian-Rad, V. V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 320–331, 12 2010.

[2] I. Atzeni, L. Ordóñez, G. Scutari, D. Palomar, and J. Fonollosa, "Noncooperative and cooperative optimization of distributed energy generation and storage in the demand-side of the smart grid," *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2454–2472, 2013.

[3] Y. Y. Yu, G. Huang, and L. Liang, "Stackelberg game-theoretic model for optimizing advertising, pricing and inventory policies in vendor managed inventory (vmi) production supply chains," *Computers & Industrial Engineering*, vol. 57, pp. 368–382, 2009.

[4] Y. Tsao, J. Lu, N. An, R. Al-Khayyai, F. amd Lu, and G. Han, "Retailer shelf-space management with trade allowance: A stackelberg game between retailer and manufacturers," *Int. J. Production Economics*, vol. 148, pp. 133–144, 2014.

[5] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 7–18, 1 2008.

[6] D. Niyato, E. Hossain, and Z. Han, "Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: A game theoretic modeling approach," *IEEE Trans. on Mobile Computing*, vol. 8, no. 8, pp. 1009–1022, 2009.

[7] T. Basar and G. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA, USA: SIAM, 1995.

[8] D. Lukes, "Equilibrium feedback control in linear games with quadratic costs," *SIAM J. on Control and Opt.*, vol. 9, pp. 234–252, 1971.

[9] T. Basar and G. Olsder, "Team-optimal closed-loop stackelberg strategies in hierarchical control problems," *Automatica*, vol. 16, pp. 409–414, 1980.

[10] E. H. K. Zhu and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Trans. on Mobile Computing*, vol. 13, no. 8, pp. 1843—1856, 2014.

[11] T. Basar and Bernhard, *H-Infinity Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Switzerland: Birkhäuser, 1995.

[12] M. Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J. Hubaux, "Game theory meets network security and privacy," *ACM Comput. Surv.*, vol. 45, no. 3, pp. 1—39, 6 2013.

[13] T. Alpcan and T. Basar, *Network Security: A Decision and Game Theoretic Approach*. Cambridge, UK: Cambridge Univ. Press, 2011.

[14] J. Moon and T. Basar, "Control over tcp-like lossy networks: A dynamic game approach," in *Proc. of American Control Conference*, 6 2013, pp. 1581–1586.

[15] ——, "Control over lossy networks: A dynamic game approach," in *Proc. of American Control Conference*, 6 2014, pp. 5367–5372.

[16] D. Shen and J. J. Cruz, "Disturbance attenuation analysis of state feedback nash strategy for linear quadratic sequential games," in *Proc. of European Control Conference*, 7 2007, pp. 3371–3378.

[17] N. Higham and S. H. Cheng, "Modifying the inertia of matrices arising in optimization," *Linear Algebra and its Applications*, vol. 275-276, pp. 261–279, 1998.

[18] D. G. Feingold and R. S. Varga, "Block Diagonally Dominant Matrices and Generalizations of the Gerschgorin Circle Theorem," *Pacific Journal of Mathematics*, vol. 12, no. 4, pp. 1241–1249, 1962.