A Nash LQ Game with an Infinite Horizon Major Player and Many Randomly Entering Minor Players of Different Time Horizons.

Ioannis P. Kordonis
PhD. Candidate
Division of Control, Signals &Robotics
Department of EE&CS
National Technical University of Athens
Athens, Greece
jkordonis1920@yahoo.com

George P. Papavassilopoulos
Professor
Division of Control, Signals &Robotics
Department of EE&CS
National Technical university of Athens
Athens, Greece
yorgos@netmode.ntua.gr

Abstract

We consider a class of Discrete-Time Linear Quadratic Games involving a major player who has an infinite time horizon and a random number of minor players of several types. Each type of minor player has its own time horizon and its own dynamic equation. At every time step, the dynamic equation of each player depends on the state vector of the currently active players and its own control. The number of new minor players of each type, entering at each time step is a random variable following a Markov chain. Sufficient conditions characterizing a Nash equilibrium of Linear Feedback Strategies are derived. Games involving a large number of minor players are then studied using a mean field approach. ε–Nash equilibrium results are derived for the case with a large number of players. Numerical examples are also given.

1. Introduction

In most of the dynamic game models the time interval during which the players are involved in the game as well as the number of players that participate in the game at each time step is quite structured. For example in finite or infinite horizon dynamic games e.x. [27], [6] all the players participate in the game for identical time intervals. In overlapping generation games [26], [5], [13] a known number of players of each generation enters the game at each time step and stays for a certain period of time. Several attempts to impose less structure on the players' time intervals or on the number of players that participate in the game have been made. For example in games with population uncertainty or in Poisson games [20] the number of players that participate in the game is not known a priori. Games with random horizon have been studied in [2] in a repeated game setting and in [28] in a differential game setting. In this class of games the time intervals that the players are involved in the game are identical, however the time horizon is random. In [14] a game with overlapping generations involving players with horizon 2 is considered. The number of players of each generation is however random. This work is a new attempt to impose less structure on the time intervals that the players participate in the game as well as on the number of the players active at each time step.

In this paper we consider a Dynamic Linear Quadratic (LQ) Stochastic Game with players having different time horizons, interacting for different time intervals. Particularly there is a player with infinite time horizon, called the major player and many players with finite time

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES. Investing in knowledge society through the European Social Fund.

horizons, called minor players, that enter the game at every time step. We suppose that each minor player belongs to one of several, but finite in multitude, categories (types). Each type of minor player has its own dynamics, cost function and time horizon (i.e. it stays in the game for a specified number of steps). The minor players enter randomly in the game. Particularly at each time step, the number of new players of each type that enter the game is a random variable with a distribution depending on the number of players of each type that participate in the game at that time step. The problem considered here is the characterization of a Nash equilibrium of Feedback Strategies. After that, the case of a game with a large number of minor players i.e. the case which the number of new minor players entering at each step is large, is considered. A mean field approximation is used to characterize strategies, which are asymptotically optimal as the number of new minor players in each step tends to infinity.

The motivation for the introduction of the major player comes from several game situations where there is a long living agent or institution that at each time step interacts with a number of agents and the interaction with each agent is maintained for a certain, rather small amount of time. For example a bank that gives loans to households may be considered as a major player with an infinite horizon and each person that assumes a loan as a minor player with a finite pre-specified time horizon. The bank issues loans also to enterprises and other entries and thus it deals with several type of customers. Another example of a game theoretic model involving players with different time horizons and different types is a liberalized energy market ([15]) in which there is a public power corporation with an infinite time horizon and many renewable energy producers that take a permission to enter the system for a certain amount of time. A third example is University Games [25], where the students of each semester stand for the minor players and the university as a major player. Other examples involve the study of repeated games with long-run and short run-players [8] such as chain store game and the study of reputation effects [18],[7].

The interest for the games with large number of players is not new. In [21] games with a continuum of players called oceanic games were introduced and a value for such games was defined. Models with a continuum of players were also studied in [4] (see also [22 ch. X]). In recent years the mean field approach in the study of games with large number of players was introduced [9]. The closely related methodology of Nash Certainty Equivalence was recently developed in order to obtain asymptotic Nash equilibrium results as the number of players tends to infinity [10]. Stochastic games with large number of players have been considered in [11]. An LQG game involving a major player and a large number of players of infinite time horizon is considered in [12] and asymptotically optimal decentralized feedback strategies were characterized.

In games with random entrance that we study in the current work, the future number of players is not known precisely and thus the dynamic equations are linear but uncertain. The problem of random entrance is thus reduced to the study of coupled finite and infinite horizon LQ problems for Markov Jump Linear Systems (MJLS). Thus the Nash equilibria are characterized using appropriate coupled Riccati type equations. For the large number of players case, the mean field approach involves the statement of approximate optimal control problems assuming an infinity of players. In that case ε-equilibrium results are proved. The method used to prove the ε-equilibrium results is based on some results connecting the stability and the LQ control of MJLS with the convergence of a sequence of Markov chains. These results are proved in the Appendix and may be of independent interest.

The rest of the paper is organized as follows: In section 2 the dynamics of major and minor players as well as the cost functions are defined. In section 3 sufficient conditions for a strategy of a player to be optimal given the strategies of the other players are obtained. In section 4 sufficient conditions on a set of linear feedback strategies to constitute a Nash equilibrium are obtained. In section 5 a numerical example is given. In section 6 the problem with a large number of players is approximated by a mean field model. Then some ε-Nash equilibrium results are obtained. In

Section 7, we conclude and propose some directions for future research. The proofs of the results in the text are relegated to the appendix.

2. Description of the Game

The random entrance of the minor players is first described. Let $(\Omega, \Im, \Pr)$ be a probability space. The number of the types of minor players is finite; denote by $1,2,...,p$ the types of minor players. For a minor player of type $j \in \{1,2,....,p\}$, let T_j be its horizon i.e. the number of steps that this player stays in the game. Consider a countably infinite set of minor players $\Lambda \equiv \mathbb{N} = \{1,2,...\}$. For any minor player $i \in \Lambda$, let $t_i : \Omega \to \mathbb{N} \cup \{\infty\}$ be a stopping time describing the time step at which the player i enters the game. At each time step k, the number of the minor players of each type that participate in the game may be described by the vector:

$$y_k = \left(N_k^{1,0}, N_k^{1,1}, ..., N_k^{1,T_1-1}, N_k^{2,0}, N_k^{2,1}, ..., N_k^{2,T_2-1}, ..., N_k^{p,0}, N_k^{p,1}, ..., N_k^{p,T_p-1}\right) / s_c \tag{1}$$

where $N_k^{j,l} = \# I_k^{j,l}$, $I_k^{j,l}$ the set of players of type j with entrance time $k-l$ for $l = 0,1,..,T_j - 1$ and s_c the maximum possible number of active players and will be called the scale variable. Let us also denote $I_k = \bigcup_{j=1}^{p} \bigcup_{l=0}^{T_j-1} I_k^{j,l}$, the set of active players at time step k. Suppose that the number of new minor players of type j that enter at time step $k+1$ i.e. $N_{k+1}^{j,0}$ is a random variable with distribution that depends on y_k. Thus the random entrance is modeled by a Markov chain y_k with a finite state space and let $1,2,...,M$ be an enumeration of the Markov chain state space. We shall use the vector form (1) and the enumeration $1,2,...,M$ interchangeably. Denote also by $\Pi = \left[p_{ij}\right]$ the transition matrix i.e. $\Pr\left(y_{k+1} = j \mid y_k = i\right) = p_{ij}$.

Each player participating in the game has its own dynamic equation. The evolution of the state vector of each player depends on the state vectors of the currently active players and its own control. The dynamics of the major player is described by:

$$x^M(k+1) = A^M x^M(k) + \sum_{i \in I_k} F^{MZ_i} x^i(k) / s_c + B^M u^M(k) + w^M(k), \tag{2}$$

where x^M and x^i are the state vector of the major player and minor player i respectively and Z_i is the type of player i. The stochastic disturbances $w^M(k)$ are zero mean iid random variables with finite variances. Denote by D^M the covariance matrix of $w^M(k)$. The initial condition for the major player $x^M(0)$ is also of zero mean and has a finite variance.

The dynamics of the minor player i is described by:

$$x^i(k+1) = A^{Z_i} x^i(k) + G^{Z_i} x^M(k) + \sum_{j \in I_k} F^{Z_i Z_j} x^j(k) / s_c + B^{Z_i} u^i(k) + w^i(k) \tag{3}$$

The initial condition for the minor player i is random and given by:

$$x^i(t_i) = w^i(t_i - 1). \tag{4}$$

The stochastic disturbances $w^i(k)$, $i \in I_k$ are independent for each time step k and independent of the previous values of the state vectors. For each minor player the disturbances

$w^i(k)$ are zero mean iid random variables with finite variances. For a minor player $i \in \Lambda$ of type $j \in \{1,...,p\}$, denote by D^j the covariance matrix of $w^i(k)$.

In order to define the cost functions of the players let us introduce the following quantities:

$$z^{j,l}(k) = \sum_{i \in I_k^{j,l}} x^i(k)/s_c\,,\quad \tilde{z}(k) = \left[z^{1,0}(k) \quad \ldots \quad z^{1,T_1-1}(k) \quad \ldots \quad z^{p,0}(k) \quad \ldots \quad z^{p,T_p-1}(k) \right],$$

where we use the convention that the sum of an empty set of vectors is the zero vector of appropriate dimensions. Let us call the vector $\tilde{z}$, the vector of mean field quantities. The cost function of the major player is given by:

$$J^M = E\left\{ \sum_{k=0}^{\infty} a^k \left[\left[\left(x^M(k)\right)^T \quad \tilde{z}^T(k) \right] Q(y_k) \left[\left(x^M(k)\right)^T \quad \tilde{z}^T(k) \right]^T + \left(u^M(k)\right)^T R u^M(k) \right] \right\} \tag{5}$$

where $Q(y)$, $y \in \{1,...,M\}$ and R are positive semidefinite and positive definite matrices respectively of appropriate dimensions and $a \in (0,1)$ is a discount factor.

For the minor player $i \in \Lambda$ the cost function is given by:

$$J^i = E\left\{ \left(\tilde{x}^i\right)^T \left(t_i + T_{Z(i)}\right) Q_f^{Z_i}\left(y_{t_i+T_{Z(i)}}\right) \left(\tilde{x}^i\right)^T \left(t_i + T_{Z(i)}\right) + \right.$$
$$\left. + \sum_{k=0}^{T_{Z(i)}-1} \left[\left(\tilde{x}^i\right)^T (t_i + k) Q^{Z_i}(y_k) x^i(t_i + k) + \left(u^i(t_i + k)\right)^T R^{Z_i} u^i(t_i + k) \right] \right\} \tag{6}$$

where $\tilde{x}^i(k) = \left[\left(x^M(k)\right)^T \quad \tilde{z}^T(k) \quad \left(x^i\right)^T(k) \right]^T$, $y \in \{1,...,M\}$, $Q_f^{Z_i}(y)$ and $Q^{Z_i}(y)$ are positive semidefinite matrices of appropriate dimensions and R positive definite matrix of appropriate dimensions.

Remark 2.1: Linear dynamics describing the state of the universe may also be included in the model as a part of the state vector of the major player. More general cases with time varying $Q^{Z_i}(y_k)$, or dependence of $A^M, F^{MZ_i}, B^M, A^{Z_i}, G^{Z_i}, F^{Z_iZ_j}, B^{Z_i}$ on y_k may be considered. The dynamic equation of the major and minor players may also depend on the weighted sum of the values of the control variables of the other players. These cases may be analyzed with the methods considered here. For notational simplicity we consider the cases given by (2), (3), (5) and (6). □

Let us suppose that a player involved in the game at time step k has access to the value of the state vector of the major player $x^M(k)$, the mean field quantities $\tilde{z}(k)$, the value of its own state vector and the value of the Markov chain y_k. The problem considered here is the characterization of a Nash equilibrium of closed loop strategies. We shall focus on feedback strategies [6 Def 5.2] i.e. on strategies that depend only on the current state measurements. Particularly a minor player $i \in \Lambda$ uses a strategy in the form:

$$u^i = L^{Z_iM}(k - t_i, y_k) x^M(k) + \sum_{j=1}^{p} \sum_{l=0}^{T_j-1} L^{Z_i}(j, l, k - t_i, y_k) z^{j,l}(k) + \overline{L}^{Z_i}(k - t_i, y_k) x^i(k) \tag{7}$$

and the major player uses a strategy in the form:

$$u^M = L^{MM}(y_k) x^M(k) + \sum_{j=1}^{p} \sum_{l=0}^{T_j-1} L^M(j, l, y_k) z^{j,l}(k) \tag{8}$$

3. Optimal control problems

In this section the optimal control problems for the players involved in the game are stated under the assumption that the other players use strategies (7) and (8). The dynamics of a player as well as its cost function under the control laws (7) and (8) depend only on $x^M(k)$, $\tilde{z}(k)$, the value of its own state vector and y_k. Thus optimal control problems for the players will be stated based on reduced order models involving only these quantities and solutions of the optimal control problems are again in the form (7) and (8). At first the evolution of the mean field vector $\tilde{z}(k)$ is computed. For the component $z^{j,l}$, $j=1,...,p$ and $l=0,...,T_{Z_j}-1$ it holds:

$$z^{j,l+1}(k+1)=\frac{1}{s_c}\sum_{i\in I_{k+1}^{j,l+1}} x^i(k+1)=\frac{1}{s_c}\sum_{i\in I_k^{j,l}} x^i(k+1)$$

For $i\in I_k^{j,l}$ it holds:

$$x^i(k+1)=A^j x^i(k)+G^j x^M(k)+\sum_{j'=1}^{p}\sum_{l'=0}^{T_{j'}-1}F^{jj'}z^{j',l'}+B^j u^i(k)+w^i(k)$$

Thus we have:

$$\begin{aligned} z^{j,l+1}(k+1)=A^j z^{j,l}(k)+\frac{N_k^{j,l}}{s_c}G^j x^M(k)+\frac{N_k^{j,l}}{s_c}\sum_{j'=1}^{p}\sum_{l'=0}^{T_{j'}-1}F^{jj'}z^{j',l'}(k)+\\ +\frac{1}{s_c}B^j\sum_{i\in I_k^{j,l}}u^i(k)+\frac{1}{s_c}\sum_{i\in I_k^{j,l}}w^i(k) \end{aligned} \tag{9}$$

To compute the evolution of $\tilde{z}$ it remains to compute $z^{j,0}(k+1)$. It holds:

$$z^{j,0}(k+1)=\frac{1}{s_c}\sum_{i\in I_{k+1}^{j,0}} x^i(k+1)=\frac{1}{s_c}\sum_{i\in I_{k+1}^{j,0}} w^i(k) \tag{10}$$

3.1 Optimal control for the major player

The dynamics and cost for the major player depend only on the quantities x^M, $\tilde{z}$ and y. Thus using (7), (9) and (10), the evolution of these may be computed:

$$x^M(k+1)=A^M x^M(k)+\sum_{j=1}^{p}\sum_{l=0}^{T_j-1}F^{Mj}z^{j,l}(k)+B^M u^M(k)+w^M(k),$$

$$z^{j,0}(k+1)=\frac{1}{s_c}\sum_{i\in I_{k+1}^{j,0}} w^i(k),$$

$$\begin{aligned} z^{j,l+1}(k+1)=A^j z^{j,l}(k)+\frac{N_k^{j,l}}{s_c}\left(G^j+B^j L^{jM}(l,y_k)\right)x^M(k)+\\ +\frac{N_k^{j,l}}{s_c}\sum_{j'=1}^{p}\sum_{l'=0}^{T_{j'}-1}\left(F^{jj'}+B^j L^j(j',l',l,y_k)\right)z^{j',l'}(k)+\frac{N_k^{j,l}}{s_c}B^j\overline{L}^j(l,y_k)z^{j,l}(k)+\frac{1}{s_c}\sum_{i\in I_k^{j,l}}w^i(k) \end{aligned}$$

The evolution of x^M and $\tilde{z}$ may be described in compact form:

$$\begin{bmatrix} x^M(k+1) \\ \tilde{z}(k+1) \end{bmatrix} = \tilde{A}^M(y_k) \begin{bmatrix} x^M(k) \\ \tilde{z}(k) \end{bmatrix} + \tilde{B}^M u^M(k) + W^M(k) \tag{11}$$

where:

$$\tilde{A}^M = \begin{bmatrix} \tilde{A}^M_{x^M x^M} & \tilde{A}^M_{x^M z^{1,0}} & \cdots & \tilde{A}^M_{x^M z^{1,T_1-1}} & \cdots & \tilde{A}^M_{x^M z^{p,0}} & \cdots & \tilde{A}^M_{x^M z^{p,T_p-1}} \\ \tilde{A}^M_{z^{1,0} x^M} & \tilde{A}^M_{z^{1,0} z^{1,0}} & \cdots & \tilde{A}^M_{z^{1,0} z^{1,T_1-1}} & \cdots & \tilde{A}^M_{z^{1,0} z^{p,0}} & \cdots & \tilde{A}^M_{z^{1,0} z^{p,T_p-1}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \tilde{A}^M_{z^{1,T_1-1} x^M} & \tilde{A}^M_{z^{1,T_1-1} z^{1,0}} & \cdots & \tilde{A}^M_{z^{1,T_1-1} z^{1,T_1-1}} & \cdots & \tilde{A}^M_{z^{1,T_1-1} z^{p,0}} & \cdots & \tilde{A}^M_{z^{1,T_1-1} z^{p,T_p-1}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \tilde{A}^M_{z^{p,0} x^M} & \tilde{A}^M_{z^{p,0} z^{1,0}} & \cdots & \tilde{A}^M_{z^{p,0} z^{1,T_1-1}} & \cdots & \tilde{A}^M_{z^{p,0} z^{p,0}} & \cdots & \tilde{A}^M_{z^{p,0} z^{p,T_p-1}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \tilde{A}^M_{z^{p,T_p-1} x^M} & \tilde{A}^M_{z^{p,T_p-1} z^{1,0}} & \cdots & \tilde{A}^M_{z^{p,T_p-1} z^{1,T_1-1}} & \cdots & \tilde{A}^M_{z^{p,T_p-1} z^{p,0}} & \cdots & \tilde{A}^M_{z^{p,T_p-1} z^{p,T_p-1}} \end{bmatrix} \tag{12}$$

$\tilde{A}^M_{x^M x^M}(y_k) = A^M$, $\qquad \tilde{A}^M_{x^M z^{j,l}}(y_k) = F^{Mj}$,

$\tilde{A}^M_{z^{j,l} x^M}(y_k) = \dfrac{N_k^{j,l-1}}{s_c}\left(G^j + B^j L^{jM}(l-1, y_k)\right)$ for $l \geq 1$,

$\tilde{A}^M_{z^{j,0} x^M}(y_k) = 0$,

$\tilde{A}^M_{z^{j,l} z^{j',l'}}(y_k) = \dfrac{N_k^{j,l-1}}{s_c}\left(F^{jj'} + B^j L^j(j', l', l-1, y_k)\right) + \delta_{l-1,l'}\delta_{j,j'}\left(A^j + \dfrac{N_k^{j,l-1}}{s_c} B^j \bar{L}^j(l', y_k)\right)$ for $l \geq 1$,

$\tilde{A}^M_{z^{j,0} z^{j',l'}}(y_k) = 0 \qquad$ and $\qquad \tilde{B}^M = \begin{bmatrix} B^M \\ 0 \end{bmatrix}$.

Thus the optimal control problem that solves the major player is a discounted infinite horizon LQ control problem for a MJLS. The problem is stated as:

<u>Minimize</u>:

$$J^M = E\left\{\sum_{k=0}^{\infty} a^k \left[\left[\left(x^M(k)\right)^T \quad \tilde{z}^T(k)\right] Q(y_k) \left[\left(x^M(k)\right)^T \quad \tilde{z}^T(k)\right]^T + \left(u^M(k)\right)^T R u^M(k)\right]\right\}$$

<u>subject to</u>:

$$\begin{bmatrix} x^M(k+1) \\ \tilde{z}(k+1) \end{bmatrix} = \tilde{A}^M(y_k) \begin{bmatrix} x^M(k) \\ \tilde{z}(k) \end{bmatrix} + \tilde{B}^M u^M(k) + W^M(k)$$

3.2 Optimal control for a minor player

Consider a minor player i_0 with entrance time t_{i_0} and suppose that other players use the feedback strategies (7) and (8). Then the evolution of the state vector and the cost of i_0 depend only on x^M, x^{i_0}, $\tilde{z}$ and y. The evolution of vector $\tilde{z}$ is computed substituting u^M from (8) and u^i from (7) for all minor players except i_0. It holds:

$$z^{j,0}\left(k+1\right)=\frac{1}{s_c}\sum_{i\in I_{k+1}^{j,0}} w^i\left(k\right),$$

for $j \neq Z_{i_0}$ or $l \neq k-t_{i_0}$ it holds:

$$z^{j,l+1}\left(k+1\right)=A^j z^{j,l}\left(k\right)+\frac{N_k^{j,l}}{s_c}\left(G^j+B^j L^{jM}\left(l,y_k\right)\right)x^M\left(k\right)+$$

$$+\frac{N_k^{j,l}}{s_c}\sum_{j'=1}^{p}\sum_{l'=0}^{T_{j'}-1}\left(F^{jj'}+B^j L^j\left(j',l',l,y_k\right)\right)z^{j',l'}\left(k\right)+\frac{N_k^{j,l}}{s_c}B^j\overline{L}^j\left(l,y_k\right)z^{j,l}\left(k\right)+\frac{1}{s_c}\sum_{i\in I_k^{j,l}} w^i\left(k\right)$$

and for $j = Z_{i_0}$ and $l = k-t_{i_0}$ it holds:

$$z^{j,l+1}\left(k+1\right)=A^j z^{j,l}\left(k\right)+\frac{N_k^{j,l}}{s_c}\left(G^j+\frac{N_k^{j,l}-1}{N_k^{j,l}}B^j L^{jM}\left(l,y_k\right)\right)x^M\left(k\right)-\frac{1}{s_c}B^j\overline{L}^j\left(l,y_k\right)x^{i_0}\left(k\right)+$$

$$+\frac{N_k^{j,l}}{s_c}\sum_{j'=1}^{p}\sum_{l'=0}^{T_{j'}-1}\left(F^{jj'}+\frac{N_k^{j,l}-1}{N_k^{j,l}}B^j L^j\left(j',l',l,y_k\right)\right)z^{j',l'}\left(k\right)+\frac{N_k^{j,l}}{s_c}B^j\overline{L}^j\left(l,y_k\right)z^{j,l}\left(k\right)+$$

$$+\frac{1}{s_c}B^j u^{i_0}\left(k\right)+\frac{1}{s_c}\sum_{i\in I_k^{j,l}} w^i\left(k\right)$$

The evolution of the state vector of the major player is described by:

$$x^M\left(k+1\right)=\left(A^M+B^M L^{MM}\left(y_k\right)\right)x^M\left(k\right)+\sum_{j=1}^{p}\sum_{l=0}^{T_j-1}\left(F^{Mj}+B^M L^M\left(j,l,y_k\right)\right)z^{j,l}\left(k\right)+w^M\left(k\right)$$

Thus the evolution of the quantities x^M, $\tilde{z}$ and x^{i_0} may be described in a more compact form:

$$\begin{bmatrix} x^M\left(k+1\right) \\ \tilde{z}\left(k+1\right) \\ x^{i_0}\left(k+1\right)\end{bmatrix}=\tilde{A}^{j_0}\left(k-t_{i_0},y_k\right)\begin{bmatrix} x^M\left(k\right) \\ \tilde{z}\left(k\right) \\ x^{i_0}\left(k\right)\end{bmatrix}+\tilde{B}^{j_0}u^{i_0}\left(k\right)+W^{i_0}\left(k\right) \tag{13}$$

where: $j_0 = Z_{i_0}$,

$$\tilde{A}^{j_0}\left(k-t_{i_0},y_k\right)=\begin{bmatrix}
\tilde{A}^{j_0}_{x^M x^M} & \tilde{A}^{j_0}_{x^M z^{1,0}} & \cdots & \tilde{A}^{j_0}_{x^M z^{1,T_1-1}} & \cdots & \tilde{A}^{j_0}_{x^M z^{p,0}} & \cdots & \tilde{A}^{j_0}_{x^M z^{p,T_p-1}} & \tilde{A}^{j_0}_{x^M x^{i_0}} \\
\tilde{A}^{j_0}_{z^{1,0} x^M} & \tilde{A}^{j_0}_{z^{1,0} z^{1,0}} & \cdots & \tilde{A}^{j_0}_{z^{1,0} z^{1,T_1-1}} & \cdots & \tilde{A}^{j_0}_{z^{1,0} z^{p,0}} & \cdots & \tilde{A}^{j_0}_{z^{1,0} z^{p,T_p-1}} & \tilde{A}^{j_0}_{z^{1,0} x^{i_0}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\tilde{A}^{j_0}_{z^{1,T_1-1} x^M} & \tilde{A}^{j_0}_{z^{1,T_1-1} z^{1,0}} & \cdots & \tilde{A}^{j_0}_{z^{1,T_1-1} z^{1,T_1-1}} & \cdots & \tilde{A}^{j_0}_{z^{1,T_1-1} z^{p,0}} & \cdots & \tilde{A}^{j_0}_{z^{1,T_1-1} z^{p,T_p-1}} & \tilde{A}^{j_0}_{z^{1,T_1-1} x^{i_0}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\tilde{A}^{j_0}_{z^{p,0} x^M} & \tilde{A}^{j_0}_{z^{p,0} z^{1,0}} & \cdots & \tilde{A}^{j_0}_{z^{p,0} z^{1,T_1-1}} & \cdots & \tilde{A}^{j_0}_{z^{p,0} z^{p,0}} & \cdots & \tilde{A}^{j_0}_{z^{p,0} z^{p,T_p-1}} & \tilde{A}^{j_0}_{z^{p,0} x^{i_0}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\tilde{A}^{j_0}_{z^{p,T_p-1} x^M} & \tilde{A}^{j_0}_{z^{p,T_p-1} z^{1,0}} & \cdots & \tilde{A}^{j_0}_{z^{p,T_p-1} z^{1,T_1-1}} & \cdots & \tilde{A}^{j_0}_{z^{p,T_p-1} z^{p,0}} & \cdots & \tilde{A}^{j_0}_{z^{p,T_p-1} z^{p,T_p-1}} & \tilde{A}^{j_0}_{z^{p,T_p-1} x^{i_0}} \\
\tilde{A}^{j_0}_{x^{i_0} x^M} & \tilde{A}^{j_0}_{x^{i_0} z^{1,0}} & \cdots & \tilde{A}^{j_0}_{x^{i_0} z^{1,T_1-1}} & \cdots & \tilde{A}^{j_0}_{x^{i_0} z^{p,0}} & \cdots & \tilde{A}^{j_0}_{x^{i_0} z^{p,T_p-1}} & \tilde{A}^{j_0}_{x^{i_0} x^{i_0}}
\end{bmatrix} \tag{14}$$

and

$$\tilde{A}^{j_0}_{x^M x^M}\left(k-t_{i_0},y_k\right)=A^M+B^M L^{MM}\left(y_k\right),\qquad \tilde{A}^{j_0}_{x^M z^{j,l}}\left(k-t_{i_0},y_k\right)=F^{Mj}+B^M L^M\left(j,l,y_k\right),$$

$$\tilde{A}^{j_0}_{z^{j,l+1}x^M}\left(k-t_{i_0},y_k\right)=\frac{N_k^{j,l}}{s_c}\left(G^j+B^j L^{jM}\left(l,y_k\right)\right)-\frac{1}{s_c}B^j L^{jM}\left(l,y_k\right)\delta_{j,Z_{i_0}}\delta_{l,k-t_{i_0}}\ \text{ for } l+1\geq 1,$$

$$\tilde{A}^{j_0}_{z^{j,0}x^M}\left(k-t_{i_0},y_k\right)=0$$

$$\tilde{A}^{j_0}_{z^{j,l+1}z^{j',l'}}\left(k-t_{i_0},y_k\right)=\left(A^j+\frac{N_k^{j,l}}{s_c}B^j\bar{L}^j\left(l,y_k\right)\right)\delta_{j,j'}\delta_{l',l}+\frac{N_k^{j,l}}{s_c}\left(F^{jj'}+B^j L^j\left(j',l',l,y_k\right)\right)-$$
$$-\frac{1}{s_c}B^j L^j\left(j',l',l,y_k\right)\delta_{j',Z_{i_0}}\delta_{l,k-t_{i_0}}$$

$$\tilde{A}^{j_0}_{z^{j,0}z^{j',l'}}\left(k-t_{i_0},y_k\right)=0,\qquad \tilde{A}^{j_0}_{x^M x^{i_0}}\left(k-t_{i_0},y_k\right)=0,$$

$$\tilde{A}^{j_0}_{z^{j,l+1}x^{i_0}}\left(k-t_{i_0},y_k\right)=-\frac{1}{s_c}B^j\bar{L}^j\left(l,y_k\right)\delta_{j,Z_{i_0}}\delta_{l,k-t_{i_0}}$$

$$\tilde{A}^{j_0}_{x^{i_0}x^M}\left(k-t_{i_0},y_k\right)=G^{Z_{i_0}},\qquad \tilde{A}^{i_0}_{x^{i_0}z^{j,l}}\left(k-t_{i_0},y_k\right)=F^{j,j'},\qquad \tilde{A}^{j_0}_{x^{i_0}x^{i_0}}\left(k-t_{i_0},y_k\right)=A^{Z_{i_0}}$$

and $\tilde{B}^{j_0}=\begin{bmatrix}0\\0\\B^{Z_{i_0}}\end{bmatrix}$.

Therefore the optimal control problem that solves the minor player i_0 is a finite horizon LQ control problem for a MJLS. The problem is stated as:

Minimize:

$$J^i=E\Big\{\left(\tilde{x}^i\right)^T\left(t_i+T_{Z(i)}\right)Q_f^{Z_i}\left(y_{t_i+T_{Z(i)}}\right)\left(\tilde{x}^i\right)^T\left(t_i+T_{Z(i)}\right)+$$
$$+\sum_{k=0}^{T_{Z(i)}-1}\left[\left(\tilde{x}^i\right)^T\left(t_i+k\right)Q^{Z_i}\left(y_k\right)x^i\left(t_i+k\right)+\left(u^i\left(t_i+k\right)\right)^T R^{Z_i}u^i\left(t_i+k\right)\right]\Big\}$$

subject to:

$$\begin{bmatrix}x^M\left(k+1\right)\\ \tilde{z}\left(k+1\right)\\ x^{i_0}\left(k+1\right)\end{bmatrix}=\tilde{A}^{j_0}\left(k-t_{i_0},y_k\right)\begin{bmatrix}x^M\left(k\right)\\ \tilde{z}\left(k\right)\\ x^{i_0}\left(k\right)\end{bmatrix}+\tilde{B}^{j_0}u^{i_0}\left(k\right)+W^{i_0}\left(k\right)$$

4. Optimality Conditions and Nash Equilibrium

The optimal control problems: minimize (5) subject to (11) and minimize (6) subject to (13) are solved in [17]. The solution of the optimal control problem for the major player is given in terms of some matrices: $K(y),\Lambda(y)$ for $y=1,...,M$. Particularly if there are matrices $K(y),\Lambda(y)$ such that:

$$K(y)=Q(y)+\left(\tilde{A}^M\right)^T(y)\left(a\Lambda(y)-a\Lambda(y)^T\tilde{B}^M\left(R/a+\left(\tilde{B}^M\right)^T\Lambda(y)\tilde{B}^M\right)^{-1}\left(\tilde{B}^M\right)^T\Lambda(y)\right)\tilde{A}^M(y)\quad (15)$$

$$\Lambda(y) = E\left[K(y_{k+1}) \mid y_k = y\right] = \sum_{j=1}^{M} p_{yj} K(j) \tag{16}$$

for $y = 1,\ldots,M$ then the control law given by:

$$u^M(k) = -\left(\left(\tilde{B}^M\right)^T \Lambda(y_k)\tilde{B}^M + R/a\right)^{-1}\left(\tilde{B}^M\right)^T \Lambda(y_k)\tilde{A}^M(y_k)\left[\left(x^M(k)\right)^T \quad \tilde{z}^T(k)\right]^T \tag{17}$$

is optimal under some stability conditions ensuring that the cost with the control law (17) is finite. To state the stability conditions consider the matrix:

$$\tilde{A}_{cl}^M(y) = \tilde{A}^M(y) - \tilde{B}^M\left(\left(\tilde{B}^M\right)^T \Lambda(y)\tilde{B}^M + R/a\right)^{-1}\left(\tilde{B}^M\right)^T \Lambda(y)\tilde{A}^M(y)$$

that corresponds to the closed loop behavior of the system. Then the stability conditions depend on the spectral radius of an $n_M(n_M+1)(M/2) \times n_M(n_M+1)(M/2)$ matrix $\tilde{T}$, where n_M is the dimension of the state vector of (12). The matrix $\tilde{T}$ is the matrix form of the operator $\tilde{T}$ defined in the Appendix A1 and may be computed using the relationships:

$$e_{v(j,i_1,i_2)}^T \tilde{T} e_{v(i,i_3,i_4)} = p_{ij} e_{i_1}^T \tilde{A}_{cl}^M(i) e_{i_3} e_{i_2}^T \tilde{A}_{cl}^M(i) e_{i_4} \tag{18}$$

where:

$$v(j,i_1,i_2) = n_M(n_M+1)(j-1)/2 + (i_1-1)(n_M+1) - i_1(i_1-1)/2 + i_2 - i_1 + 1.$$

If the spectral radius of $\tilde{T}$ is less than $1/a$ then the control law (17) is optimal.

For a minor player i_0, the optimal controller is computed using recursively the following relations:

$$K_{T_{j_0}}^{j_0}(y) = Q_f^{j_0}(y) \tag{19}$$

$$\Lambda_{k+1}^{j_0}(y) = E\left[K_{k+1}^{j_0}\left(y_{k+1+t_{i_0}}\right) \mid y_{k+t_{i_0}} = y\right] = \sum_{j=0}^{M} p_{yj} K_{k+1}^{j_0}(j) \tag{20}$$

$$\begin{aligned} K_k^{j_0}(y) = Q(y) + \left(\tilde{A}^{j_0}(k,y)\right)^T \Big[\Lambda_{k+1}^{j_0}(y) - \\ -\Lambda_{k+1}^{j_0}(y)\tilde{B}^{j_0}\left(R(y) + \left(\tilde{B}^{j_0}\right)^T \Lambda_{k+1}^{j_0}(y)\tilde{B}^{j_0}\right)^{-1}\left(\tilde{B}^{j_0}\right)^T \Lambda_{k+1}^{j_0}(y)\Big]\tilde{A}^{j_0}(k,y) \end{aligned} \tag{21}$$

The optimal control is then given by:

$$u^{i_0}\left(k+t_{i_0}\right) = L_k^{j_0}\left(y_{k+t_{i_0}}\right)\left[\left(x^M\left(k+t_{i_0}\right)\right)^T \quad \tilde{z}^T\left(k+t_{i_0}\right) \quad \left(x^{i_0}\left(k+t_{i_0}\right)\right)^T\right] \tag{22}$$

where:

$$L_k^{j_0}(y) = -\left(R(y) + \left(\tilde{B}^{i_0}\right)^T \Lambda_{k+1}(y)\tilde{B}^{i_0}\right)^{-1}\left(\tilde{B}^{i_0}\right)^T \Lambda_{k+1}(y)\tilde{A}^{i_0}(k,y) \tag{23}$$

In order a set of feedback strategies to constitute a Nash equilibrium, each strategy should be optimal given the other strategies. Thus some consistency conditions are stated in the next definition.

Definition 4.1: Consider the set of feedback strategies (7) and (8). Compute the matrix $\tilde{A}^M$ and the set of matrices $\tilde{A}^j(k,y)$ for $j = 1,\ldots,p$, $k = 0,\ldots,T_j - 1$ and $y = 1,\ldots,M$. Then the set of feedback strategies it is called consistent if:

(a) There exist a set of matrices $K(y), L(y)$, $y = 1,\ldots,M$ satisfying (15) and (16) and moreover it holds:

$$L^{MM}(y)=-\left(\left(\tilde{B}^{M}\right)^{T}\Lambda(y)\tilde{B}^{M}+R/a\right)^{-1}\left(\tilde{B}^{M}\right)^{T}\Lambda(y)\tilde{A}^{M}(y)\begin{bmatrix}I_{n_M} & 0 & \ldots & 0\end{bmatrix}^{T} \quad (24)$$

and

$$L^{M}(j,l,y)=-\left(\left(\tilde{B}^{M}\right)^{T}\Lambda(y)\tilde{B}^{M}+R/a\right)^{-1}\left(\tilde{B}^{M}\right)^{T}\Lambda(y)\tilde{A}^{M}(y)\begin{bmatrix}0 & \ldots & 0 & I_{n_j} & 0 & \ldots & 0\end{bmatrix}^{T} \quad (25)$$

for $j=1,\ldots,p$, $k=0,\ldots,T_j-1$ and $y=1,\ldots,M$, where n_j is the size of the state vector of a minor player of type j and I_{n_j} appears after $1+T_1+\ldots+T_{j-1}+l$ zero matrices of appropriate dimensions (in the same place as $z^{j,l}$ in $\begin{bmatrix}x^T & \tilde{z}^T\end{bmatrix}^T$).

(b) The spectral radius of the matrix $\tilde{T}$ computed using (18) is less than $1/a$.

(c) For any type j_0 if the vectors $L_k^{j_0}(y)$ for $y=1,\ldots,M$ and $k=0,\ldots,T_j-1$ computed by (19 - 23) satisfy:

$$L^{j_0M}(k,y)=L_k^{j_0}(y)\begin{bmatrix}I_{n_M} & \begin{bmatrix}0 & \ldots & 0\end{bmatrix} & 0\end{bmatrix}^{T} \quad (26)$$

$$L^{j_0}(j,l,k,y)=L_k^{j_0}(y)\begin{bmatrix}0 & \begin{bmatrix}0 & \ldots & 0 & I_{n_j} & 0 & \ldots & 0\end{bmatrix} & 0\end{bmatrix}^{T} \quad (27)$$

$$\bar{L}^{j_0}(k,y)=L_k^{j_0}(y)\begin{bmatrix}0 & \begin{bmatrix}0 & \ldots & 0\end{bmatrix} & I_{n_{j0}}\end{bmatrix}^{T} \quad (28)$$

I_{n_j} appears after $1+T_1+\ldots+T_{j-1}+l$ zero matrices of appropriate dimensions (in the same place as $z^{j,l}$ in $\begin{bmatrix}x^T & z^T & \left(x^{i_0}\right)^T\end{bmatrix}^T$).

Proposition 4.2: Consider a set of strategies in the form (7) and (8) and suppose that this set of strategies is consistent. Then it constitute a Nash equilibrium.

Proof:

The strategy of the major player is optimal due to the Proposition 5.1 of [17]. The strategy of any minor player is optimal due to the Proposition 4.1 of [17]. □

Remark 4.3: The consistency conditions (12), (14) and (24 - 28) are Riccati equations with two types of coupling. The first type of coupling is though the A matrices according to equations (12) and (14) and has the same nature as the coupled Riccati equations of the LQ games [27 section 3], [24]. The second is through the Λ variables and has the same nature as the interconnected Riccati equations in the study of LQ control of MJLS [1]. □

5. Numerical Example

In this section an algorithm for solving the consistency conditions is developed. The algorithm is applied to a simple example involving a major player and one type of minor players having time horizon 2. The dependence of the feedback gains on the Markov chain state variable is then numerically studied for several values of the parameters.

The algorithm initially guesses a value for the feedback gains. With the assumed values it computes the matrices for the optimal control problems. Then the optimal control problems are solved and new feedback gains are computed. The new feedback gains are used to compute the system matrices and solve the optimal control problems and so on. The algorithm is the following:

Algorithm 5.1:

Step 1: Take an initial guess for the vectors $L^{1M}(k,y)$, $L^{1}(1,l,l',y)$, $\overline{L}^{1}(k,y)$, $L^{MM}(y)$ and $L^{M}(1,l,y)$ for $y=1,...,M$ $k=0,1$ and $l,l'=0,1$.

Step 2: Compute the matrices $\tilde{A}^{1}(k,y)$ for $y=1,...,M$ $k=0,1$ using (14).

Step 3: Compute $L^{1M}(k,y)$, $L^{1}(1,l,l',y)$ and $\overline{L}^{1}(k,y)$ using (19 - 23) and (26 - 28).

Step 4: Compute the matrix $\tilde{A}^{M}$ using (12)

Step 5: Set $L^{MM}_{old}(y)=L^{MM}(y)$ and $L^{M}_{old}(1,l,y)=L^{M}(1,l,y)$ for $y=1,...,M$ $k=0,1$ and $l=0,1$.

Step 6: Compute matrices $K(y),\Lambda(y)$ for $y=1,...,M$ to satisfy (15) and (16)

Step 7: Use (24) and (25) to update the values of $L^{MM}(y)$ and $L^{M}(1,l,y)$.

Step 8: If $\left(\sum_{y=1}^{M}\left\|L^{MM}_{old}(y)-L^{MM}(y)\right\|^{2}+\sum_{y=1}^{M}\sum_{l=0}^{1}\left\|L^{M}_{old}(1,l,y)-L^{M}(1,l,y)\right\|^{2}\right)^{1/2}$ is small enough then halt. Else go to step 2. □

Remark 5.2: Step 6 may be implemented in several ways. Probably the simpler is to use the value iteration algorithm described in the proof of proposition A.1.4.1. Other numerical methods involve [1]. □

Let us then apply the algorithm to the following example:

Example 5.3: In this example major and minor players have an one dimensional state equation. The time horizon of each of the minor players is 2 and the maximum number of minor players participating in the game at some time step is 4. At each time step either one or two new minor players enter the game. Thus the entrance dynamics is described by a Markov chain with maximum number of active players 4 and 4 states: $(1/4,1/4),(1/4,2/4),(2/4,1/4)$ and $(2/4,2/4)$. We use the enumeration $(1/4,1/4)\to 1$, $(1/4,2/4)\to 2$, $(2/4,1/4)\to 3$ and $(2/4,2/4)\to 4$. The transition matrix of the Markov chain is given by:

$$\Pi=\begin{bmatrix}0.9 & 0 & 0.1 & 0\\ 0.2 & 0 & 0.8 & 0\\ 0 & 0.3 & 0 & 0.7\\ 0 & 0.8 & 0 & 0.2\end{bmatrix}$$

The dynamic equation of the major player is given by:

$$x^{M}(k+1)=x^{M}(k)+\frac{c_1}{4}\sum_{i\in I_k}x^{i}(k)+u^{M}(k)+w^{M}(k)$$

and the dynamic equation for a minor player is given by:

$$x^{i}(k+1)=x^{i}(k)+c_1x^{M}(k)+\frac{c_1}{4}\sum_{j\in I_k}x^{j}(k)+u^{i}(k)+w^{i}(k)$$

where by c_1 we denote all the coupling coefficients. Thus the parameters of the state equations are given by: $A^{M}=1$, $F^{M1}=c_1$, $B^{M}=1$, $A^{1}=1$, $G^{1}=c_1$, $F^{11}=c_1$, and $B^{1}=1$.

The Q matrices are given by $Q(1)=Q(1)=Q(3)=I_3$ and $Q(4)=(1+c_2)I_3$. Assume also that the matrices $Q^1(y)$ are given by: $Q_f^1(y)=Q^1(y)=I_4$ for $y=1,...,4$ and the R matrices are all units. The discount factor is $a=0.95$.

For example if $c_1=c_2=1$, after 20 steps of the algorithm, the gain matrices change less than 10^{-12}. The gain matrices for the major player are given by:

$L^{MM}(1)=-0.6461$, $L^{MM}(2)=-0.6982$, $L^{MM}(3)=-0.7862$, $L^{MM}(4)=-0.7301$

$L^{M}(0,1)=-0.8668$, $L^{M}(0,2)=-0.9867$, $L^{M}(0,3)=-0.9836$, $L^{M}(0,4)=-0.9361$

$L^{M}(1,1)=-0.6461$, $L^{M}(1,2)=-0.6982$, $L^{M}(1,3)=-0.7862$, $L^{M}(1,4)=-0.7301$

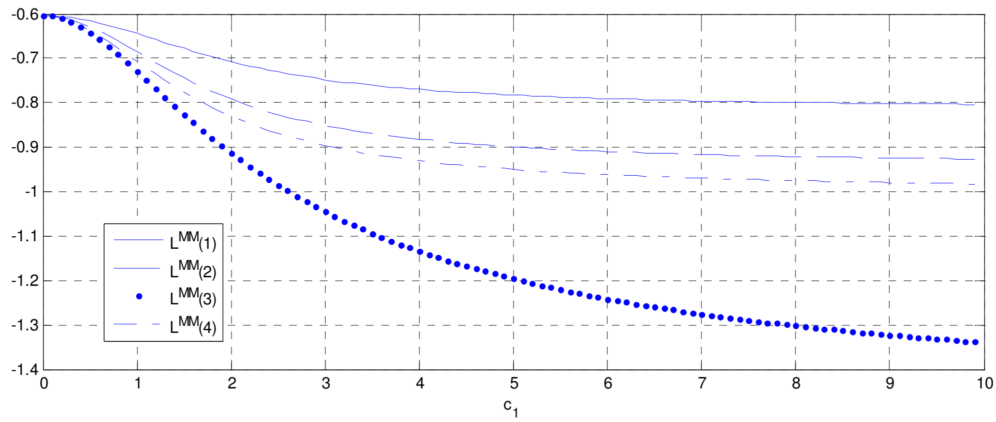

Figure 5.1: Dependence of the feedback gains on the coupling coefficients

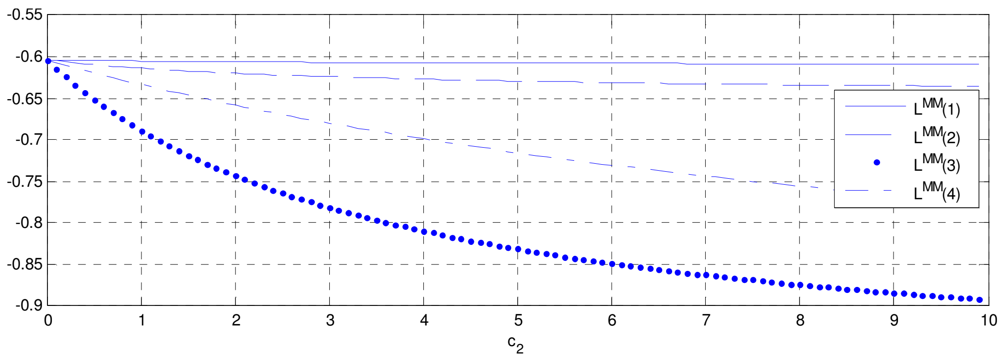

Figure 5.2: Dependence of the feedback gains on the Q matrices.

We then study the dependence of the feedback gains on the Markov chain state space for several values of c_1, c_2. The Algorithm 5.1 is applied for $c_1 \in [0,10]$ and $c_2=0$ as well as for $c_2 \in [0,10]$ and $c_1=0$. The results are presented in the Figures 5.1 and 5.2 respectively.

The numerical results illustrate that the feedback gains depend on the Markov chain state variable due to two reasons. The first is the coupling in the dynamic equations and the second is the dependence of the Q matrices on y. □

6. Large Population Case

In this section we use a mean filed approximation in order to study games with a very large number of players. This approach involves the statement of some approximate optimal control problems that correspond to the limit of those in section 3 as the scale variable tends to infinity i.e. the number of new minor players at each time step tends to infinity. The Markov chain with a large number of states is approximated by a Markov process with a continuum of states and thus a notion of convergence of Markov processes is first recalled in section 6.1. Then the solution of the approximate optimal control problems for the major and minor players is characterized by appropriate Riccati equations and consistency conditions analogous of those of section 4 are stated. Finally it is proved that a set of feedback strategies satisfying those consistency conditions, constitute an ε–Nash equilibrium for a game with a very large number of players.

Another motivation for the use of the continuous approximation is computational. The state space of the Markov chain that describes the random entrance grows fast as the maximum number of players increases. For example if there is only one type of minor players that has a time horizon 5 and the new minor players in each step belong to the set $\{1,2,...,N\}$ then the state space of the Markov chain describing the entrance has N^5 points. Thus the equations characterizing a Nash equilibrium (Ded. 4.1) depend on many parameters and thus are very complicated. On the other hand in several cases the situation is much simplified using the continuous approximation (example 6.3 and section 6.4).

6.1 Convergence of stochastic kernels

The approximation of the Markov chain with a Markov process with a continuum of sates is based on the notion of weak convergence of Markov processes (Def 6.2). The continuous state space of the Markov process is defined as a subset of $\mathbb{R}^n$ that contains the values of the Markov chain.

The dimension of the vector y_k as defined by (1) depends on the time horizon of each of the types of players involved in the game. Particularly it holds $\dim(y_k) = T_1 + ... + T_p = T_{tot}$. Thus the state space of the Markov chain is a subset of:

$$D = \left\{ \left(y_1, ..., y_{T_{tot}} \right) \in \mathbb{R}^{T_{tot}} : \sum_{i=1}^{T_{tot}} y_i \le 1, y_i \ge 0, i = 1, ..., T_{tot} \right\} \tag{29}$$

The Markov chain may be described using the notion of a stochastic kernel.

Definition 6.1: Let $D' = \{d_1, ..., d_M\} \subset D$ and $P = [p_{ij}]$ a $M \times M$ stochastic matrix. The stochastic kernel that corresponds to the Markov chain with state space D' and transition matrix P is defined as:

$$\bar{K}(y, B) = \Pr(y_{k+1} \in B \mid y_k = y) = \sum_{j: d_j \in B} p_{ij}$$

where $i = \min\left\{ \arg\min_l \left\{ \| y - d_l \| \right\} \right\}$, $y \in D$ and B a Borel subset of D. □

A notion of convergence of stochastic kernels is then recalled form [16] and a notion of continuity of stochastic kernels from [19].

Definition 6.2: (i) We shall say that a sequence of stochastic kernels $\bar{K}_v$ converges weekly to a stochastic kernel $\bar{K}$ if for any sequence y_v of elements of D converging to an element y of D and any continuous function g it holds:

$$\int_D g(y')\bar{K}_\nu(y_\nu, y') \to \int_D g(y')\bar{K}(y, y'). \tag{30}$$

(ii) A stochastic kernel $\bar{K}$ is called Feller continuous if $\bar{K}(y_\nu, \cdot) \Rightarrow \bar{K}(y, \cdot)$ when $y_\nu \to y$. □

Let us turn back to games described by the relationships (2–6) and a large number of minor players. To do so we consider a sequence of games with increasing number of minor players. Let us denote the ν-th game by g^ν and its parameters by $x^{M,\nu}, x^{i,\nu}, u^{i,\nu}, w^{m,\nu}, w^{i,\nu}, s_c^\nu, \tilde{z}^\nu, J^{M,\nu}, J^{i,\nu}$ in the place of $x^M, x^i, u^i, w^m, w^i, s_c, \tilde{z}, J^M, J^i$. For the scale variable it holds: $s_c^\nu \to \infty$ as $\nu \to \infty$. The state of the Markov chain describing the entrance is denoted by y_k^ν, the number of states of the Markov chain by M^ν and the corresponding stochastic kernel is denoted by $\bar{K}^\nu$. Conclusions about the final part of this sequence of games are obtained under the assumption that the stochastic kernels $\bar{K}^\nu$ converge weakly to a stochastic kernel $\bar{K}$ of a Feller continuous Markov process. The stochastic kernel $\bar{K}$, thus, approximates the final part of the sequence of Markov chains. We also suppose that the matrices $Q(y_i^\nu), Q_f^j(y_i^\nu), Q^j(y_i^\nu)$ for $i = 1, ..., M^\nu$ are samples of continuous matrix functions $Q(y), Q_f^j(y), Q^j(y)$, for $y \in D$.

Example 6.3: Consider games involving one type of minor players of time horizon 2. At each time step each one of ν players enters the game with probability p. Thus the new minor players at each step follow a binomial distribution. The entrance dynamics is thus described by the Markov chain $y_k^\nu = \left[N_k^{1,0,\nu} \quad N_k^{1,1,\nu}\right] / s_c^\nu$, $s_c^\nu = 2\nu$ and it holds:

$$\Pr\left(N_{k+1}^\nu = i\right) = \binom{\nu}{i} p^i (1-p)^{\nu-i}.$$

Let us denote by $\bar{K}^\nu$ the stochastic kernel describing the Markov chain y_k^ν.

The random variable $N_k^{1,0,\nu} / s_c$ converges weakly to the value $p\nu$. Thus the Markov chain y_k^ν may be approximated by a Markov process which has a stochastic kernel:

$$\bar{K}((y_1, y_2), \cdot) = \delta(p/2, y_1),$$

where δ denotes the Dirac measure and it holds $\bar{K}^\nu \to \bar{K}$ weakly. This example shows that for the large number of players case the approximate description of the Markov process in several cases may be much simpler than the original □

6.2 Approximate Optimal Control Problems

In this subsection an approximate optimal control problem for each one of the players involved in the game is stated. We suppose that every player assumes that the other players use feedback strategies in the form (7) and (8) as well as that the game has an infinite scale i.e. the Markov process has the approximate stochastic kernel and the scale variable s_c has an infinite value. The approximate optimal control problems are, thus, stated in terms of the approximate dynamics for which, for simplicity, we use again the symbols $x^M, x^i, u^i, w^m, w^i, \tilde{z}$.

<u>Approximate optimal control problem for the major player:</u>

The optimal control problem is stated in terms of the state vector $\left[\left(x^M\right)^T \quad z^T\right]^T$. The evolution of this vector is given by (11). The optimal control problem for the major player is the same as in section 3 except that $y_{k+1} \sim \bar{K}(y_k, \cdot)$. The solution of the problem is given in terms of the following quantities [17]:

$$K(y) = Q(y) + \left(\tilde{A}^M\right)^T (y) \left(a\Lambda(y) - a\Lambda(y)^T \tilde{B}^M \left(R/a + \left(\tilde{B}^M\right)^T \Lambda(y) \tilde{B}^M \right)^{-1} \left(\tilde{B}^M\right)^T \Lambda(y) \right) \tilde{A}^M(y) \quad (31)$$

$$\Lambda(y) = E\left[K(y_{k+1}) \mid y_k = y \right] = \int_D K(y') \bar{K}(y, dy') \quad (32)$$

The policy given by:

$$u_k = -\left(\left(\tilde{B}^M\right)^T \Lambda(y_k) \tilde{B}^M + R/a \right)^{-1} \left(\tilde{B}^M\right)^T \Lambda(y_k) \tilde{A}^M(y_k) \left[\left(x^M(k)\right)^T \quad \bar{z}^T(k) \right]^T \quad (33)$$

is optimal provided that with this control law the cost is finite. A sufficient condition for the finiteness of the cost is derived in terms of the spectral radius [3] of a certain operator $T = T_{\tilde{A}_{cl}^M, \bar{K}}$ defined in the Appendix A.1 where $\tilde{A}_{cl}^M(y)$ is the closed loop matrix:

$$\tilde{A}_{cl}^M(y) = \tilde{A}^M(y) - \tilde{B}^M \left(\left(\tilde{B}^M\right)^T \Lambda(y) \tilde{B}^M + R/a \right)^{-1} \left(\tilde{B}^M\right)^T \Lambda(y) \tilde{A}^M(y).$$

<u>Approximate optimal control problem for a minor player:</u>

Consider a minor player i_0 with entrance time t_{i_0}. The evolution of the quantities x^M, $\tilde{z}$ and x^{i_0} may be described by:

$$\begin{bmatrix} x^M(k+1) \\ \tilde{z}(k+1) \\ x^{i_0}(k+1) \end{bmatrix} = \tilde{A}^{j_0}\left(k - t_{i_0}, y_k\right) \begin{bmatrix} x^M(k) \\ \tilde{z}(k) \\ x^{i_0}(k) \end{bmatrix} + \tilde{B}^{j_0} u^{i_0}(k) + W^{i_0}(k) \quad (34)$$

where: $j_0 = Z_{i_0}$ and

$\tilde{A}^{j_0}_{x^M x^M}\left(k - t_{i_0}, y_k\right) = A^M + B^M L^{MM}(y_k)$, $\tilde{A}^{j_0}_{x^M z^{j,l}}\left(k - t_{i_0}, y_k\right) = F^{Mj} + B^M L^M(j, l, y_k)$

$\tilde{A}^{j_0}_{z^{j,l+1} x^M}\left(k - t_{i_0}, y_k\right) = y_k^{j,l}\left(G^j + B^j L^{jM}(l, y_k)\right)$ for $l+1 \geq 1$,

$\tilde{A}^{j_0}_{z^{j,0} x^M}\left(k - t_{i_0}, y_k\right) = 0$

$$\tilde{A}^{j_0}_{z^{j,l+1} z^{j',l'}}\left(k - t_{i_0}, y_k\right) = \left(A^j + y_k^{j,l} B^j \bar{L}^j(l, y_k)\right)\delta_{j,j'}\delta_{l',l} + y_k^{j,l}\left(F^{jj'} + B^j L^j(j', l', l, y_k)\right) \quad (35)$$

$\tilde{A}^{j_0}_{z^{j,0} z^{j',l'}}\left(k - t_{i_0}, y_k\right) = \tilde{A}^{j_0}_{x^M x^{i_0}}\left(k - t_{i_0}, y_k\right) = \tilde{A}^{j_0}_{z^{j,l} x^{i_0}}\left(k - t_{i_0}, y_k\right) = 0$,

$\tilde{A}^{j_0}_{x^{i_0} x^M}\left(k - t_{i_0}, y_k\right) = G^{Z_{i_0}}$ $\tilde{A}^{i_0}_{x^{i_0} z^{j,l}}\left(k - t_{i_0}, y_k\right) = F^{j,j'}$ $\tilde{A}^{j_0}_{x^{i_0} x^{i_0}}\left(k - t_{i_0}, y_k\right) = A^{Z_{i_0}}$,

$\tilde{B}^{j_0} = \begin{bmatrix} 0 \\ 0 \\ B^{Z_{i_0}} \end{bmatrix}$ and $y_k^{j,l} = ye_{T_1 + \ldots + T_{j-1} + l + 1}$

We observe that the matrix $\tilde{A}^{j_0}$ does not depend on time as well as that it has simpler expressions than (14).

The solution may be computed recursively using the following relations [17]:

$$K^{j_0}_{T_{j_0}}(y) = Q_f^{j_0}(y) \quad (36)$$

$$\Lambda_{k+1}^{j_0}(y)=E\left[K_{k+1}^{j_0}\left(y_{k+1+t_{i_0}}\right)\mid y_{k+t_{i_0}}=y\right]=\int_D K_{k+1}^{j_0}(y')\bar{K}(y,dy') \tag{37}$$

$$\begin{aligned}K_k^{j_0}(y)=Q(y)+\left(\tilde{A}^{j_0}(y)\right)^T\Big[\Lambda_{k+1}^{j_0}(y)-&\\ -\Lambda_{k+1}^{j_0}(y)\tilde{B}^{j_0}\left(R^{j_0}+\left(\tilde{B}^{j_0}\right)^T\Lambda_{k+1}^{j_0}(y)\tilde{B}^{j_0}\right)^{-1}\left(\tilde{B}^{j_0}\right)^T\Lambda_{k+1}^{j_0}(y)\Big]\tilde{A}^{j_0}(y)&\end{aligned} \tag{38}$$

The optimal control is then given by:

$$u_{k+t_{i_0}}^{i_0}=L_k^{j_0}\left(y_{k+t_{i_0}}\right)\left[\left(x^M\left(k+t_{i_0}\right)\right)^T \quad \tilde{z}^T\left(k+t_{i_0}\right) \quad \left(x^{i_0}\left(k+t_{i_0}\right)\right)^T\right] \tag{39}$$

where:

$$L_k^{j_0}(y)=-\left(R_{j_0}+\left(\tilde{B}^{i_0}\right)^T\Lambda_{k+1}(y)\tilde{B}^{i_0}\right)^{-1}\left(\tilde{B}^{i_0}\right)^T\Lambda_{k+1}(y)\tilde{A}^{i_0}(y) \tag{40}$$

6.3 Consistency conditions and ε–Nash equilibrium

Consider a set of strategies in the form that is given by (7) and (8). We first give consistency conditions for the approximate optimal control problems.

Definition 6.4: Consider the set of feedback strategies (7) and (8) assume that depend continuously y. Compute the matrix $\tilde{A}^M$ and the set of matrix functions $\tilde{A}^j(y)$ given by (12) and (35) for $j=1,...,p$ and $y\in D$. Then the set of feedback strategies is called approximately consistent if:

(a) There exist matrix functions $K(y),L(y)$, $y\in D$ satisfying (31) and (32) and moreover it holds:

$$L^{MM}(y)=-\left(\left(\tilde{B}^M\right)^T\Lambda(y)\tilde{B}^M+R/a\right)^{-1}\left(\tilde{B}^M\right)^T\Lambda(y)\tilde{A}^M(y)\begin{bmatrix}I_{n_M} & 0 & \dots & 0\end{bmatrix}^T \tag{41}$$

and

$$L^M(j,l,y)=-\left(\left(\tilde{B}^M\right)^T\Lambda(y)\tilde{B}^M+R/a\right)^{-1}\left(\tilde{B}^M\right)^T\Lambda(y)\tilde{A}^M(y)\begin{bmatrix}0 & \dots & 0 & I_{n_j} & 0 & \dots & 0\end{bmatrix}^T \tag{42}$$

for $j=1,...,p$ and $y\in D$. In (42), I_{n_j} appears after $1+T_1+...+T_{j-1}+l$ zero matrices of appropriate dimensions (in the same place as $z^{j,l}$ in $\begin{bmatrix}x^T & \tilde{z}^T\end{bmatrix}^T$).

(b) The operator $T=T_{\tilde{A}_{cl}^M,\bar{K}}$ has a spectral radius less that $1/a$.

(c) For any type $j_0\in\{1,...,p\}$ if the vectors $L_k^{j_0}(y)$ for $y\in D$ and $k=0,...,T_j-1$ computed by (19 - 23) satisfy:

$$L^{j_0M}(k,y)=L_k^{j_0}(y)\begin{bmatrix}I_{n_M} & [0 \;\; \dots \;\; 0] & 0\end{bmatrix}^T \tag{43}$$

$$L^{j_0}(j,l,k,y)=L_k^{j_0}(y)\begin{bmatrix}0 & [0 \;\; \dots \;\; 0 \;\; I_{n_j} \;\; 0 \;\; \dots \;\; 0] & 0\end{bmatrix}^T \tag{44}$$

$$\bar{L}^{j_0}(k,y)=L_k^{j_0}(y)\begin{bmatrix}0 & [0 \;\; \dots \;\; 0] & I_{n_{j_0}}\end{bmatrix}^T \tag{45}$$

where I_{n_j} in (44) appears after $1+T_1+...+T_{j-1}+l$ zero matrices of appropriate dimensions (in the same place as $z^{j,l}$ in $\begin{bmatrix}x^T & z^T & \left(x^{i_0}\right)^T\end{bmatrix}^T$). □

A set of approximately consistent policies is applied a game with a large number of players g^{ν}. This set of policies constitute an ε-Nash equilibrium i.e. for any player, given the strategies of other players, its strategy has a cost at most ε far from the optimal provided that the closed loop system is mean square stable. This property is illustrated by the following Theorem 6.5 and its Corollary 6.6. Let us denote by $J^{M,\nu}\left(\pi_{(7),(8)}\right)$ and $J^{i,\nu}\left(\pi_{(7),(8)}\right)$ the value of the cost functions using the policies (7) and (8) for the major player and player i respectively. Denote also by $\left(J^{M,\nu}\left(\pi_{(7),-M}\right)\right)^*$ the optimal value of the cost function for the major player when the minor players apply the policy given by (7) and $\left(J^{i,\nu}\left(\pi_{(7),(8),-i}\right)\right)^*$ is the optimal value of the cost function for the minor player i when the major player uses strategy (8) and other minor players use the strategy given by (7).

Theorem 6.5: Consider a set of feedback strategies given by (7) and (8) and assume that it is approximately consistent. Then for any positive ε, there exist a positive integer ν_0 such that:

$$J^{M,\nu}\left(\pi_{(7),(8)}\right) \le \left(J^{M,\nu}\left(\pi_{(7),-M}\right)\right)^* + \varepsilon$$

$$J^{i,\nu}\left(\pi_{(7),(8)}\right) \le \left(J^{i,\nu}\left(\pi_{(7),(8),-i}\right)\right)^* + \varepsilon\left(1 + E\left[\left\|\left[\left(x^M\right)^T(t_i) \quad \tilde{z}^T(t_i) \quad \left(x^{i_0}\right)^T(t_i)\right]\right\|^2\right]\right)$$

for every $\nu \ge \nu_0$.

Structure of the proof:
The proof of the second inequality is based on the fact that the optimal policies for a minor player involve continuous functions of the state vector and Markov chain and properties of the weak convergence.
A basic step in the proof of the first inequality is given in Appendix A.2 where it is shown that some stability properties of MJLS are preserved under weak convergence. It is then shown that the final part of the series involved in the costs is small in some sense, uniformly in the initial conditions, and thus it suffices to compare finite series. The result for finite series is similar to the proof of the second inequality.
The detailed proof is relegated to Appendix. More general results are first shown in section A.3 and particularly in Propositions A.3.1 and A.3.2. Theorem 6.5 is then proved as a consequence of these Propositions in section A.4. □

Corollary 6.6: If in addition of the assumptions of Theorem 6.5, the spectral radius of the operator $T = T_{\tilde{A}_{cl}^M, \bar{K}}$ is less than 1 i.e. it holds $r(T) < 1$, then the set of strategies given by (7) and (8) constitute an ε-Nash equilibrium for large ν i.e.: for any positive ε, there exist a positive integer ν_0 such that:

$$J^{M,\nu}\left(\pi_{(7),(8)}\right) \le \left(J^{M,\nu}\left(\pi_{(7),-M}\right)\right)^* + \varepsilon$$

$$J^{i,\nu}\left(\pi_{(7),(8)}\right) \le \left(J^{i,\nu}\left(\pi_{(7),(8),-i}\right)\right)^* + \varepsilon$$

for every $\nu \ge \nu_0$.

Proof: The first inequality is the same as in Theorem 6.5. For the proof of the second inequality we apply the stability results from Corollary A.2.3. □

Remark 6.7: Approximate consistency conditions involve nonlinear matrix integral equations and in general are not simpler than the consistency conditions of Section 4. However in several cases the situation is extremely simplified as illustrated in the example of the next section. □

6.4 Numerical Example

This example considers a game with a large number of minor players and no major player. There is only one type of minor players. Each one of the minor players $i \in \Lambda$ has time horizon two and has a dynamic equation of the form:

$$x^i(k+1) = x^i(k) + \sum_{i \in I_k} x^i / s_c + u^i(k) + w_k .$$

At each time step, each one of ν minor players tosses a fair coin and with probability 1/2 enters the game. The cost function is given by (6), the vector $\tilde{x}^i(k)$ by $\tilde{x}^i(k) = \left[\tilde{z}^T(k) \quad (x^i)^T(k) \right]^T$ and the matrices by: $Q_f = 4y^1 I_3$, $Q = 12yy^T I_3$ and $R = 1$ respectively where we use the notation Q_f, Q, R, y^1, y^0 instead of $Q_f^1, Q^1, R^1, y^{1,1}, y^{1,0}$. The scale variable has a value $s_c = 2\nu$. The approximate description of the Markov chain when $\nu \to \infty$ is given by the stochastic kernel:

$$\bar{K}((y_1, y_2), \cdot) = \delta_{(1/4, y_1)}(\cdot),$$

where δ is the Dirac measure. We first find a policy that satisfies the approximate consistency conditions. Due to the absence of a major player the consistency conditions involve (35 - 38), (40), (44) and (45). The unknown quantities may be expressed in terms of the functions $L_1(y) = L^1(1,0,0,y)$, $L_2(y) = L^1(1,1,0,y)$, $L_3(y) = L^1(1,0,1,y)$, $L_4(y) = L^1(1,1,1,y)$, $L_5(y) = \bar{L}^1(0,y)$ and $L_6(y) = \bar{L}^1(1,y)$.

Due to the form of the approximate Markov process the integral in the equation (36) is simplified as following:

$$\Lambda_{k+1}((y_1, y_2)) = E\left[K_{k+1}\left(y_{k+1+t_{i_0}}\right) | \, y_{k+t_{i_0}} = y \right] = \int_D K_{k+1}(y') \bar{K}(y, dy') = K_{k+1}((\bar{y}, y_1))$$

where $\bar{y} = 1/4$.

Thus the form of the Markov process implies a decoupling in the consistency conditions. Particularly for $y = (\bar{y}, \bar{y})$ the consistency conditions do not depend on other values of y. Writing the consistency conditions for some $y = (\bar{y}, y_1)$ the equations depend only on $L_1((\bar{y}, y_1)), \ldots, L_6((\bar{y}, y_1))$ and $L_1((\bar{y}, \bar{y})), \ldots, L_6((\bar{y}, \bar{y}))$. Furthermore for some $y = (y_1, y_2)$ the consistency equations depend only on $L_1((y_1, y_2)), \ldots, L_6((y_1, y_2))$ and $L_1((\bar{y}, y_1)), \ldots, L_6((\bar{y}, y_1))$. This structure of the consistency conditions suggests the following procedure: Compute the values of $L_1((\bar{y}, \bar{y})), \ldots, L_6((\bar{y}, \bar{y}))$ solving a system of six equations with six unknowns. Then for each $y \in D$, in the form $y = (\bar{y}, y_1)$ the values of $L_1(y), \ldots, L_6(y)$ may be computed as a solution of six equations with six unknown variables involving $L_1(y), \ldots, L_6(y)$ and $L_1((\bar{y}, \bar{y})), \ldots, L_6((\bar{y}, \bar{y}))$. Finally for some $(y_1, y_2) = y \in D$ the values of $L_1(y), \ldots, L_6(y)$ may be computed as a solution of six equations with six unknown variables involving $L_1(y),....,L_6(y)$ and $L_1((\bar{y},y_1)),....,L_6((\bar{y},y_1))$.

The solution for $y=(\bar{y},\bar{y})$ is given by:

$L_1(y)=-0.8133$, $L_2(y)=-0.68$, $L_3(y)=-0.5$, $L_4(y)=-0.5$, $L_5(y)=-0.6667$ and $L_5(y)=-0.5$.

As an example to compute the values of $L_1(y),....,L_6(y)$ for $y=[0.2 \quad 0.6]$, we first compute $L_1(y'),....,L_6(y')$ for $y'=[\bar{y} \quad 0.2]$.

$L_1(y')=-0.8133$, $L_2(y')=-0.68$, $L_3(y')=-0.4444$, $L_4(y')=-0.4444$, $L_5(y')=-0.6667$ and $L_5(y')=-0.4444$.

The values of $L_1(y),....,L_6(y)$ are thus given by:

$L_1(y)=-0.7543$, $L_2(y)=-0.66$, $L_3(y)=-0.4444$, $L_4(y)=-0.4444$, $L_5(y)=-0.6261$ and $L_5(y)=-0.4444$.

The behavior of the system with the control law obtained is described by the closed loop matrix $\tilde{A}_{cl}^M(y)$. It holds $y_k=\bar{y}$ a.s. for $k\geq 3$. On $y=(\bar{y},\bar{y})$, $\tilde{A}_{cl}^M(y)$ is:

$$\tilde{A}_{cl}^M(y)=\begin{bmatrix} 0 & 0 \\ 0.88 & 0.08 \end{bmatrix}.$$

It holds $\left|\lambda_{\max}\left(\tilde{A}_{cl}^M(y)\right)\right|=0.08<1$. Thus operator $T=T_{\tilde{A}_{cl}^M,\bar{K}}$ has spectral radius less than one and the control laws obtained are approximately consistent. Thus for any $\varepsilon>0$ there exist an $\nu_0\in\mathbb{N}$ such that the control laws obtained constitute an ε Nash equilibrium for any $\nu\geq\nu_0$.

Remark 6.8: When there is an independent random entrance, the approximate consistency conditions are decoupled. The solutions to the approximate optimal control problems could, thus, be obtained taking into advance this special form. □

7. Conclusion and future work

Games with a major player and many minor players of several time horizons, a random number of which enters at each time step were considered. Sufficient conditions for a set of symmetric linear feedback strategies to constitute a Nash equilibrium were derived. For the large number of players case, a notion of convergence of Markov processes was used in order to approximate a Markov chain which has a large number of states with a Markov process with a continuum of states. Approximate optimal control problems are then stated. The form of approximate control problems is Linear Quadratic problems for a MJLS with a Markov process with a continuous state space. A set of symmetric linear feedback strategies that is optimal for the approximate system is proved to constitute an ε-Nash equilibrium when the scale is sufficiently large and under some additional stability conditions. An example of a game with a large number of players is also studied where the use of mean field approximation simplifies the analysis.

Some possible directions of future work involve extensions to the case that there exists also a random exit or the case where the players have the choice to enter or leave the game and the study of algorithms that solve the equations for the consistency conditions for the finite as well as the mean field case. Existence and uniqueness of a Nash equilibrium questions for games in the form studied in this paper are clearly of interest.

Appendix

The Appendix contains the proof of Theorem 6.5 and some propositions needed to prove the result. In the first section A.1 we recall some results from [17] about the stability of MJLS. Section A.2 studies the properties of a sequence of MJLS systems when the sequence of Markov chains converges weakly to a Feller continuous limit. The basic result of section A.2 is Proposition A.2.1 which shows that if the limit system is stable then a tail of the sequence of systems consists of stable systems. Section A.3 proves that policies optimal for the limit system are ε-optimal for a tail of the sequence. The basic results are Proposition A.3.1 and Proposition A.3.2, where the result is proved for the finite and infinite horizon problems respectively. In Section A.4 the proof of theorem 6.5 is completed.

A.1 MJLS with general state space

In this section we recall some results for MJLS with general state space. The proofs of the results as well as further details could be found in [17].

Let D be a compact set and $\bar{K}(\cdot,\cdot)$ a stochastic kernel on D. Consider a system in the form:

$x_{k+1} = A(y_k)x_k, \quad y_{k+1} \sim \bar{K}(y_k,\cdot)$

The exponential mean square stability of a system in this form is equivalent to the fact that the spectral radius of an operator $T = T_{A,\bar{K}}$ is less than 1. The operator is defined using the following quantities $P_k : \Re(D) \to R^{n\times n}$ where $\Re(D)$ is the set of Borel measurable subsets of D and $P_k(C) = E\left[x_k x_k^T \chi_{y_k \in C}\right]$ for any C Borel measureable subset of D. The operator is defined such that $P_{k+1} = TP_k$. In the case of a Markov chain with finite state space, the operator T takes the form of the matrix defined by (18).

The exponential mean square stability is a uniform on the initial conditions. Thus it is equivalent to the existence of a constant $a \in (0,1)$ and a positive integer k_0 such that $E\left[x_{k_0}^T x_{k_0}\right] < a x_0^T x_0$ for any x_0, y_0 non-random initial conditions. Furthermore it is equivalent to the existence of positive constants $M > 0$ and $0 < a < 1$ such that $E\left[x_k^T x_k\right] < M a^k E\left[x_0^T x_0\right]$ for any initial conditions.

A.2 Weak convergence and mean square stability

Consider a sequence of systems:

$$x_{k+1}^{\nu} = A(y_k)x_k, \quad y_{k+1}^{\nu} \sim \bar{K}^{\nu}(y_k^{\nu},\cdot) \tag{A1}$$

and a limit system:

$$x_{k+1} = A(y_k)x_k, \quad y_{k+1} \sim \bar{K}(y_k,\cdot). \tag{A2}$$

Suppose that $\bar{K}^{\nu} \to \bar{K}$ weakly, $\bar{K}$ is Feller continuous and that for the limit system it holds $r\left(T_{A,\bar{K}}\right) < 1$. Suppose also that $A(\cdot)$ is a continuous matrix function. Finally assume that (A2) is exponentially stable. It will be shown that (A1) is also exponentially stable for large ν.

For any $a \in (0,1)$, there exist an integer k such that:

$$E\left[x_k^T x_k\right] < aE\left[x_0^T x_0\right] \tag{A3}$$

for any x_0, y_0 initial conditions. Choosing x_0 to be any non-random initial condition the last inequality may be written as:

$$x_0^T E\left[A^T(y_0)...A^T(y_{k-1})A^T(y_{k-1})...A^T(y_0)\right]x_0 < ax_0^T x_0.$$

Thus it holds:

$$aI - E\left[A^T(y_0)...A^T(y_{k-1})A^T(y_{k-1})...A^T(y_0)\right] > 0. \tag{A4}$$

The positive definiteness of the matrix in (A4) is equivalent due to the Sylvester criterion to a set of inequalities in the form:

$$f_j\left(E\begin{bmatrix}\bar{f}_1(y_0,...,y_{k-1}) \\ \vdots \\ \bar{f}_{n^2}(y_0,...,y_{k-1})\end{bmatrix}\right) > 0 \tag{A5}$$

for $j = 1,...,n$, where $\bar{f}_i$ $i = 1,...,n^2$ correspond to the elements of the matrix in (A4) and are continuous and f_j are multinomials arising using Sylvester criterion. Inverse process shows that conditions (A5) imply a relationship in the form of (A3) and thus imply $r\left(T_{A,\bar{K}}\right) < 1$.

Proposition A.2.1: Under the assumptions stated above, $r\left(T_{A,\bar{K}}\right) < 1$ implies the existence of ν_0 such that: $r\left(T_{A,\bar{K}^\nu}\right) < 1$ for every $\nu \ge \nu_0$.

Before proving the proposition a lemma will be stated. This lemma illustrates a uniformity property of the weak convergence. The uniformity is expressed in terms of the Bounded Lipschitz metric [23 section 17] which is defined by:

$$\beta(P_1, P_2) = \sup\left\{\left|\int f dP_1 - \int f dP_2\right| : \|f\|_{BL} \le 1\right\}$$

where $\|f\|_{BL} = \sup_{y \in D}\{f(y)\} + \inf\{L : f \text{ is } L\text{- Lipschitz}\}$. D is separable and thus $\beta(\cdot,\cdot)$ metrizes the weak convergence of probability measures [23 section 17].

To state the lemma consider the functions:

$$\Xi^\nu, \Xi : (D, \|\ \|) \to \left(\prod(D^k), \beta\right)$$

where $\prod(D^k)$ is the space of probability measures on D^k and for any $A \in \Re(D^k)$ has the form $\left(\Xi^\nu(y)\right)(A) = \Pr\left((z_0,...,z_{k-1}) \in A\right)$ where z_0 has the a distribution concentrated to y and $z_{i+1} \sim K^\nu(z_i,\cdot)$. In the same way the values of $\left(\Xi(y)\right)(A)$ are defined. Thus Ξ maps the initial condition y_0 to the distribution of $(y_0, y_1, ..., y_{k-1})$.

We may show that Ξ is continuous. To do so we observe that it holds $\bar{K} \to \bar{K}$ weakly. Thus $y^\nu \to y$ implies $\Xi(y^\nu) \Rightarrow \Xi(y)$ [16]. Therefore Ξ is continuous. The next Lemma A.2.2 shows a uniformity property on the convergence of Ξ^ν to Ξ. Particularly it is shown that for sufficiently large ν, the distribution of $\left(y_0^\nu, y_1^\nu, ..., y_{k-1}^\nu\right)$ is close to $(y_0, y_1, ..., y_{k-1})$ for $y_0^\nu = y_0$, uniformly in y_0.

Lemma A.2.2: For any positive constant ε there exist a positive integer ν_0 such that $\beta\left(\Xi^\nu(y),\Xi(y)\right)<\varepsilon$ for any $\nu \ge \nu_0$ and any $y \in D$.

Proof: To contradict suppose that there exist a positive constant ε such that for any $\nu_0 \in \mathbb{N}$, there exist a $\nu \ge \nu_0$ with $\beta\left(\Xi^\nu(y),\Xi(y)\right)>\varepsilon$. Then there exist sequences m_ν, y_{m_ν} such that $m_\nu \ge \nu$, $m_\nu \ge m_{\nu-1}$ and $\beta\left(\Xi^{m_\nu}\left(y_{m_\nu}\right),\Xi\left(y_{m_\nu}\right)\right)>\varepsilon$. y_{m_ν} is a sequence on a compact set and thus there exist a converging subsequence $y_{m_{\nu_l}} \to \bar{y}$. Theorem 1 of [16] implies $\beta\left(\Xi^{m_{\nu_l}}\left(y_{m_{\nu_l}}\right),\Xi(\bar{y})\right)\to 0$. However triangle inequality implies:

$$\begin{aligned}\beta\left(\Xi^{m_{\nu_l}}\left(y_{m_{\nu_l}}\right),\Xi(\bar{y})\right) &\ge \beta\left(\Xi^{m_{\nu_l}}\left(y_{m_{\nu_l}}\right),\Xi\left(y_{m_{\nu_l}}\right)\right)-\beta\left(\Xi\left(y_{m_{\nu_l}}\right),\Xi(\bar{y})\right)\\ &>\varepsilon-\beta\left(\Xi\left(y_{m_{\nu_l}}\right),\Xi(\bar{y})\right)\end{aligned}$$

Continuity of Ξ implies that $\beta\left(\Xi^{m_{\nu_l}}\left(y_{m_{\nu_l}}\right),\Xi(\bar{y})\right)>\varepsilon/2$ which contradicts $\beta\left(\Xi^{m_{\nu_l}}\left(y_{m_{\nu_l}}\right),\Xi(\bar{y})\right)\to 0$. □

Proof of Proposition A.2.1: The quantities:

$$g_j(y_0)=f_j\left(E\begin{bmatrix}\bar{f}_1(y_0,\dots,y_{k-1})\\ \vdots\\ \bar{f}_{n^2}(y_0,\dots,y_{k-1})\end{bmatrix}\right)$$

may be seen as continuous functions of y_0. Thus, due to the compactness of D there exist a positive constant ε_1 such that:

$g_j(y_0)>\varepsilon_1$,

for any $y_0 \in D$. The functions f_j are uniformly continuous. Thus there exist a constant $\delta_1>0$ such that $f_j(v_1)>\varepsilon_1$ implies $f_j(v_2)>0$ for any $v_2 \in D^k$ such that $\|v_1-v_2\|<\delta_1$ and any $j=1,\dots,n$.

Choose $y_0=y_0^\nu$. Then

$$E\begin{bmatrix}\bar{f}_1(y_0,\dots,y_{k-1})\\ \vdots\\ \bar{f}_{n^2}(y_0,\dots,y_{k-1})\end{bmatrix} \text{ and } E\begin{bmatrix}\bar{f}_1(y_0^\nu,\dots,y_{k-1}^\nu)\\ \vdots\\ \bar{f}_{n^2}(y_0^\nu,\dots,y_{k-1}^\nu)\end{bmatrix}$$

may be written in the form:

$$\begin{bmatrix}\int \bar{f}_1(w)\left(\Xi(y_0)\right)(dw)\\ \vdots\\ \int \bar{f}_{n^2}(w)\left(\Xi(y_0)\right)(dw)\end{bmatrix} \text{ and } \begin{bmatrix}\int \bar{f}_1(w)\left(\Xi^\nu(y_0)\right)(dw)\\ \vdots\\ \int \bar{f}_{n^2}(w)\left(\Xi^\nu(y_0)\right)(dw)\end{bmatrix}.$$

Let us first prove the claim: For large ν it holds:

$$\left|\int \bar{f}_i(w)\left(\Xi(y_0)\right)(dw)-\int \bar{f}_i(w)\left(\Xi^\nu(y_0)\right)(dw)\right|<\delta_1/n^2$$

for any $y_0 \in D$ and $i=1,\ldots,n^2$.

To prove the claim recall that any uniformly continuous function may be approximated by a Lipschitz one. Let $\bar{f}_i': D^k \to \mathbb{R},\ i=1,\ldots,n^2$ be Lipschitz functions such that $\left\|\bar{f}_i'-\bar{f}_i\right\|_\infty<\delta_1/\left(4n^2\right)$. Denote by $\bar{L}$ the maximum bounded Lipschitz norm of the functions $\bar{f}_i'$ i.e $\max_{i=1,\ldots,n^2}\left\|\bar{f}_i'\right\|_{BL}$. Then

$$\begin{aligned}\left|\int \bar{f}_i(w)\left(\Xi(y_0)\right)(dw)-\int \bar{f}_i(w)\left(\Xi^\nu(y_0)\right)(dw)\right| &\le \int\left|\bar{f}_i(w)-\bar{f}_i'(w)\right|\left(\Xi(y_0)\right)(dw)+\\ &+\left|\int \bar{f}_i(w)\left(\Xi(y_0)\right)(dw)-\int \bar{f}_i(w)\left(\Xi^\nu(y_0)\right)(dw)\right|+\int\left|\bar{f}_i(w)-\bar{f}_i'(w)\right|\left(\Xi^\nu(y_0)\right)(dw)\\ &\le\left|\int \bar{f}_i(w)\left(\Xi(y_0)\right)(dw)-\int \bar{f}_i(w)\left(\Xi^\nu(y_0)\right)(dw)\right|+\delta_1/\left(2n^2\right)\end{aligned}$$

Appling Lemma A.2.2 an integer ν_0 may be found such that:

$$\beta\left(\Xi^\nu(y),\Xi(y)\right)<\delta_1/\left(2n^2\bar{L}\right)$$

for any $\nu \ge \nu_0$ and any $y_0 \in D$. This completes the proof of the claim.

Thus for $\nu \ge \nu_0$ it holds:

$$f_j\left(E\begin{bmatrix}\bar{f}_1\left(y_0^\nu,\ldots,y_{k-1}^\nu\right)\\ \vdots\\ \bar{f}_{n^2}\left(y_0^\nu,\ldots,y_{k-1}^\nu\right)\end{bmatrix}\right)>0$$

for any $y_0^\nu \in D$ and $j=1,\ldots,n$. Thus for any $y_0^\nu \in D$ it holds:

$$aI-E\left[A^T\left(y_0^\nu\right)\ldots A^T\left(y_{k-1}^\nu\right)A^T\left(y_{k-1}^\nu\right)\ldots A^T\left(y_0^\nu\right)\right]>0.$$

Therefore:

$$E\left[\left(x_k^\nu\right)^T x_k^\nu \mid x_0^\nu, y_0^\nu\right]<aE\left[\left(x_0^\nu\right)^T x_0^\nu \mid x_0^\nu, y_0^\nu\right].$$

Integrating over the distribution of x_0^ν, y_0^ν we conclude:

$$E\left[\left(x_k^\nu\right)^T x_k^\nu\right]<aE\left[\left(x_0^\nu\right)^T x_0^\nu\right].$$

Thus $r\left(T_{A,\bar{K}^\nu}\right)<1$ for any $\nu \ge \nu_0$. □

Corollary A.2.3: Consider the systems described by (A1) and (A2). Assume that $\bar{K}^\nu \to \bar{K}$ weakly and $\bar{K}$ is Feller continuous. Let $a<1$ and suppose that $T=T_{A,\bar{K}}$ has spectral radius less than $1/a$. Then for any $\varepsilon>0$, there exist positive integers k_0,ν_0 such that:

$$E\left\{\sum_{k=k_0}^{\infty}a^k\left[\left(x_k^\nu\right)^T x_k^\nu\right]\right\}<\varepsilon_1\left(E\left[\left(x_0^\nu\right)^{\mathrm{T}} x_0^\nu\right]+1\right)$$

and:

$$E\left\{\sum_{k=k_0}^{\infty} a^k\left[\left(x_k\right)^T x_k\right]\right\}<\varepsilon_1\left(E\left[x_0^T x_0\right]+1\right)$$

Proof: Consider the operators $T=T_{A,\bar{K}}$ and $T_\nu=T_{A,\bar{K}_\nu}$. By the spectral formula and proposition A.2.1, there exist positive integers m_0 and ν_0 such that $\left\|T^{m_0}\right\|<\left(\frac{1}{a}-\varepsilon\right)^{m_0}$ such that $\left\|T_\nu^{m_0}\right\|<\left(\frac{1}{a}-\varepsilon\right)^{m_0}$ for some $\varepsilon>0$ and any $\nu\geq\nu_0$. Denote by $c_1,\ldots,c_{m_0-1}$ some constants such that $\left\|T_\nu^i\right\|\leq c_i>1$, for $i=1,\ldots,m_0-1$. Thus

$$\left\|T_\nu^i\right\|\leq c_1\ldots c_{m_0-1}\left(\frac{1}{a}-\varepsilon\right)^{m_0[i/m_0]}\leq c_1\ldots c_{m_0-1}\frac{1}{a^{m_0}}\left(\frac{1}{a}-\varepsilon\right)^i.$$

Denote by $C=c_1\ldots c_{m_0-1}\frac{1}{a^{m_0}}$. Let $P_k^\nu(B)=E\left[x_k^\nu\left(x_k^\nu\right)^T\chi_{y_k\in B}\right]$. Then it holds:

$$\left\|P_{k+1}^\nu\right\|\leq\left\|T_\nu^{k+1}\right\|\left\|P_0^\nu\right\|+\|W\|\sum_{i=1}^{k+1}\left\|T_\nu^i\right\|\leq C\left\|P_0^\nu\right\|\left(\frac{1}{a}-\varepsilon\right)^{k+1}+\|W\|C\frac{1-(1/a-\varepsilon)^{k+1}}{1+\varepsilon-1/a}.$$

Therefore it holds:

$$\begin{aligned}\sum_{k=k_0}^{\infty}a^k\left\|P_k^\nu\right\|&\leq\sum_{k=k_0}^{\infty}\left[C\left\|P_0^\nu\right\|(1-a\varepsilon)^k+\frac{\|W\|C}{1+\varepsilon-1/a}(1-a\varepsilon)^k+\frac{\|W\|C}{1+\varepsilon-1/a}a^k\right]\\&\leq C\left\|P_0^\nu\right\|\frac{(1-a\varepsilon)^{k_0}}{a\varepsilon}+\frac{\|W\|C}{1+\varepsilon-1/a}\frac{(1-a\varepsilon)^{k_0}}{a\varepsilon}+\frac{\|W\|C}{1+\varepsilon-1/a}\frac{a^{k_0}}{1-a}\end{aligned}$$

On the other hand $E\left[\left(x_0^\nu\right)^T x_0^\nu\right]\leq\left\|P_0^\nu\right\|$. Thus:

$$\sum_{k=k_0}^{\infty}a^k\left\|P_k^\nu\right\|\leq\left(C\frac{(1-a\varepsilon)^{k_0}}{a\varepsilon}\right)E\left[\left(x_0^\nu\right)^T x_0^\nu\right]+\left(\frac{\|W\|C}{1+\varepsilon-1/a}\frac{(1-a\varepsilon)^{k_0}}{a\varepsilon}+\frac{\|W\|C}{1+\varepsilon-1/a}\frac{a^{k_0}}{1-a}\right)$$

Each of the terms $\left(C\frac{(1-a\varepsilon)^{k_0}}{a\varepsilon}\right)$ and $\left(\frac{\|W\|C}{1+\varepsilon-1/a}\frac{(1-a\varepsilon)^{k_0}}{a\varepsilon}+\frac{\|W\|C}{1+\varepsilon-1/a}\frac{a^{k_0}}{1-a}\right)$ could become smaller than ε_1 for sufficiently large k_0. □

A.3 Convergence of stochastic kernels and ε-Optimality

In the following suppose that $\bar{K}^\nu\to\bar{\mathrm{K}}$ weakly, $\bar{\mathrm{K}}$ is Feller continuous and the functions $A^\nu(y,k),A(y,k),A(y)$ are continuous on the y argument.

Notation: Consider the system:

$x_{k+1} = A'(y_k, k)x_k + B(y_k)u_k + w_k, \qquad y_{k+1} \sim \bar{K}'(y_k, \cdot),$

and the feedback control law $u_k = L_k(y_k)x_k$. Then we denote by:

$$J_{\bar{K}',k_0,A',u_k=L_k(y_k)x_k}(x_0, y_0) = E\left[x_{k_0+1}^T Q_{k_0+1}(y_{k_0+1})x_{k_0+1} + \sum_{k=0}^{k_0} a^k x_k^T \left[L_k^T(y_k)R(y_k)L_k(y_k) + Q(y_k)\right]x_k\right]$$

and $J^*_{\bar{K}',A',k_0}(x_0, y_0)$ the optimal value, where $\bar{K}'$ may take the values $\bar{K}$ or $\bar{K}^\nu$. The time horizon is allowed to take the infinity value. We use the notation $J_{\bar{K}',k_0,u_k=L_k(y_k)x_k}(x_0, y_0)$ and $J^*_{\bar{K}',k_0}(x_0, y_0)$ for $A'(y_k, k) = A(y_k)$. □

The basic topic of this section is the proof of the following two propositions about ε-optimality in the finite and infinite horizon case respectively.

Proposition A.3.1: Let $A^\nu(y_k, k)$ for each k be a sequence of continuous matrix functions. Suppose that $A^\nu(y_k, k) \to A(y_k, k)$ as $\nu \to \infty$. Denote by $u_k = L_k(y_k)x_k$ the optimal feedback strategy that attains $J_{\bar{K},k_0,A(y_k,k),u_k=L_k(y_k)x_k} = J^*_{\bar{K},k_0}$. Then for any $\varepsilon > 0$ there exists a positive integer ν_0 such that:

$$J_{\bar{K}^\nu,k_0,A^\nu(y_k,k),u_k=L_k(y_k)x_k}(x_0, y_0) < J^*_{\bar{K}^\nu,A(y_k,k),k_0} + \varepsilon\left(x_0^T x_0 + 1\right)$$

for any $\nu \geq \nu_0$. □

Proposition A.3.2: Denote by $u_k = L(y_k)x_k$ the optimal feedback strategy that attains $J_{\bar{K},\infty,u_k=L(y_k)x_k} = J^*_{\bar{K},\infty}$ and suppose that is continuous on y_k. Then for every $\varepsilon > 0$ there exist a positive integer ν_0 such that $J_{\bar{K}^\nu,\infty,u_k=L(y_k)x_k}(x_0, y_0) \leq J^*_{\bar{K}^\nu,\infty}(x_0, y_0) + \varepsilon\left(x_0^T x_0 + 1\right)$

for any $\nu \geq \nu_0$. □

The proof of propositions A.3.1 and A.3.2 depends on the following lemmas.

Lemma A.3.3: Consider the feedback control law: $u_k = L_k(y_k)x_k$. Then for any $\varepsilon > 0$, there exist a positive integer ν_0 such that:

$$\left|J_{\bar{K},k_0,u_k=L_k(y_k)x_k}(x_0, y_0) - J_{\bar{K}^\nu,k_0,u_k=L_k(y_k)x_k}(x_0, y_0)\right| \leq \varepsilon\left(x_0^T x_0 + 1\right).$$

for any $\nu \geq \nu_0$.

Proof: Denoting by $\tilde{y} = \begin{bmatrix} y_0 & y_1 & \ldots \end{bmatrix}$ and $\tilde{y}^\nu = \begin{bmatrix} y_0^\nu & y_1^\nu & \ldots \end{bmatrix}$ with straightforward calculation we may compute:

$$J_{\bar{K},k_0,u_k=L_k(y_k)x_k}(x_0, y_0) = x_0^T\left(E\left[\tilde{a}_1(\tilde{y})\right]\right)x_0 + E\left[\tilde{a}_2(\tilde{y})\right].$$

and:

$$J_{\bar{K}^\nu,k_0,u_k=L_k(y_k)x_k}(x_0, y_0) = x_0^T\left(E\left[\tilde{a}_1(\tilde{y}^\nu)\right]\right)x_0 + E\left[\tilde{a}_2(\tilde{y}^\nu)\right],$$

where functions $\tilde{a}_1(\cdot)$ and $\tilde{a}_2(\cdot)$ are continuous functions of the variables: $\left(y_0, y_1, \ldots, y_{k_0}\right)$ and $\left(y_0^\nu, y_1^\nu, \ldots, y_{k_0}^\nu\right)$ respectively. Thus they are continuous functions $\tilde{a}_1, \tilde{a}_2 : (D, \Im_D)^{\mathbb{N}} \to \mathbb{R}$ with the product topology. Thus [16] implies that $E\left[\tilde{a}_1(\tilde{y}^\nu)\right] \to E\left[\tilde{a}_1(\tilde{y})\right]$ and

$E\left[\tilde{a}_2\left(\tilde{y}^\nu\right)\right] \to E\left[\tilde{a}_2\left(\tilde{y}\right)\right]$. Therefore for any $\varepsilon > 0$, there exist a positive integer ν_0 such that $\left\| E\left[\tilde{a}_1\left(\tilde{y}^\nu\right)\right] - E\left[\tilde{a}_1\left(\tilde{y}\right)\right] \right\| < \varepsilon$ and $\left| E\left[\tilde{a}_2\left(\tilde{y}^\nu\right)\right] - E\left[\tilde{a}_2\left(\tilde{y}\right)\right] < \varepsilon \right|$ for any $\nu \geq \nu_0$. Thus for any $\nu \geq \nu_0$ it holds:

$$\left| J_{\bar{K},k_0,u_k=L_k(y_k)x_k}\left(x_0, y_0\right) - J_{\bar{K}_\nu,k_0,u_k=L_k(y_k)x_k}\left(x_0, y_0\right) \right| \leq \varepsilon\left(x_0^T x_0 + 1\right) \qquad \square$$

Lemma A.3.4: Let $f_\nu : D \to \mathbb{R}^l$ a sequence of continuous functions and $f : D \to \mathbb{R}^l$ a continuous function such that: $f_\nu\left(y\right) \to f\left(y\right)$ for any $y \in D$. Then it holds:

$$\int_D f_\nu\left(y'\right)\bar{K}^\nu\left(y,dy'\right) \to \int_D f\left(y'\right)\bar{K}\left(y,dy'\right)$$

Proof: Since D is compact, the convergence is uniform. Thus for any $\varepsilon > 0$, there exist a positive integer ν_{01} such that $\left\| f_\nu\left(y\right) - f\left(y\right) \right\| < \varepsilon / 2$ for any $y \in D$ and $\nu \geq \nu_{01}$. Thus for $\nu \geq \nu_{01}$ it holds:

$$\left| \int_D f_\nu\left(y'\right)\bar{K}^\nu\left(y,dy'\right) - \int_D f\left(y'\right)\bar{K}\left(y,dy'\right) \right| \leq \frac{\varepsilon}{2} + \left| \int_D f\left(y'\right)\bar{K}^\nu\left(y,dy'\right) - \int_D f\left(y'\right)\bar{K}\left(y,dy'\right) \right|$$

Weak convergence implies the existence of a positive integer ν_{02} such that:

$$\left| \int_D f\left(y'\right)\bar{K}^\nu\left(y,dy'\right) - \int_D f\left(y'\right)\bar{K}\left(y,dy'\right) \right| < \varepsilon / 2$$

for any $\nu \geq \nu_{02}$. Choose $\nu_0 = \max\left\{\nu_{01}, \nu_{02}\right\}$. $\square$

Lemma A.3.5: For any $\varepsilon > 0$, there exist a positive integer ν_0 such that:

$$\left| J^*_{\bar{K},k_0}\left(x_0, y_0\right) - J^*_{\bar{K}^\nu,k_0}\left(x_0, y_0\right) \right| \leq \varepsilon\left(x_0^T x_0 + 1\right).$$

for any $\nu \geq \nu_0$.

Proof: The optimal costs may be computed recursively using the relations (10–14) in [17]. The computation involve the functions:, $K_k^\nu\left(y_k\right)$, $K_k\left(y_k\right)$, $\Lambda_{k+1}^\nu\left(y_k\right)$, $\Lambda_{k+1}\left(y_k\right)$, $c_k^\nu\left(y_k\right)$, $c_k\left(y_k\right)$, $J^*_{\bar{K}^\nu,k_0}\left(x_0, y_0\right)$, $J^*_{\bar{K},k_0}\left(x_0, y_0\right)$. The functions involved are continuous on their arguments. Thus using Lemma A.3.4 we may show inductively that $\Lambda_{k+1}^\nu\left(y_k\right) \to \Lambda_{k+1}\left(y_k\right)$ as $\nu \to \infty$, thus $K_k^\nu\left(y_k\right) \to K_k\left(y_k\right)$ and $c_k^\nu\left(y_0\right) \to c_k\left(y_0\right)$. Thus for any $\varepsilon > 0$, there exist a positive integer ν_0 such that: $\left| c_k^\nu\left(y_0\right) - c_k\left(y_0\right) \right| < \varepsilon$ and $\left\| K_k^\nu\left(y_k\right) - K_k\left(y_k\right) \right\| < \varepsilon$ for any for any $\nu \geq \nu_0$. Thus

$$\left| J^*_{\bar{K}^\nu,k_0}\left(x_0, y_0\right) - J^*_{\bar{K},k_0}\left(x_0, y_0\right) \right| < \varepsilon\left(x_0^T x_0 + 1\right) \qquad \square$$

Proof of proposition A.3.1:

Lemma A.3.3 implies that there exists a positive integer ν_{01} such that:

$$J_{\bar{K}^\nu,k_0,A^\nu,u_k=L_k(y_k)x_k}\left(x_0, y_0\right) < J_{\bar{K},k_0,A,u_k=L_k(y_k)x_k}\left(x_0, y_0\right) + \varepsilon\left(x_0^T x_0 + 1\right)/2 =$$
$$= J^*_{\bar{K},A,k_0}\left(x_0, y_0\right) + \varepsilon\left(x_0^T x_0 + 1\right)/2$$

for any $\nu \geq \nu_{01}$. Lemma A.3.5 implies that there exists a positive integer ν_{02} such that:

$$J^*_{\bar{K},A,k_0}\left(x_0, y_0\right) < J^*_{\bar{K}^\nu,A^\nu,k_0}\left(x_0, y_0\right) + \varepsilon\left(x_0^T x_0 + 1\right)/2$$

for any $\nu \geq \nu_{02}$. Thus choosing $\nu_0 = \max\{\nu_{01}, \nu_{02}\}$ we have:

$$J_{\bar{K}^\nu, k_0, A^\nu, u_k = L_k(y_k)x_k}(x_0, y_0) < J^*_{\bar{K}^\nu, A, k_0} + \varepsilon\left(x_0^T x_0 + 1\right). \qquad \square$$

Proof of Proposition A.3.2:

The proof is based on a series of comparisons:

$$J_{\bar{K}^\nu, \infty, u_k = L(y_k)x_k} \xrightarrow{1} J_{\bar{K}^\nu, k_0, u_k = L(y_k)x_k} \xrightarrow{2} J_{\bar{K}, k_0, u_k = L(y_k)x_k} \xrightarrow{3} J^*_{\bar{K}, \infty} \xrightarrow{4} J^*_{\bar{K}, k_0'} \xrightarrow{5} J^*_{\bar{K}^\nu, k_0'} \xrightarrow{6} J^*_{\bar{K}^\nu, \infty}$$

Set $Q_{k_0+1} = 0$. For comparison 1, using continuity of K, Corollary A.2.3 imply that for any $\varepsilon > 0$ there exist positive integers k_0 and ν_{01} such that:

$$J_{\bar{K}^\nu, \infty, u_k = L(y_k)x_k}(x_0, y_0) \leq J_{\bar{K}^\nu, k_0, u_k = L(y_k)x_k}(x_0, y_0) + \varepsilon\left(x_0^T x_0 + 1\right)/4,$$

for any $\nu \geq \nu_{01}$.

For comparison 2, Lemma A.2.3 implies the existence of a positive integer ν_{02} such that:

$$J_{\bar{K}^\nu, k_0, u_k = L(y_k)x_k}(x_0, y_0) < J_{\bar{K}, k_0, u_k = L(y_k)x_k}(x_0, y_0) + \varepsilon\left(x_0^T x_0 + 1\right)/4$$

Comparison 3 holds with $\leq$ i.e. it holds:

$$J_{\bar{K}, k_0, u_k = L(y_k)x_k}(x_0, y_0) \leq J^*_{\bar{K}, \infty}(x_0, y_0)$$

To obtain an inequality for comparison 4, we observe that $K_k \to K$ and $c_k \to c$ uniformly as $k \to \infty$, where K_k and c_k as in the proof of the Proposition 5.1 of [17]. We also have: $J^*_{\bar{K}, k}(x_0, y_0) = x_0^T K_k(y_0) x_0 + c_k(y_0)$ and $J^*_{\bar{K}, \infty}(x_0, y_0) = x_0^T K(y_0) x_0 + c(y_0)$. Thus for any $\varepsilon > 0$, there exist a positive integer k_0' such that $J^*_{\bar{K}, \infty}(x_0, y_0) < J^*_{\bar{K}, k}(x_0, y_0) + \varepsilon\left(1 + x_0^T x_0\right)/4$.

Using this value k_0' Lemma A.3.5 implies the existence of a positive integer ν_{03} such that

$$J^*_{\bar{K}, k_0'}(x_0, y_0) < J^*_{\bar{K}^\nu, k_0'}(x_0, y_0) + \varepsilon\left(x_0^T x_0 + 1\right)/4.$$

Comparison 6 holds with $\leq$ i.e.

$$J^*_{\bar{K}_\nu, k_0'}(x_0, y_0) \leq J^*_{\bar{K}_\nu, \infty}(x_0, y_0).$$

Thus choosing appropriate k_0, k_0' and $\nu_0 = \max\{\nu_{01}, \nu_{02}, \nu_{03}\}$ we have:

$$J_{\bar{K}_\nu, \infty, u_k = L(y_k)x_k}(x_0, y_0) \leq J^*_{\bar{K}_\nu, \infty}(x_0, y_0) + \varepsilon\left(x_0^T x_0 + 1\right)$$

for any $\nu \geq \nu_0$. $\square$

A.4 Proof of Theorem 6.5

At first observe that optimal control problems involve continuous matrix functions. For the first inequality apply proposition A.3.2. For the proof of the second inequality observe that $\tilde{A}^{j_0,\nu}\left(k - t_{i_0}, y_k\right) \to \tilde{A}^{j_0}\left(k - t_{i_0}, y_k\right)$ as $\nu \to \infty$. Then apply proposition A.3.1 and take expectations. $\square$

References

[1] H. Abou-Kandil, G. Freiling and G. Jank "On the solution of discrete-time Markovian jump linear quadratic control problems", Automatica, Vol. 31, pp. 765-768, 1995.

[2] I. Arribas, A. Urbano "Repeated games with probabilistic horizon", Mathematical Social Sciences, Vol. 50, Iss. 1, pp. 39–60, 2005.

[3] W. Arveson “A Short Course on Spectral Theory” Springer-Verlag, New York, 2002.
[4] R. J. Aumann “Markets with a Continuum of Traders”, Econometrica, Vol. 32, pp. 39 - 50, 1964.
[5] Y. Balasko and K. Shell “The Overlapping-Generations Model, I: The case of Pure Exchange without Money”, Journal of Economic Theory, Vol. 23, pp. 281-306, 1980.
[6] T. Basar and G. J. Olsder “Dynamic Noncooperative Game theory”, SIAM 2nd edition 1999.
[7] D. Fudenberg , J. Tirole “Game Theory”, MIT press, 1991.
[8] D. Fudenberg, D. M. Kreps, and E. S. Maskin, “Repeated games with long-run and short-run players”, Review of Economic Studies, vol 57, no. 4, pp. 555-573, 1990.
[9] O. Gueant, J.M. Lasry and P.L. Lions, “Mean field games and applications”, Springer, Paris-Princeton Lectures on Mathematical Finance, 2010.
[10] M. Huang, R. P. Malhamé and P. E. Caines “Nash Equilibria for Large-Population Linear Stochastic Systems of Weakly Coupled Agents”, Analysis, Control and Optimization of Complex Dynamical Systems, Chapter 9, pp. 215-252, Springer, New York, 2005.
[11] M. Huang, R.P. Malhamé, and P.E. Caines "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle", Communications in Information and Systems, vol. 6, no. 3, pp. 221-251, 2006.
[12] M. Huang. “Large-population LQG games involving a major player: the Nash certainty equivalence principle”, SIAM J. Control Optimization, vol. 48, no. 5, pp. 3318-3353, 2010.
[13] S. Jørgensen and D.W.K. Yeung “Inter- and intragenerational renewableresource extraction”, Annals of Operations Research, Vol. 88, 1999.
[14] S. Jørgensen, DWK Yeung “An Overlapping Generations Stochastic Deferential Game”, Automatica, Vol. 41, pp. 69–74, 2005.
[15] N.Ch. Kakogiannis and G.P. Papavassilopoulos “Games with Players Having Long Term and Short term Horizons”, 14th International Symposium of Dynamic Games, Bannf Center, Bannf, Alberta, Canada, 19-23 July, 2010.
[16] A.F. Karr “Weak Convergence of a Sequence of Markov Chains”, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete. Vol 33 pp. 41-48, 1975.
[17] I. Kordonis and G.P. Papavassilopoulos “On Stability and Linear Quadratic Control of MJLS with General State Space” submitted to the IEEE AC.
[18] D. Kreps and R. Wilson “Reputation and Imperfect Information”, Journal of Economic Theory, vol 27 pp 253-279, 1982.
[19] S. Meyn and R. L. Tweedie “Markov Chains and Stochastic Stability”, 2nd edition, Cambridge University Press, 2009.
[20] R. Myerson “Population Uncertainty and Poisson Games”, International Journal of Game Theory, Vol. 27, Iss. 3, Nov. 1998.
[21] J. W. Milnor and L. S. Shapley “Values of Large Games II: Oceanic Games”, Mathematics of Operations Research Vol. 3, No. 4, Nov. 1978.
[22] G. Owen “Game Theory”, W.B. Saunders Company, 1969.
[23] D. Panchenko “18.175 Theory of Probability fall-2008”, section 17 (Massachusetts Institute of Technology: MIT OpenCourseWare), http://ocw.mit.edu (Accessed March 10, 2008). License: Creative Commons BY-NC-SA.
[24] G.P. Papavassilopoulos J.V. Medanic and J.B. Cruz “On the Existence of Nash Strategies and Solutions of Riccati Equations in Linear Quadratic Games” Journal of Optimization Theory and Applications, vol. 42 1984.
[25] G.P. Papavassilopoulos “University Games”, submitted for publication
[26] P. Samuelson “An exact consumption loan model of interest with or without the social contrivance of money”, Journal of Political Economy, 66(6), Dec.1958.
[27] A. W. Starr Y.C. Ho “Nonzero Sum Differential Games”, Journal of Optimization Theory and Applications vol. 3 no. 3, 1969.
[28] D. W. K. Yeung and L. A. Petrosyan “Subgame Consistent Cooperative Solution of Dynamic Games with Random Horizon” Journal of Optimization Theory and Applications, Vol. 150, No. 1, 2011