# LQ Nash Games With Random Entrance: An Infinite Horizon Major Player and Minor Players of Finite Horizons

Ioannis Kordonis, *Member, IEEE*, and George P. Papavassilopoulos, *Senior Member, IEEE*

***Abstract*—We study Dynamic Games with randomly entering players, staying in the game for different lengths of time. Particularly, a class of Discrete Time Linear Quadratic (LQ) Games, involving a major player who has an infinite time horizon and a random number of minor players is considered. The number of the new minor players, entering at some instant of time, is random and it is described by a Markov chain. The problem of the characterization of a Nash equilibrium, consisting of Linear Feedback Strategies, is reformulated as a set of coupled finite and infinite horizon LQ optimal control problems for Markov Jump Linear Systems (MJLS). Sufficient conditions characterizing Nash equilibrium are then derived. The problem of Games involving a large number of minor players is then addressed using a Mean Field (MF) approach and asymptotic $\varepsilon$—Nash equilibrium results are derived. Sufficient conditions for the existence of a MF Nash equilibrium are finally stated.**

***Index Terms*—Game theory, Markov jump linear systems (MJLS), random entrance, stochastic optimal control.**

## I. Introduction

FOR the 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 instant of time are quite structured. For example, in finite or infinite horizon dynamic games (e.x. [1], [2]) all the players participate in the game for identical time intervals. In overlapping generation games ([3]–[5]), a known number of players of a new generation enters into 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 [6] the number of players that participate in the game is not known a priori. Games with random horizon have been studied in a repeated game setting in [7] and in a differential game setting in [8]. In this class of games, the time intervals in which the players are involved in the game are identical; however, the duration of the game is random. In [9], a game with overlapping generations involving players which remain in the game for two time steps is considered. The number of players of each generation is however random.

The current work studies games with random entrance in a LQ setting and imposes less structure on the time intervals during which the players participate in the game, as well as on the number of the active players at each time step. In particular we, consider a player with infinite time horizon, called the major player and many players with finite time horizons, called minor players. The number of new players entering the game at any time step is a random variable that has a distribution which depends only on the number of active players at that moment. The random entrance is, thus, described by a Markov chain. The problem considered here is the characterization of the Perfect Nash equilibria. After that, we study the case where the number of minor players is very large. A Mean Field (MF) approximation is used to characterize strategies, which are asymptotically optimal as the number of new minor players in each step tends to infinity. The equations derived using MF approximation are often much easier.

The structure proposed for the participation of the minor players in the game is not unusual in practice. There are several examples of game situations where there is a long living agent or institution which, 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 instance, 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. Another example is a liberalized energy market in which there is a public power corporation with an infinite time horizon and many renewable energy producers that have a permission to enter the system for a certain amount of time [10]. A third example is University-Student Games [11], where the students of each semester stand for the minor players and the university as a major player. Cases involving players with different time horizons were studied also in [12], [13]. Other examples involve the study of repeated games with long-run and short run-players [14], such as the chain store game and the study of reputation effects (ex. [15], [16]).

The interest for the games with large number of players is not new. In [17], games with a continuum of players, called oceanic games, were introduced and a value for such games

Manuscript received July 12, 2013; revised February 27, 2014; accepted November 18, 2014. Date of publication January 29, 2015; date of current version May 21, 2015. This work was supported 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 and the Action: ARISTIA, Project: HEPHAESTUS: Towards Next Generation Intelligent Energy Systems which is implemented under the Operational Programme "Education and Lifelong Learning" which is co-funded by the European Union (European Social Fund) and Greece (National Sources). Recommended by Associate Editor D. Bauso.

The authors are with the School of Electrical and Computer Engineering, National Technical University of Athens, Athens 106 82, Greece (e-mail: jkordonis1920@yahoo.com; yorgos@netmode.ntua.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2015.2396642

was defined. Models with a continuum of players were also studied in [18] (see also [19], Ch. X). Dynamic Games with a continuum of players in discrete time were studied in [20] and [21], under the name Anonymous Sequential Games and in [22] a Mean Field approximation in Discrete Time Dynamic Games was used and the Oblivious Equilibrium notion was defined. Recently, the Mean Field approach in the study of games with large number of players was introduced [23], [24]. The closely related methodology of Nash Certainty Equivalence was recently developed in order to derive asymptotic Nash equilibrium results as the number of players tends to infinity [25], [26]. Linear Quadratic Games with large number of players were studied in [27]–[29]. Risk Sensitive and Robust Mean Field Games were analyzed in [30] and [31]. A quite general setting is analyzed in [32] using Operator Theoretic techniques. An LQG game involving a major player and a large number of minor players of infinite time horizon is considered in [33] where asymptotically optimal decentralized feedback strategies are characterized. Connections among the discrete and continuous time, as well as several generalizations are studied in [34].

In the current work, the problem of random entrance is 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. There are two types of coupling; the first corresponds to the Markov Jump character of the optimal control problems and the second to the LQ Game coupling. In the case of a large number of players, the Mean Field approach involves the statement of approximate optimal control problems assuming an infinity of players. In that case, $\varepsilon$—equilibrium results are proved. The method used to prove the $\varepsilon$—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 are also of independent interest.

The rest of the paper is organized as follows: In Section II, the dynamics and the cost functions of major and minor players are defined. In Section III, the optimal control problems that the participants of the game face are reformulated as a set of coupled finite and infinite horizon LQ problems for MJLS. In Section IV, sufficient conditions on a set of linear feedback strategies to constitute a Nash equilibrium are derived. In Section V, the problem with a large number of players is approximated using a Mean Field model. Then some $\varepsilon$—Nash equilibrium results are obtained. In Section VI, an algorithm for computing a Nash equilibrium is stated and it is shown to converge under certain conditions. Furthermore, some numerical examples are studied. In Section VII, we conclude. The proofs of some results in the text are relegated to the Appendix. The notation used in the main text is, also, summarized in the Appendix.

*Notation:* The transpose of a matrix is denoted by $\cdot^T$. In all the text, except Section VI-B, $\|\cdot\|$ denotes the usual 2-norm. The underlying probability space is denoted by $(\Omega, \mathcal{F}, Pr)$ and the spectral radius of a matrix or an operator by $r(\cdot)$. The Borelian subsets of a set $D$ are denoted by $\mathcal{B}(D)$. The notion of a stochastic kernel is also used to describe the evolution of a Markov chain. Particularly, for a Markov chain $y_k$ with state space $D$, we denote by $\bar{K}(\cdot,\cdot) : D \times \mathcal{B}(D) \to \mathbb{R}$ the stochastic kernel, i.e., $\bar{K}(y, B) = Pr(y_{k+1} \in B | y_k = y)$, for $y \in D$ and $B \in \mathcal{B}(D)$. The fact that the random variable $y$ has probability distribution $F$ is denoted by $y \sim F$ and the weak convergence of probability measures is denoted by "$\Rightarrow$." The Kronecker delta $\delta_{ij}$ is also used, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. A matrix function $A : D \to \mathbb{R}^{n\times n}$ is called strictly positive definite if there exists a positive constant $c$ such that $A(y) > cI$ for any $y \in D$. Finally, the $i, j$ element of a matrix $A$ is denoted by $A^{i,j}$. The dependence of a function on the Markov chain state variable $y_k$ will be omitted, in several points in the text.

## II. Description of the Game

At first, the random entrance of the minor players is described. The minor players have time horizon $T$, i.e., each one of them stays in the game for $T$ time steps. Consider a countably infinite set of minor players $\Delta = \{1, 2, \ldots\}$. For any minor player $i \in \Delta$, let $t_i : \Omega \to \mathbb{N}$ be a stopping time describing the time step at which the player $i$ enters the game. At the time step $k$, the number of the minor players that participate in the game may be described by the vector

$$y_k = \left(N_k^0, \ldots, N_k^{T-1}\right) / s_c \tag{1}$$

where $N_k^l = \# I_k^l$ and $I_k^l$ is the set of players with entrance time $k - l$ and and $s_c$, which will be called the "scale variable," is the maximum possible number of active players. Let us finally denote by $I_k$, the set of active players at time step $k$.

The number of new minor players that enter the game at the time step $k + 1$ is a random variable with a distribution depending on $y_k$. Thus, the random entrance is modeled by the Markov chain $y_k$ having a finite state space. Let $1, 2, \ldots, M$ be an enumeration of the state space and $\Pi = [p_{ij}]$ the transition matrix of the Markov chain. We shall use the vector form (1) and the enumeration interchangeably.

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 in a symmetric manner. The dynamic equation of the major player is given by

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

where $x^M$ and $x^i$ are the state vectors of the major player and minor player $i$ respectively. The stochastic disturbances $w^M(k)$ are zero mean, finite variance, i.i.d. random variables, independent of the state vectors. The initial condition for the major player is given by $x^M(0) = w^M(-1)$.

The dynamics of the minor player $i$ is described by

$$x^i(k+1) = Ax^i(k) + \sum_{j \in I_k} Fx^j(k)/s_c + Gx^M(k) + Bu^i(k) + w^i(k) \tag{3}$$

where the stochastic disturbances $w^i(k)$ are zero mean finite variance random variables, independent of the state vectors $x^M(k)$ and $x^i(k)$, $i \in I_k$. The initial values of the state vectors of the minor players are given by $x^i(t_i) = w^i(t_i - 1)$. The dependence of $w^i(k)$ with $w^j(k)$, $i \neq j$ is not disallowed.

In order to define the cost functions of the players, let us introduce the mean field quantities $z^l(k) = \sum_{i \in I_k^l} x^i(k)/s_c$ and the vector of the mean field quantities

$$\tilde{z}(k) = \left[z^0, \ldots, z^{T-1}\right].$$

The cost function of the major player is given by

$$J^M = E\left\{\sum_{k=0}^{\infty} a^k \left[\left[\left(x^M(k)\right)^T \tilde{z}^T(k)\right] Q^M(y_k) \cdot \left[\left(x^M(k)\right)^T \tilde{z}^T(k)\right]^T + (u^M(k))^T R^M u^M(k)\right]\right\} \quad (4)$$

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

For the minor player $i$, the cost function is given by

$$J^i = E\left\{\sum_{k=0}^{N} a(\tilde{x}^i(t_i+T))^T Q_f(y_{t_i+T})\tilde{x}^i(t_i+T) + \sum_{k=t_i}^{t_i+T-1} (\tilde{x}^i(k))^T Q(y_k)\tilde{x}^i(k) + (u^i(k))^T R u^i(k)\right\} \quad (5)$$

where $\tilde{x}^i = [(x^M)^T \tilde{z}^T (x^i)^T]^T$, $Q_f(y)$ and $Q(y)$ positive semidefinite matrices of appropriate dimensions for any $y \in \{1, \ldots, M\}$ and $R$ positive definite matrix of appropriate dimensions.

The problem considered here is the characterization of a Nash equilibrium that satisfies the Dynamic Programming (Perfect equilibrium [35]). We shall focus on Linear Feedback Strategies (i.e., strategies with no memory; see [2, Def. 5.2]). Furthermore, due to the symmetry of the dynamic equations and cost functions, we shall further concentrate to strategies in the following form:

$$u^i = L^{1M}(k-t_i, y_k)x^M + \sum_{l=0}^{T_j-1} L(l, k-t_i, y_k)z^l + \bar{L}(k-t_i, y_k)x^i(k) \quad (6)$$

and

$$u^M = L^{MM}(y_k)x^M + \sum_{l=0}^{T-1} L^M(l, y_k)z^l. \quad (7)$$

The equations (6) and (7) serve only as a general form of the feedback strategies. Equations characterizing the gains $L^{1M}, L, \bar{L}, L^{MM}$, and $L^M$ are determined in the next sections.

For the compactness of the presentation, the following notation will be used:

$$\tilde{L}^M(y) = [L^{MM}(y)L^M(0,y)\ldots L^M(T-1,y)] \quad (8)$$

$$\hat{L}_k(y) = [L^{1M}(k,y)L(0,k,y)\ldots L(T-1,k,y)\bar{L}(k,y)]$$

$$\tilde{L}(y) = [\hat{L}_0(y), \ldots, \hat{L}_{T-1}(y)] \quad (9)$$

*Remark 1:* A set of strategies in the form (6), (7) has two types of symmetries. At first, the feedback gains are the same for all the minor players. Consider a strategy in that form. Then the control values depend on the mean field quantities $z^l$. Thus, the feedback gains corresponding to the players of the same entrance time are the same, which is a second form of symmetry. These symmetry assumptions are justified by the structure of the dynamics and the cost functions. □

*Remark 2:* Although we know that for Linear Quadratic games, closed loop Nash equilibria in nonlinear strategies may also exist, it is only the linear ones that survive if we introduce noise in the state equation or the measurements [36]. This is the reason due to which we restrict our attention to Linear Feedback Nash equilibria. □

*Remark 3:* An interesting extension is to study games involving minor players with different time horizons.[1] This does not make the problem more difficult and all the results in this work can be immediately generalized to that case.

*Remark 4:* The major player may be viewed as a coordinator helping the stabilization of the overall system. An interesting alternative is to see the major player as a common adversary of all the minor players in the spirit of [31].[2] The techniques used in the current work should, however, be adapted in order to study this alternative.

*Remark 5:* An interesting extension would involve the continuous time analog of the current formulation. This can be done either by taking the limit of the discrete problems as the discretization time tends to zero in the spirit of [37] or by stating the corresponding problem in continuous time directly.

*Remark 6:* The study of Stackelberg equilibria with the major player as a leader is a related interesting problem. The same techniques can be used in order to characterize feedback Stackelberg equilibria.

*Remark 7:* The only necessary measurements for a player to implement a strategy in the form of (6) or (7) are the value of the state vector of the major player $x^M$, the mean field quantities $\tilde{z}(k)$, the value of its own state vector and the value of the Markov chain state variable $y_k$. Thus, we shall make the following assumption:

*Assumption 1:* All the players have access to the current values of $x^M$, $\tilde{z}$, and $y_k$. Furthermore, each player can measure its own state vector. □

## III. Optimal Control Problems

The problem of the Nash equilibrium characterization for LQ games with random entrance is converted to the problem of finding a solution to a set of coupled LQ control problems for MJLS. Particularly, the optimal control problems are stated in spaces of smaller dimensions and the random entrance problem is transformed to a random coefficients problem of a linear dynamic equation, depending on the Markov chain given by (1). This reduction is possible, due to the symmetric form of the dynamic equations, the cost functions and the control strategies. We shall assume that the players follow strategies in the general form (6), (7).

[1] In fact, the first version of this work involved several types of minor players having different time horizons. However, for simplicity and clarity of the presentation reasons, after reviewers' recommendation, we restrict ourselves to the case where all the minor players have the same time horizon.

[2] This alternative was proposed by an anonymous reviewer.

### A. Optimal Control Problem for the Major Player

The evolution of the state vector of $x^M$ and the cost function $J^M$ depend only on $x^M$, $\tilde{z}$ and $u^M$. Assuming that the minor players use the strategies in the general form (6), the evolution of the components $z_k^l$ of $\tilde{z}$ depend only on and $x^M$ and $\tilde{z}$, as well. Hence, symmetry implies that the the evolution of the quantities in the cost function (4) can be described by a state vector of smaller dimension: $[(x^M)^T,\ \tilde{z}^T]^T$. The dynamics, after straightforward manipulations, is given by

$$\begin{bmatrix} x^M(k+1) \\ \tilde{z}(k+1) \end{bmatrix} = \tilde{A}^M(y_k) \begin{bmatrix} x^M(k) \\ \tilde{z}(k) \end{bmatrix} + \begin{bmatrix} B^M \\ 0 \end{bmatrix} u^M(k) + W^M(k) \quad (10)$$

where

$$\tilde{A}^M(y_k) = \begin{bmatrix} \tilde{A}^M_{x^M x^M} & \tilde{A}^M_{x^M z^{1,0}} & \dots & \tilde{A}^M_{x^M z^{p,\bar{T}}} \\ \tilde{A}^M_{z^0 x^M} & \tilde{A}^M_{z^0 z^0} & \dots & \tilde{A}^M_{z^0 z^{\bar{T}}} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{A}^M_{z^{\bar{T}} x^M} & \tilde{A}^M_{z^{\bar{T}} z^0} & \dots & \tilde{A}^M_{z^{\bar{T}} z^{\bar{T}}} \end{bmatrix}$$

$\bar{T} = T - 1$. The entries of the first row of the matrix are given by: $\tilde{A}^M_{x^M x^M} = A^M$ and $\tilde{A}^M_{x^M z^l} = F^M$. The second row consists of zeros. For the rest of the entries it holds

$$\tilde{A}^M_{z^{l+1} x^M} = \frac{N_k^l}{s_c}(G + BL^{1M}(l, y_k)), \text{ and}$$
$$\tilde{A}^M_{z^{l+1} z^{l'}} = \frac{N_k^l}{s_c}(F + BL(l', l, y_k)) + \delta_{l,l'}(A + B\bar{L}(l, y_k)).$$

The matrix $\tilde{A}^M$ depends on $y_k$ through the terms $N_k^l/s_c$. Thus, the study of the optimal control problem of the major player, under the random entrance of the minor players is reduced to the study of the following infinite horizon LQ control problem for a MJLS.

*OC Problem 1:* "Minimize the cost function (4) subject to the dynamics (10) and (1)." □

### B. Optimal Control Problem for the Minor Players

For the minor players, a similar reasoning applies. Consider a minor player $i_0$ with entrance time $t_{i_0}$. Assume that the other players use the feedback strategies in the general form (6) and (7). The evolution of the state vector and the cost of the player $i_0$ depend only on the quantities: $x^M$, $x^{i_0}$, $\tilde{z}$, and $y$. Thus, the cost function of player $i_0$ can be described using the state vector $\tilde{x}^{i_0} = [(x^M)^T \tilde{z}^T (x^{i_0})^T]^T$, which evolves according to

$$\tilde{x}^{i_0}(k+1) = \tilde{A}(k - t_{i_0}, y_k)\tilde{x}^{i_0}(k) + \tilde{B}(k - t_{i_0}, y_k)u^{i_0} + W^{i_0}(k) \quad (11)$$

where

$$\tilde{A}(k - t_{i_0}, y_k) = \begin{bmatrix} \tilde{A}_{x^M x^M} & \tilde{A}_{x^M z^0} & \dots & \tilde{A}_{x^M z^{\bar{T}}} & \tilde{A}_{x^M x^{i_0}} \\ \tilde{A}_{z^0 x^M} & \tilde{A}_{z^0 z^0} & \dots & \tilde{A}_{z^0 z^{\bar{T}}} & A_{z^0 x^{i_0}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{A}_{z^{\bar{T}} x^M} & \tilde{A}_{z^{\bar{T}} z^0} & \dots & \tilde{A}_{z^{\bar{T}} z^{\bar{T}}} & A_{z^{\bar{T}} x^{i_0}} \\ \tilde{A}_{x^{i_0} x^M} & \tilde{A}_{x^{i_0} z^0} & \dots & \tilde{A}_{x^{i_0} z^{\bar{T}}} & A_{x^{i_0} x^{i_0}} \end{bmatrix}$$

$\bar{T} = T - 1$. The entries of the matrix are computed by simple but lengthy calculations. For the first row, it holds

$$\tilde{A}_{x^M x^M} = A^M + B^M L^{MM}(y_k),$$
$$\tilde{A}_{x^M z^l} = F^M + B^M L^M(l, y_k).$$

The entries of the second row are zero. For the rest of the rows except the last one, the entries are given by

$$\tilde{A}_{z^{l+1} x^M} = \frac{N_k^l}{s_c}(G + BL^{1M}(l, y_k)) - \frac{\delta_{l,k-t_{i_0}}}{s_c} BL^{1M}(l, y_k)$$
$$\tilde{A}_{z^{l+1} z^{l'}} = \delta_{l,l'}\left(A + \frac{N_k^l}{s_c} B\bar{L}(l, y_k)\right) + \frac{N_k^l}{s_c}(F + BL(l', l, y_k)) - BL(l', l, y_k)\delta_{l,k-t_{i_0}}/s_c$$
$$\tilde{A}_{z^{l+1} x^{i_0}} = -\frac{\delta_{l,k-t_{i_0}}}{s_c} B\bar{L}(l, y_k).$$

The entries of the last row are determined by (3). Thus, $\tilde{A}_{x^{i_0} x^M} = G$, $\tilde{A}_{x^{i_0} z^l} = F$ and $\tilde{A}_{x^{i_0} x^{i_0}} = A$.

The $\tilde{B}$ matrix is given by

$$\tilde{B}(k - t_{i_0}, y_k) = \begin{bmatrix} 0 & \tilde{B}^T_{z^0} & \dots & \tilde{B}^T_{z^{T-1}} & B^T \end{bmatrix}^T$$

where $\tilde{B}_{z^{l+1}} = \delta_{l,k-t_{i_0}} B/s_c$.

The matrix $\tilde{A}$ is time varying and depends on the Markov chain through the terms $N_k^l/s_c$. Hence, the random entrance problem is transformed to a MJLS problem for the minor players, as well. The optimal control problem that a minor player faces is the following finite LQ control problem for a MJLS:

*OC Problem 2:* "Minimize the cost function (5) subject to the dynamics (11) and (1)." □

*Remark 8:* The state vectors of the dynamics of the major and minor players (10) and (11) have a much smaller dimention than the state vector consisting of all the active players. Furthermore, the dimensions of the state vectors do not depend on the number of players. □

*Remark 9:* The Markov jump character of the OC problems has two origins. The first is the random entrance and affects several terms such as $A^M_{z^{l+1} x^M}$, through the factor $N_k^l/s_c$. The second is the dependence of the $Q$ matrices on $y_k$. Hence, the dependence of the $Q$ matrices on $y_k$, does not make the problem more difficult. □

## IV. Optimality Conditions and Nash Equilibrium

The optimality conditions for the OC Problems 1 and 2 are derived and then used to characterize a perfect Nash equilibrium. The general form of the solutions of the finite and infinite horizon discounted LQ control problems for MJLS can be found in [38].

For the infinite horizon OC Problem 1, the optimality conditions are given in terms of a set of coupled Riccati type equations. Particularly, let us consider a set of matrices $K(y)$, $\Lambda(y)$, for $y = 1, \ldots, M$ such that

$$K^M(y) = Q^M + (\tilde{A}^M)^T \Big[ a\Lambda^M - a\Lambda^M \tilde{B}^M \cdot \left( R^M/a + (\tilde{B}^M)^T \Lambda^M \tilde{B}^M \right)^{-1} (\tilde{B}^M)^T \Lambda^M \Big] \tilde{A}^M \quad (12)$$

$$\Lambda^M(y) = E[K^M(y_{k+1})|y_k = y] = \sum_{j=1}^{M} p_{yj} K^M(j). \quad (13)$$

Let us also consider the control law given by

$$u^M(k) = \tilde{L}^{M,\star}(y_k) \left[ \left( x^M(k) \right)^T \tilde{z}^T(k) \right]^T \quad (14)$$

where

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

The control law given by (14) is optimal, provided that it makes the cost function (4) finite. The criterion for the finiteness of the cost function (4) is given in terms of the closed loop matrix

$$\tilde{A}^M_{cl} = \tilde{A}^M - \tilde{B}^M \tilde{L}^{M,\star}. \quad (16)$$

The finiteness criterion is based on the existence of a positive definite solution to a set of coupled Lyapunov equations (see Appendix A, (51)).

*Finiteness Criterion 1:* There exist positive definite matrices $S(i)$, $i = 1, \ldots, M$ that solve the following set of coupled Lyapunov equations:

$$S(i) - \sum_{j=1}^{M} a p_{ij} (\tilde{A}^M_{cl}(i))^T S(j) \tilde{A}^M_{cl}(i) = I \quad (17)$$

for $i = 1, \ldots, M$. □

The optimality conditions for a minor player $i_0$ are given in terms of the solution of the following set of coupled Riccati type difference equations:

$$K_T(y) = Q_f(y) \quad (18)$$

$$\Lambda_{k+1}(y) = E\left[ K_{k+1}(y_{k+1+t_{i_0}}) | y_{k+t_{i_0}} \right] = \sum_{j=0}^{M} p_{yj} K_{k+1}(j) \quad (19)$$

$$K_k(y) = Q + \left( \tilde{A}(k) \right)^T \Big[ \Lambda_{k+1} - \Lambda_{k+1} \tilde{B} \cdot (R + \tilde{B}^T \Lambda_{k+1} \tilde{B})^{-1} \tilde{B}^T \Lambda_{k+1} \Big] \tilde{A} \quad (20)$$

for $k = T-1, \ldots, 0$. The optimal control law is then given by

$$u^{i_0}(k + t_{i_0}) = \hat{L}^\star_k \tilde{x}^{i_0}(k + t_{i_0}) \quad (21)$$

where

$$\hat{L}^\star_k(y) = -(R + \tilde{B}^T \Lambda_{k+1} \tilde{B})^{-1} \Lambda_{k+1} \tilde{A} \quad (22)$$

We are interested in Nash equilibria, i.e., a set of strategies, each of which is optimal given the others. Therefore, we define the notion of a consistent set of strategies.

*Definition 1:* Consider a set of strategies in the general form (6), (7), with gains $\tilde{L}$ and $\tilde{L}^M$ and the set of matrices $\tilde{A}^M(y)$ and $\tilde{A}(k, y)$ for $k = 0, \ldots, T$ and $y = 1, \ldots, M$. Then, the set of strategies is called consistent if:

(i) There exists a set of matrices $K^M(y)$, $\Lambda^M(y)$ for $y = 1, \ldots, M$ that satisfies (12), (13). Moreover, it holds

$$\tilde{L}^M(y) = L^{M,\star}(y) \quad (23)$$

for any $y = 1, \ldots, M$.

(ii) The Finiteness Criterion 1 is satisfyied.

(iii) It holds

$$\tilde{L}(y) = \left[ \hat{L}^\star_0(y) \ldots \hat{L}^\star_{T-1}(y) \right] \quad (24)$$

for any $y = 1, \ldots, M$. □

It is not difficult to show that a consistent set of strategies constitutes a perfect Nash equilibrium.

*Proposition 1:* Consider a consistent set of strategies in the general form (6) and (7). Then it constitutes a Perfect Nash equilibrium.

*Proof:* The strategies of the minor players are optimal. Using equation (51) of the Appendix, with $\sqrt{a} A^M_{cl}$ in the place of the $A$ matrix, we conclude that equation (17) implies the finiteness of the total cost of the Major player and thus the optimality of the control law (14). Thus, the strategy of each player is optimal and the consistent set of strategies constitutes a Nash equilibrium. Furthermore, the strategies of the players satisfy the Dynamic Programming and thus the Nash equilibrium is perfect. □

Sufficient conditions for the existence of a Nash equilibrium are given in Section VI, for a special case. They are expressed as sufficient conditions for the convergence of an algorithm approximating the Nash equilibrium.

*Remark 10:* The optimality conditions given by the equations (12)–(15) and (18)–(22) are Riccati type equations with two kinds of coupling. The first is due to the involvement of the gains in the $\tilde{A}$ matrices and has the same nature as the coupled Riccati equations of the LQ games [1], [39]. The second kind of coupling is through the $\Lambda$ matrices and it has the same nature as the interconnected Riccati equations in the study of LQ control of MJLS [40]. □

## V. Large Number of Players Case

In this section we use a Mean Field approximation in order to study games with a very large number of players. This approach assumes a continuum of players. A set of optimal control problems that correspond to the limit of those in Section III, as the scale variable tends to infinity, is then stated. The Markov chain with a large number of states is approximated by a Markov chain with a continuum of states and thus a notion of convergence of Markov chains is first recalled in the Section V-A. Then the solution of the approximate optimal control problems for the major and minor players is characterized by appropriate Riccati integral equations and consistency conditions analogous of those of Section IV are stated. Finally, it is proved that a set of feedback strategies satisfying those consistency conditions constitutes an $\varepsilon$—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 the minor players have a time horizon 5 and the new minor players in each step belong to the set $1, 2, \dots, \mathrm{N}$ then the state space of the Markov chain describing the entrance has $N^5$ points. Thus, the equations characterizing a Nash equilibrium depend on many parameters and therfore, they are very complicated. On the other hand, in several cases the situation is much simplified using the continuous approximation.

### *A. Convergence of Markov Chains*

In this section the vector representation of $y_k$ will be used. The state space of the Markov chain is contained in the set

$$D = \left\{ (y^0, \dots, y^{T-1}) \in \mathbf{R}^T : \sum_{i=0}^{T-1} y^i \leq 1, y^i \geq 0, \right\}. \quad (25)$$

The continuous approximation will be defined on the set $D$.

A Markov chain could be described using the notion of the stochastic kernel.

*Definition 2:* Let $D' = \{d_1, \dots, d_M\} \subset D$ and $P = [p_{ij}]$ be an $M \times M$ stochastic matrix. The stochastic kernel that corresponds to the Markov chain with state space $D'$ and transition matrix $P$ is defined as

$$\bar{K}(y, B) = Pr(y_{k+1} \in B | y_k = y) = \sum_{j: d_j \in B} p_{ij} \quad (26)$$

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

Let us then recall a notion of convergence of stochastic kernels from [41] and a notion of continuity from [42].

*Definition 3:*

(i) We shall say that a sequence of stochastic kernels $\bar{K}_\nu$ converges weakly to a stochastic kernel $\bar{K}$ if for any sequence $y_\nu$ of elements of $D$ converging to an element $y$ of $D$ and any (bounded) continuous function $g$, it holds

$$\int_D g(y')\bar{K}_\nu(y_\nu, y') \to \int_D g(y')\bar{K}(y, y') \quad (27)$$

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

Let us turn back to games described by the relationships (2)–(5) and a large number of minor players. To do so, we consider a sequence $g^\nu$ of games with increasing number of minor players; i.e., for the scale variable we assume $s_c \to \infty$ as $\nu \to \infty$. The state of the Markov chain describing the entrance is denoted by $y_k^\nu$, the number of states of the Markov chain by $M^\nu$ 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 sequence of stochastic kernels $\bar{K}^\nu$ converges weakly to a Feller continuous stochastic kernel $\bar{K}$. The stochastic kernel $\bar{K}$, hence, approximates the final part of the sequence of Markov chains. We finally assume that the matrix functions $Q^M(\cdot), Q_f(\cdot), Q(\cdot)$ are continuous.

The following example shows that the continuum approximation often simplifies a lot the description of the random entrance.

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

$$Pr(N_{k+1}^{\nu,0} = i) = \binom{\nu}{i} p^i (1-p)^{\nu-i}.$$

The random variable $N_k^{\nu,0}/s_c$ converges weakly to the deterministic constant $p/2$. Thus, the Markov chain with large $\nu$ may be approximated by a Markov chain with continuous state space and a stochastic kernel given by $\bar{K}((y_1, y_2), \cdot) = \delta(p/2, y_1)$, where $\delta$ denotes the Dirac measure. Thus, for a large number of players, the approximate description of the Markov chain is much simpler than the original. □

### *B. Approximate Optimal Control Problems*

The approximate optimal control problems are then stated. These problems correspond to the limits of the OC Problems 1 and 2 of Section III as the scale variable tends to infinity.

The reduced order dynamics for the major player, given by (10), remains unchanged under the limiting procedure. Thus, the limit optimal control problem for the major player is stated as follows:

*OC Problem 3:* "Minimize the cost function (4) subject to the dynamics (10) and $y_{k+1} \sim \bar{K}(y_k, \cdot)$." □

The solution of the optimal control Problem 3 depends on the solution of a Riccati integral equation given by

$$K^M(y) = Q^M + (\tilde{A}^M)^T \Big[ a\Lambda^M - a\Lambda^M \tilde{B}^M \cdot \Big( R^M/a + (\tilde{B}^M)^T \Lambda^M \tilde{B}^M \Big)^{-1} (\tilde{B}^M)^T \Lambda^M \Big] \tilde{A}^M \quad (28)$$

$$\Lambda^M(y) = E[K^M(y_{k+1}) | y_k = y] = \int_D K^M(y')\bar{K}(y, dy'). \quad (29)$$

Consider the matrix functions $K(\cdot), \Lambda(\cdot)$ satisfying the Riccati integral equation (28), (29). Then the control law given by

$$u^M(k) = \tilde{L}^{M,\star}(y_k) \left[ \left(x^M(k)\right)^T \tilde{z}^T(k) \right]^T \quad (30)$$

where

$$\tilde{L}^{M,\star}(y) = -\Big( (\tilde{B}^M)^T \Lambda \tilde{B}^M + R^M/a \Big)^{-1} \Lambda \tilde{A}^M \quad (31)$$

solves the optimal control Problem 3, under the following finiteness criterion:

*Finiteness Criterion 2:* Consider the closed loop matrix given by

$$\tilde{A}_{cl}^M = \tilde{A}^M - \tilde{B}^M \tilde{L}^{M,\star}.$$

There exists a strictly positive definite matrix function $S(\cdot)$ satisfying

$$\int_{y'\in D} a\left(\tilde{A}_{cl}^{M}(y)\right)^{T} S(y')\tilde{A}_{cl}^{M}(y)\bar{K}(y,dy') - S(y) = -I \quad (32)$$

for any $y \in D$. □

The reduced order dynamics for a minor player is simplified under the limiting procedure. Specifically, consider a minor player $i_0$ with entrance time $t_{i_0}$. The limit dynamics is given by

$$\tilde{x}^{i_0}(k+1) = \tilde{A}(k-t_{i_0}, y_k)\tilde{x}^{i_0}(k) + \tilde{B}(k-t_{i_0}, y_k)u^{i_0} + W^{i_0}(k). \quad (33)$$

The entries of the matrix are computed by simple but lengthy calculations. The entries of the first row are given by

$$\tilde{A}_{x^M x^M} = A^M + B^M L^{MM}(y_k),$$
$$\tilde{A}_{x^M z^l} = F + B^M L^M(l, y_k).$$

The entries of the second row are zero. For the rest of the rows except the last one, the entries are given by

$$\tilde{A}_{z^{l+1}x^M} = y_k^l(G + BL^{1M}(l, y_k))$$
$$\tilde{A}_{z^{l+1}z^{l'}} = \delta_{l,l'}\left(A + y_k^l B\bar{L}(l, y_k)\right) + (F + BL(l', l, y_k))y_k^l$$
$$\tilde{A}_{z^{l+1}x^{i_0}} = 0.$$

The entries of the last row remain the same, i.e., $\tilde{A}_{x^{i_0}x^M} = G$, $\tilde{A}_{x^{i_0}z^l} = F$ and $\tilde{A}_{x^{i_0}x^{i_0}} = A$.

The limit optimal control problem for a minor player $i_0$ is, thus, the following:

*OC Problem 4:* "Minimize the cost function (5) subject to the dynamics (33) and $y_{k+1} \sim \bar{K}(y_k, \cdot)$." □

The solution of the limit optimal control Problem 4 depends on the following Riccati type difference integral equaitons:

$$K_T(y) = Q_f(y) \quad (34)$$

$$\Lambda_{k+1}(y) = E\left[K_{k+1}(y_{k+1+t_{i_0}})|y_{k+t_{i_0}}\right] = \int_D K_{k+1}(y')\bar{K}(y, dy') \quad (35)$$

$$K_k(y) = Q + \left(\tilde{A}(k)\right)^T \left[\Lambda_{k+1} - \Lambda_{k+1}\tilde{B} \cdot (R + \tilde{B}^T\Lambda_{k+1}\tilde{B})^{-1}\tilde{B}^T\Lambda_{k+1}\right]\tilde{A}. \quad (36)$$

The optimal control law is then given by

$$u^{i_0}(k+t_{i_0}) = \hat{L}_k^{\star}\tilde{x}^{i_0}(k+t_{i_0}) \quad (37)$$

where

$$\hat{L}_k^{\star}(y) = -(R + \tilde{B}^T\Lambda_{k+1}\tilde{B})^{-1}\Lambda_{k+1}\tilde{A}. \quad (38)$$

### C. Consistency Conditions and $\varepsilon$—Nash Equilibrium

The consistency conditions for the solutions of the OC Problems 3 and 4 are stated and then used to characterize approximate Nash equilibrium.

*Definition 4:* Consider a set of strategies that belong in the general form (6), (7) with gains $\tilde{L}$ and $\tilde{L}^M$. Assume that they depend continuously on $y \in D$. Compute the matrix functions $\tilde{A}^M(y)$ and $\tilde{A}(y)$. The set of strategies will be called consistent if:

(i) There exist continuous matrix functions $K^M(y), \Lambda^M(y)$ satisfying (28), (29). Moreover it holds

$$\tilde{L}^M = \tilde{L}^{M,\star} \quad (39)$$

where $\tilde{L}^{M,\star}(y)$ is given by (31).

(ii) The closed loop matrix $\tilde{A}_{cl}^M(y)$ satisfies the Finiteness Criterion 2.

(iii) The matrix functions $\hat{L}_k^{\star}(y)$ computed by (34)–(38) satisfy

$$\tilde{L}(y) = [\hat{L}_0^{\star}(y), \ldots, \hat{L}_{T-1}^{\star}(y)] \quad (40)$$

for any $y \in D$. □

Let us consider a game with a large number of players $g^\nu$, where the participants use a set of approximately consistent strategies, characterized by gains $\tilde{L}$ and $\tilde{L}^M$. Under certain conditions, this set of strategies is shown to constitute an $\varepsilon$—Nash equilibrium, i.e., the cost of any player is at most $\varepsilon$—far from the optimal cost. This property is illustrated by the following Theorem 1 and its Corollary 1.

Before stating the Theorem 1, let us introduce some notation:

*Notation:* For the game $g^\nu$, let us denote by $J^{M,\nu}(\pi_{\tilde{L},\tilde{L}^M})$ and $J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M})$ be the values of the cost functions, (4) and (5), when all the players use the policies given by (6) and (7) with gains $\tilde{L}, \tilde{L}^M$.

Let also $(J^{M,\nu}(\pi_{\tilde{L},-M}))^\star$ be the minimum value of the cost function of the major player, assuming that the other players use the policies given by (6) with gains $\tilde{L}$. Finally, denote by $(J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M,-i}))^\star$ the minimum value of the cost function of the player $i$, assuming that the other players use the strategies given by (6) and (7) with gains $\tilde{L}, \tilde{L}^M$. □

*Theorem 1:* Consider an approximately consistent set of strategies given by (6) and (7) with gains $\tilde{L}, \tilde{L}^M$. Then for any positive constant $\varepsilon$, there exists a positive integer $\nu_0$ such that

$$J^{M,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \le (J^{M,\nu}(\pi_{\tilde{L},-M}))^\star + \varepsilon \quad (41)$$

$$J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \le (J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M}, -i))^\star + \varepsilon\left(1 + E\left[\left(\tilde{x}^i(t_i)\right)^2\right]\right), \quad (42)$$

for all the minor players $i \in \Delta$ and any $\nu \ge \nu_0$.

*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 the Markov chain and some properties of the weak convergence.

A basic step in the proof of the first inequality is given in Appendix B, where it is shown that some stability properties of the MJLS are preserved under weak convergence. It is then shown that the final part of the series involved in the cost function 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 the Appendix. More general results are first shown in section C and particularly in Propositions 4 and 5. Theorem 1 is then proved as a consequence of the Propositions of section C. □

*Corollary 1:* Consider a set of strategies in the form (6), (7). In addition to the assumptions of Theorem 1, assume that the closed loop system is mean square exponentially stable, i.e., the following Lyapunov equation:

$$\int_{y' \in D} (\tilde{A}_{cl}^M(y))^T S(y') \tilde{A}_{cl}^M(y) \bar{K}(y, dy') - S(y) = -I$$

admits a strictly positive definite solution $S(\cdot)$. Then for any positive constant $\varepsilon$, there exists a positive integer $\nu_0$ such that (6) and (7) constitute an $\varepsilon$-Nash equilibrium for any $\nu \geq \nu_0$, i.e., it holds

$$J^{M,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \leq (J^{M,\nu}(\pi_{\tilde{L},-M}))^\star + \varepsilon \tag{43}$$

$$J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \leq (J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M,-i}))^\star + \varepsilon. \tag{44}$$

*Proof:* The inequality (44) is a consequence of the inequality (42) and the mean square stability. □

*Remark 11:* The approximate consistency conditions (Def. 4) involve nonlinear matrix integral equations and in general are not simpler than the consistency conditions of Section IV. However, in several cases the situation is extremely simplified as illustrated in the Example 3 of the next section. □

## VI. Computing the Nash Equilibria

An algorithm for solving the consistency conditions derived in Section IV, is described in subsection A. Conditions under which the algorithm converges are stated and thus sufficient conditions for the existence of a Nash equilibrium are then derived. In the third subsection, some numerical examples are given.

### A. Algorithm

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 then used to compute the system matrices and solve the optimal control problems and so forth. This algorithm is presented in Algorithm 1 table.

**Algorithm 1**

1: Guess $\tilde{L}$, $\tilde{L}^M$.
2: Compute the matrices $\tilde{A}(\cdot,\cdot)$ using (11).
3: Compute the new values for the gains $\hat{L}_1^\star, \ldots, \hat{L}_{T-1}^\star$, using equations (18)–(22).
4: Set $\tilde{L} = [\hat{L}_1^\star, \ldots, \hat{L}_{T-1}^\star]$
5: Compute the matrix $\tilde{A}^M$ using (10).
6: Set $\tilde{L}_{old}^M = \tilde{L}^M$.
7: Compute matrices $K^M(\cdot), \Lambda^M(\cdot)$ to satisfy (12), (13).
8: Use (15) to update the values of $\tilde{L}^M$, i.e., set $\tilde{L}^M = \tilde{L}^{M,\star}$.
9: If the difference $\max_y \|\tilde{L}^M - \tilde{L}_{old}^M\|$ is small enough then halt. Else go to Step 2.

An analogous algorithm can be used for the Mean field case, as well.

*Remark 12:* Step 7 of the algorithm may be implemented in several ways. Probably the simpler one is to use the value iteration algorithm. A variant of the algorithm would involve the use of a single step of the value iteration method, instead of the steps 7 and 8 of the Algorithm 1. □

### B. Convergence of the Algorithm

The convergence of the Algorithm 1 depends on the existence of a Nash solution, as well as on some stability properties. Such problems are hard to solve and remain open even in simpler settings (ex. [43]).

In this subsection, we study the convergence of the Mean Field variant of Algorithm 1. Particularly, sufficient conditions for convergence of the algorithm and hence, the existence of a Mean Field Nash solution are stated, based on contraction mapping ideas. A special class of games is analyzed. Specifically, it is assumed that there are only minor players coupled only though costs having state vectors of dimension one. The time horizon of the players is three.

In what follows, for vectors $\|\cdot\|$ denotes the 1-norm, for a matrix $A$, $\|A\|$ denotes the induced 1-norm and for a matrix function $A(\cdot)$, $\|A\|$ denotes the essential supremum of the induced 1-norm. Assuming that $A = \alpha$, $B = 1$ and $r = 1$, the matrices $\tilde{A}$ and $\tilde{B}$ take the form

$$\tilde{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \alpha + (L(0,0) + \bar{L}(0)) y_k^0 & L(1,0) y_k^0 & L(2,0) y_k^0 & 0 \\ L(0,1) y_k^1 & a + (L(1,1) + \bar{L}(1)) y_k^1 & L(2,1) y_k^1 & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix}$$

$\tilde{B} = [0\ 0\ 0\ 1]^T$. Let us denote by

$$\begin{aligned} \tilde{L}^{\tilde{A}} &= [\tilde{L}^{\tilde{A}}(0), \tilde{L}^{\tilde{A}}(1)] \\ &= [[L(0,0) \quad L(1,0) \quad L(2,0) \quad \bar{L}(0)], \\ &\quad [L(0,1) \quad L(1,1) \quad L(2,1) \quad \bar{L}(1)]] \end{aligned}$$

i.e., all the entries of $\tilde{L}$ that affect $\tilde{A}$.

We consider the following mappings:

$$\tilde{L}^{\tilde{A}} \xrightarrow{T_1} \left(\tilde{A}; \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}\right) \xrightarrow{T_2} \left(\tilde{A}; \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix}\right) \xrightarrow{T_3} \left(\tilde{L}^{\tilde{A}}\right)^\star. \tag{45}$$

This mappings compute the best response of a player if the other players use strategies described by $L^{\tilde{A}}$. Sufficient conditions for the contractivity of the mapping $T = T_1 \circ T_2 \circ T_3$ are then derived.

The mappings $T_1$ and $T_2$ are computed according to the following equations:

$$\begin{aligned} K_3(y) &= Q_f(y) \\ K_2(y) &= Q + \tilde{A}^T[\Lambda_3 - \Lambda_3 \tilde{B}\tilde{B}^T \Lambda_3/(1 + \Lambda_3^{4,4})]\tilde{A} \\ K_1(y) &= Q + \tilde{A}^T[\Lambda_2 - \Lambda_2 \tilde{B}\tilde{B}^T \Lambda_2/(1 + \Lambda_2^{4,4})]\tilde{A} \\ \Lambda_i(y) &= \int_{y' \in D} K_i(y') \bar{K}(y, dy'). \end{aligned}$$

The mapping $T_3$ is given by

$$(\tilde{L}^{\tilde{A}}(i))^\star = -\frac{1}{1 + \Lambda_{i+1}^{4,4}} \Lambda_{i+1} \tilde{A}, \quad i = 0, 1. \tag{46}$$

*Lemma 1:* It holds

$$\left\| \int_{y' \in D} (K(y') - K'(y'))\bar{K}(y, dy') \right\| \leq \|K(y) - K'(y)\|$$

*Proof:* Immediate □

The mapping $T_2$ is, thus, non-expansive (weakly contractive). Sufficient conditions for the contractivity of $T$ are found using the following technical result.

*Lemma 2:* If $\|A\|, \|A'\| < d_1$, $\|K\|, \|K'\| < d_2$ and $\|A - A'\| < c_1$, $\|K - K'\| < c_2$, then it holds

$$\|f(A, K) - f(A', K')\| < 2(d_1 d_2 + d_1 d_2^2)c_1 + d_1^2(1 + d_2^2 + 2d_2 + c_2)c_2$$

where: $f(A, K) = Q + A^T[K - K\tilde{B}\tilde{B}^T K/(1 + K^{4,4})]A$.

*Proof:* The proof is long but straightforward. It uses repeatedly the matrix identity $XY - X'Y' = (X - X')Y + X'(Y - Y')$ and the sub-multiplicative property of the matrix norm. □

For a constant $0 < \rho < 1$, a region, $\bar{R}_\rho$, containing 0 that the mapping $T$ is $\rho$-contractive, will be determined. Assuming that the algorithm starts with zero gains, $\tilde{L}^{\tilde{A}} = 0$, after the application of $T$ we have

$$\begin{aligned} &\|(\tilde{L}^{\tilde{A}})^\star\| \leq \beta \\ &\quad = 2q\alpha + \alpha^3(q + (1 + \alpha^2)(q_f + q_f^2) + (q + \alpha^2(q_f + q_f^2))^2) \end{aligned}$$

where $q = \|Q\|$ and $q_f = \|Q_f\|$. This inequality can be derived using the sub-multiplicative property of the matrix norm. If after a number of iterations of $T$, $\tilde{L}^{\tilde{A}}$ remains in $\bar{R}_\rho$, then it holds

$$\|\tilde{L}^{\tilde{A}}\| < \beta/(1 - \rho),$$
$$\|\tilde{A}\| < d_1 = a + \beta/(1 - \rho) \text{ and}$$
$$\left\| \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \right\| < d_2$$

where

$$d_2 = 2q + d_1^2(q + (1 + d_1^2)(q_f + q_f^2) + (q + d_1^2(q_f + q_f^2))^2).$$

The last inequality is also shown using the sub-multiplicative property of the matrix norm.

*Proposition 2:* Assume that the parameters $\alpha, q, q_f$ are such that

$$4(d_1 d_2 + d_1 d_2^2) + 2d_1^2[1 + d_2^2 + 2d_2 + 2(d_1 d_2 + d_1 d_2^2)\beta](d_1 d_2 + d_1 d_2^2) < \rho_1. \quad (47)$$

Furthermore, assume that $\rho_2 = d_1 + d_2 + 2d_1 d_2$ is such that

$$\rho = (1 + \rho_1)\rho_2 < 1. \quad (48)$$

Then $T$ is $\rho$-contractive, the Algorithm 1 converges and there exists a MF Nash solution for the game.

*Proof:* We first determine a Lipschitz constant for $T_1$. Assuming that $\|\tilde{L}^{\tilde{A}} - \tilde{L}'^{\tilde{A}}\| < c$, Lemma 2 implies

$$\begin{aligned} &\|\Lambda_2 - \Lambda_2'\| \leq \|K_2 - K_2'\| < 2(d_1 d_2 + d_1 d_2^2)c \\ &\|K_1 - K_1'\| < 2(d_1 d_2 + d_1 d_2^2)c \\ &\quad + 2d_1^2(1 + d_2^2 + 2d_2 + 2(d_1 d_2 + d_1^2 d_2)c)(d_1 d_2 + d_1 d_2^2)c. \end{aligned}$$

Hence, inequality (47) implies that

$$\|K_1 - K_1'\| + \|K_2 - K_2'\| < \rho_1 c.$$

Thus, $T_1$ has Lipschitz constant less than $1 + \rho_1$. Using (46) and the sub-multiplicative property, it is straightforward to show that $T_3$ is $\rho_2$—contractive. These, in combination with the non-expansive property of $T_2$, complete the proof. □

*Remark 13:* The generalization to the many steps case does not add any further difficulties. The existence of a dynamic coupling except the cost coupling makes the matrix $A$, time varying. The generalization to the multi-dimensional case is also immediate. However, only this special case is analyzed in order to keep the results as simple as possible.

### C. Numerical Examples

The following example illustrates the convergence of the Algorithm 1. Furthermore, it studies the dependence of the Nash feedback gains on the coupling of the dynamic equations and costs of the major and minor players.

*Example 2:* In this example major and minor players have scalar state equations. The minor players have time horizon 2 and the maximum possible number of minor players participating in the game at some time step is 4. The number of new minor players that enter the game at each instant of time is either 1 or 2. Thus, the entrance dynamics is described by a Markov chain with state space: $\left(\frac{1}{4}, \frac{1}{4}\right)$, $\left(\frac{1}{4}, \frac{2}{4}\right)$, $\left(\frac{2}{4}, \frac{1}{4}\right)$, and $\left(\frac{2}{4}, \frac{2}{4}\right)$. For the states of the Markov chain, we shall use the enumeration 1,2,3,4 respectively. The entrance dynamics is described by the following transition matrix:

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

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 for a minor player by

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

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

The cost function matrices $Q$ for the major player are given by $Q(1) = Q(2) = Q(3) = I_3$ and $Q(4) = (1 + c_2)I_3$. The cost matrices $Q^1$ for the minor players are all units, i.e., $Q(y) = Q_f(y) = I_4$ and $R = R^M = 1$.

TABLE I
GAIN MATRICES FOR THE MAJOR PLAYER, $c_1 = c_2 = 1$

| $y_k$ | $L^{MM}(y_k)$ | $L^{M}(0, y_k)$ | $L^{M}(0, y_k)$ |
|---|---|---|---|
| 1 | -0.6411 | -0.6618 | -0.6326 |
| 2 | -0.6560 | -0.6787 | -0.6467 |
| 3 | -0.7338 | -0.7211 | -0.7240 |
| 4 | -0.6825 | -0.6682 | -0.6715 |

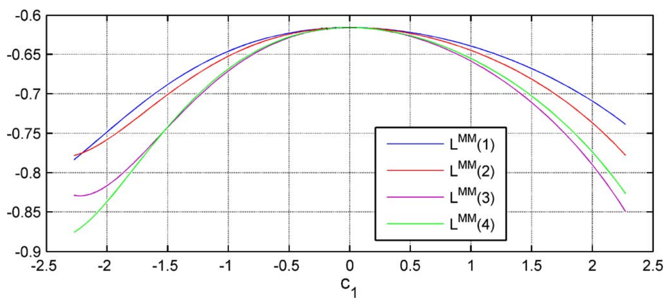


Fig. 1. The dependence of the feedback gains $L^{MM}(y)$ on $c_1$.

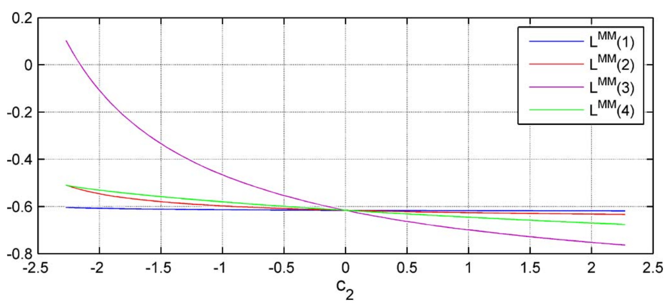


Fig. 2. The dependence of the feedback gains $L^{MM}(y)$ on $c_2$.

For example, if $c_1 = c_2 = 1$, after 30 steps of the Algorithm 1, the gain matrices change less than $10^{-15}$. The feedback gains for the major player are gain in Table I.

In what follows, we study the dependence of the feedback gains on the coupling parameters $c_1$ and $c_2$. Let us first consider the case where $c_2 = 0$ and $c_1 \neq 0$. Then, each player will interact dynamically with an unknown number of minor players. The distribution of the number of the minor players in the next step depends on $y_k$. Thus, the dependence of the feedback gains on $y$ is larger, for larger values of $|c_1|$. The dependence of $L^{MM}(y)$ for $y = 1, \ldots, 4$ on $c_1$ is illustrated in Fig. 1.

We next assume no dynamic coupling, i.e., $c_1 = 0$ and $c_2 \neq 0$. The dependence of $L^{MM}(y)$ for $y = 1, \ldots, 4$ on $c_2$ is illustrated in Fig. 2. Again, the dependence on $y_k$ is larger, for larger values of $|c_2|$.

The Algorithm 1 does not always converge. For example, for $c_1 = c_2 = 10$, the Algorithm 1 does not converge.

Finally we fix a dynamics and cost function for the participants of the game and present the sample paths of a run. The dynamics of the players are as before, with $c_1 = 1$ and the cost functions slightly different. Particularly, $Q^M(y) = I_3$, $R^M = 1$, $R = 0.75$ and

$$Q(y) = Q_f(y) = \begin{bmatrix} 10 & 0 & 0 & -10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -10 & 0 & 0 & 10 \end{bmatrix}$$

The sample paths are shown in Fig. 3. □

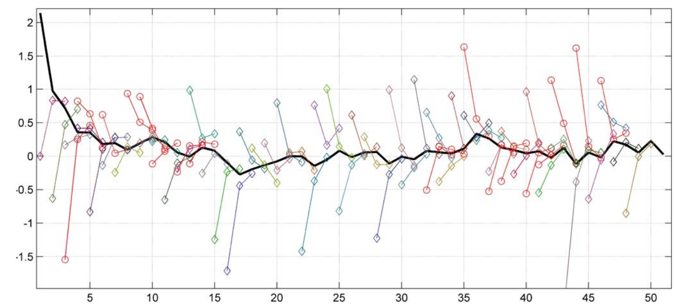

Fig. 3. The sample paths of the major and minor players. The black line corresponds to the major player. The red lines correspond to cases where there is a second player entering at some instant of time.

The following example studies a very simple game with a large number of minor players. It illustrates that, in certain cases, the mean field approximation simplifies very much the analysis of the game. For simplicity reasons, we assume that there is no major player.

*Example 3:* There is only one type of minor players with time horizon two. At each time step, each one of $\nu$ minor players tosses a fair coin and with probability 1/2 enters the game.

The dynamic equation of the minor players is given by

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

The cost function matrices are given by $Q_f = 4y^1 I_3$, $Q = 12yy^T I_3$ and $R = 1$ (where the notation $Q_f, Q, R, y^1, y^0$ is used instead of $Q_f^1, Q^1, R^1, y^{1,1}, y^{1,0}$).

The scale variable has a value $s_c = 2\nu$ and 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 $\delta$ is the Dirac measure.

Due to the absence of a major player the approximate consistency conditions involve only (33)–(36). The unknown quantities can be expressed in terms of the functions $L_1(y) = L(0, 0, y)$, $L_2(y) = L(1, 0, y)$, $L_3(y) = L(0, 1, y)$, $L_4(y) = L(1, 11, y)$, $L_5(y) = \bar{L}(0, y)$, and $L_6(y) = \bar{L}(1, y)$.

Due to the special form of the stochastic kernel, the integral equation (35) becomes

$$\Lambda_{k+1}((y_1, y_2)) = K_{k+1}(\bar{y}, y_1)$$

where $\bar{y} = 1/4$. Hence, the form of the stochastic kernel implies a decoupling in the consistency conditions. Particularly, for $\tilde{y} = (\bar{y}, \bar{y})$, the consistency conditions do not depend on the other values of y. Writing the consistency conditions for some $y' = (\bar{y}, y_1)$, the equations depend only on $L_1(y'), \ldots, L_6(y')$ and $L_1(\tilde{y}), \ldots, L_6(\tilde{y})$. Furthermore, for some $y = (y_1, y_2)$, the consistency conditions depend only on $L_1(y), \ldots, L_6(y)$ and $L_1(y'), \ldots, L_6(y')$.

This structure of the consistency conditions suggests the following procedure: Compute the values of $L_1, \ldots L_6$ on $\tilde{y}$, solving six equations with six unknowns. Then, for each $y' = (\bar{y}, y_1)$, compute the values of $L_1, \ldots L_6$ on $y'$. Finally, for any $y = (y_1, y_2)$, compute the values of $L_1, \ldots L_6$ on $y$, again, by solving six equations with six unknowns involving $L_1, \ldots L_6$ on $y$ and $y'$.

TABLE II
GAIN MATRICES COMPUTATION

| $y$ | $L_1(y)$ | $L_2(y)$ | $L_3(y)$ | $L_4(y)$ | $L_5(y)$ | $L_6(y)$ |
|---|---|---|---|---|---|---|
| $\tilde{y}$ | −0.813 | −0.68 | −0.5 | −0.5 | −0.667 | −0.5 |
| $y'$ | −0.813 | −0.68 | −0.444 | −0.444 | −0.667 | −0.444 |
| $y$ | −0.754 | −0.66 | −0.444 | −0.444 | −0.626 | −0.444 |

As an example, we compute the values of the feedback gains on $y = (0.2, 0.6)$. The computations are shown in Table II.

It remains to show the stability of the limit system. For $k \geq 3$, it holds $y_k = \tilde{y}$ a.s. Furthermore, it holds

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

Thus, Corollary 1 applies, i.e., for any $\varepsilon > 0$ the strategies computed constitute an $\varepsilon$—Nash equilibrium for large $\nu$. □

*Remark 14:* Example 3 shows that when the random entrance is independent, the approximate consistency conditions are decoupled. The solutions to the approximate optimal control problems could, thus, be obtained using this special form. □

## VII. CONCLUSION

Games with a major player and Randomly Entering minor players were considered. The problem of the characterization of Symmetric Linear Feedback Strategies that constitute a Nash equilibrium, was converted to a set of coupled finite and infinite horizon LQ control problems for MJLS. Appropriate coupled Riccati type equations were derived to characterize Nash equilibrium. The case where there exists a very large number of minor players was addressed using a Mean Field approach. Particularly, the evolution of the number of players is approximately described using a Markov chain having a continuum of states. Some limit optimal control problems were then stated. A set of Symmetric Linear Feedback Strategies that solves the limit optimal control problems was proved to constitute an $\varepsilon$—Nash equilibrium, when the scale is sufficiently large. A sufficient condition for the existence of a Mean Field Nash equilibrium was derived using contraction mapping ideas. Numerical examples were also presented. It occurs that, in several cases, the Mean Field approximation simplifies considerably the analysis.

## APPENDIX

For easy reference the notation used in the paper is summarized in Table III.

The rest of the Appendix contains the proof of Theorem 1 and some propositions needed to prove that result. In Section A we recall some results from [38]about the stability of MJLS. Section B 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 B is Proposition 3, which shows that if the limit system is stable then a tail of the sequence of systems consists of stable systems. Section C proves that policies which are optimal for the limit system, are $\varepsilon$—optimal for the tail of the sequence. The basic results of Section C are Proposition 4 and Proposition 5, where the result is proved for the finite and infinite horizon problems respectively. In Section D the proof of Theorem 1 is completed.

TABLE III
NOTATION

| | |
|---|---|
| $t_i$ | Entrance time of player $i$ |
| $\Delta$ | The set of minor players |
| $y_k, [\pi_{ij}]$ | M.C. describing random entrance (1) and tr. matrix |
| $s_c$ | The scale variable |
| $x^M, x^i$ | State vectors for the major player and player $i$ (2), (3) |
| $\tilde{z}, \tilde{x}^i$ | Augmented state vectors before (4), after (5) |
| $A^M, F^M, B^M$ | Parameters of the major player dynamics (2) |
| $A, F, B, G$ | Parameters of the minor players dynamics (3) |
| $L^{1M}, L, \bar{L}$ | Feedback gains for the minor players (6) |
| $L^{MM}, L^M$ | Feedback gains for the major player (7) |
| $\hat{L}_k, \tilde{L}, \tilde{L}^M$ | Feedback gains: compact form (8), (9) |
| $\tilde{A}^M$ | $A$-matrix of the major player's reduced dynamics (10) |
| $\tilde{A}$ | $A$-matrix of the minor player's reduced dynamics (11) |
| $K^M, \Lambda^M$ | Riccati equation matrices (12), (13), (28), (29) |
| $\tilde{L}^{M,\star}$ | Opt. gain for major pl. if others use $\tilde{L}$ (15) or (30) |
| $\hat{L}_k^\star$ | Opt. gain for minor if others use $\tilde{L}, \tilde{L}^M$ (22) or (38) |
| $D$ | State space of the continuous Markov chain (25) |
| $\bar{K}(\cdot,\cdot)$ | The stochastic kernel (26) |

Let us note that the results of Sections B and C, dealing with the approximation of Markov chains in MJLS, are also of independent interest.

### A. Stability of MJLS With General State Space

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

$$\begin{aligned} x_{k+1} &= A(y_k)x_k \\ y_{k+1} &\sim \bar{K}(y_k,\cdot). \end{aligned} \tag{49}$$

The exponential mean square stability of a system in this form is equivalent to the fact that the spectral radius [44] of an operator $T = T_{A,\bar{K}}$ is less than one. The operator is defined using the quantity $P_k : \mathcal{B}(D) \to \mathbb{R}^{n\times n}$, where $\mathcal{B}(D)$ is the set of the Borel subsets of $D$ and $P_k$ has the form

$$P_k(C) = E[x_k x_k^T \chi_{y_k\in C}]. \tag{50}$$

The operator $T$ is defined such that $P_{k+1}(\cdot) = T(P_k(\cdot))$.

The exponential mean square stability of a system in the form (49) has the property to be uniform on the initial conditions. Specifically, the exponential mean square stability of (49) is equivalent to the existence of a constant $a \in (0,1)$ and of a positive integer $k_0$ such that $E[x_{k_0}^T x_{k_0}] < a x_0^T x_0$, for any nonrandom initial conditions $x_0, y_0$. Furthermore, it is equivalent to the existence of positive constants $M$ and $a < 1$ such that $E[x_k^T x_k] < M a^k E[x_0^T x_0]$ for any initial conditions.

Another test for the Mean Square stability is given via the Lyapunov equation. Particularly, consider a strictly positive definite, bounded matrix function $Q(y)$. The mean square stability of the system given by (49) is equivalent to the existence of a strictly positive definite matrix function $S(y)$ satisfying the Lyapunov equation

$$\int_{y'\in D} A^T(y)S(y')A(y)\bar{K}(y,dy') - S(y) = -Q(y). \tag{51}$$

### B. Weak Convergence and Mean Square Stability

Consider a sequence of systems

$$x_{k+1}^\nu = A(y_k)x_k^\nu, \quad y_{k+1}^\nu \sim \bar{K}^\nu(y_k^\nu,\cdot) \tag{52}$$

and a limit system

$$x_{k+1} = A(y_k)x_k, \quad y_{k+1} \sim \bar{K}(y_k, \cdot). \tag{53}$$

Assume that $\bar{K}^\nu \to \bar{K}$ weakly, $\bar{K}$ is Feller continuous and that $A(\cdot)$ is a continuous matrix function. Assume also that the limit system, given by (53), is exponentially mean square stable. The basic topic of this section is to show that the system given by (52) is exponentially mean square stable, for large $\nu$.

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

$$E\left[x_k^T x_k\right] < aE\left[x_0^T x_0\right] \tag{54}$$

for any $x_0, y_0$ initial conditions. Choosing $x_0$ to be any nonrandom initial condition, the last inequality becomes

$$aI - E\left[A^T(y_0)\dots A^T(y_{k-1})A(y_{k-1})\dots A(y_0)\right] > 0. \tag{55}$$

The positive definiteness of this matrix is equivalent, due to the Sylvester criterion, to a set of inequalities in the form

$$f_j\left(E\begin{bmatrix}\bar{f}_1(y_0,\dots,y_{k-1})\\ \vdots \\ f_{n^2}(y_0,\dots,y_{k-1})\end{bmatrix}\right) > 0 \tag{56}$$

for $j = 1,\dots,n$, where $\bar{f}_i$, $i = 1,\dots,n^2$ correspond to the elements of the matrix in (55) and are continuous and $f_j$ are the multinomials derived using the Sylvester criterion.

The inverse procedure shows that the conditions in (56) imply the mean square exponential stability of the limit system (53).

Let us then state the basic result of this section.

*Proposition 3:* Under the assumptions stated above, there exists a positive integer $\nu_0$ such that $r(T_{A,\bar{K}^\nu}) < 1$, for any $\nu \geq \nu_0$.

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 ([45] section 17) which is defined by

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

where $P_1, P_2$ are probability measures on $D$ and $\|f\|_{BL} = \sup_{y\in D}\{f(y)\} + \inf\{L : f \text{ is } \mathrm{L-Lipschitz}\}$. The metric $\beta(\cdot,\cdot)$ metrizes the weak convergence, due to the fact that $D$ is seperable ([45] section 17).

In order to state the lemma, let us consider the functions

$$\Xi^\nu, \Xi : (D, \|\cdot\|) \to (\Pi(D^k), \beta)$$

where $\Pi(D^k)$ is the space of probability measures on $D^k$ and for any $C \in \mathcal{B}(D^k)$ the functions $\Xi^\nu$ have the form $\Xi^\nu(y)(C) = Pr((z_0,\dots,z_{k-1}) \in C)$, where $z_0$ has a distribution concentrated on $y$ and $z_{i+1} \sim \bar{K}^\nu(z_i, \cdot)$. In the same way the values of $\Xi$ are defined. Thus, $\Xi$ maps the initial condition $y_0$ to the distribution of $(y_0,\dots,y_{k-1})$.

Due to the Feller continuity of $\bar{K}$, it is not difficult to show that the function $\Xi$ is continuous ([41]). The following lemma illustrates a uniformity property of the convergence of $\Xi^\nu$ to $\Xi$. Particularly, it is shown that for the same initial condition $y_0^\nu = y_0$ and large $\nu$, the distribution of $(y_0^\nu,\dots,y_{k-1}^\nu)$ is $\beta$—close to $(y_0,\dots,y_{k-1})$, uniformly in $y_0$.

*Lemma 3:* Under the assumptions stated above, for any $\varepsilon > 0$, there exists a positive integer $\nu_0$ such that $\beta(\Xi^\nu(y), \Xi(y)) < \varepsilon$ for any $\nu \geq \nu_0$ and any $y \in D$.

*Proof:* To contradict, assume that there is a positive constant $\varepsilon$ such that for any $\nu_0 \in \mathbb{N}$, there exists a $\nu \geq \nu_0$ and a $y \in D$ with $\beta(\Xi^\nu(y), \Xi(y)) > \varepsilon$. Then there exist sequences $m_\nu, y_{m_\nu}$ such that $m_\nu \geq \nu$, $m_\nu > m_{\nu-1}$ and $\beta(\Xi^{m_\nu}(y_{m_\nu}), \Xi(y_{m_\nu})) > \varepsilon$. The compactness of $D$ implies the existence of a further subsequence $y_{m_{\nu_l}}$ that converges to a value $\bar{y}$. Theorem 1 of [41] implies that $\beta(\Xi^{m_{\nu_l}}(y_{m_{\nu_l}}), \Xi(\bar{y})) \to 0$.

However, the triangle inequality implies

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

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

*Remark 15:* If the functions $\Xi^\nu$ are continuous for large $\nu$, the proof of Lemma 3 becomes trivial. □

Let us then turn back to the proof of Proposition 3.

*Proof of Proposition 3:* The quantities

$$g_j(y_0) = f_j\left(E\begin{bmatrix}\bar{f}_1(y_0,\dots,y_{k-1})\\ \vdots \\ f_{n^2}(y_0,\dots,y_{k-1})\end{bmatrix}\right) \tag{57}$$

are continuous functions of $y_0$. Thus, due to the compactness of $D$, there is a constant $\varepsilon_1 > 0$, with $g_j(y) > \varepsilon_1$ for any $y \in D$ and any $j = 1,\dots,n$. The functions $f_j(\cdot)$ are uniformly continuous in $D$ and thus there exists a positive constant $\delta_1$ such that $f_j(v_1) > \varepsilon_1$ implies $f_j(v_2) > 0$, for any $v_2 \in D^k$ with $\|v_1 - v_2\| < \delta_1$ and any $j = 1,\dots,n$.

Choose $y_0 = y_0^\nu$. The entries of the functions $f_j$, i.e., $E[\bar{f}_i(y_0^\nu,\dots,y_{k-1}^\nu)]$ and $E[\bar{f}_i(y_0,\dots,y_{k-1})]$, $i = 1,\dots,n^2$ can be written in the form

$$\int \bar{f}_i(w)\,(\Xi(y_0))\,(dw) \quad \text{and} \quad \int \bar{f}_i(w)\,(\Xi^\nu(y_0))\,(dw).$$

We claim that:

*Claim 1:* For large $\nu$, it holds

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

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

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

$$\begin{aligned}&\left|\int \bar{f}_i(w)\,(\Xi(y_0))\,(dw) - \int \bar{f}_i(w)\,(\Xi^\nu(y_0))\,(dw)\right| \\ &\quad\leq \int \left|\bar{f}_i(w) - \bar{f}_i'(w)\right|(\Xi(y_0) + \Xi^\nu(y_0))(dw) \\ &\qquad + \left|\int \bar{f}_i'(w)(\Xi(y_0))(dw) - \int \bar{f}_i'(w)(\Xi^\nu(y_0))(dw)\right|.\end{aligned}$$

The first term is bounded by $\delta_1/(2n^2)$, due to the fact that $\Xi(y_0)(\cdot)$ and $\Xi^\nu(y_0)(\cdot)$ are probability measures. To bound the second term, let us observe that Lemma 3 implies the existence of a positive integer $\nu_0$ such that $\beta\left(\Xi^\nu(y_0), \Xi(y_0)\right) < \delta_1/(2n^2\bar{L})$, for any $\nu \geq \nu_0$. This completes the proof of Claim 1.

Therefore, for $\nu \geq \nu_0$ it holds

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

for any $y_0^\nu \in D$ and $j = 1,\ldots,n$. Hence

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

Integrating over the distribution of $x_0^\nu, y_0^\nu$, we conclude

$$E\left[\left(x_k^\nu\right)^T x_k^\nu\right] < aE\left[\left(x_0^\nu\right)^T x_0^\nu\right] \tag{58}$$

for any initial condition and any $\nu \geq \nu_0$, which completes the proof of the proposition. □

Let us then state a corollary of the Proposition 3, dealing with systems having also additive stochastic disturbance. Particularly, consider the systems given by

$$x_{k+1}^\nu = A\left(y_k^\nu\right)x_k^\nu + w_k^\nu, \quad y_{k+1}^\nu \sim \bar{K}^\nu\left(y_k^\nu,\cdot\right) \tag{59}$$

and

$$x_{k+1} = A(y_k)x_k + w_k, \quad y_{k+1} \sim \bar{K}(y_k,\cdot) \tag{60}$$

where $w_k^\nu$ and $w_k$ are zero mean i.i.d. random variables with finite variances.

*Corollary 2:* Consider the systems described by (59) and (60). Assume that $\bar{K}^\nu \to \bar{K}$ weakly and that $\bar{K}$ is Feller continuous. Let $a < 1$ and assume that $T = T_{A,\bar{K}}$ has spectral radius less than $1/a$. Then for any $\varepsilon > 0$, there exist positive integers $k_0, \nu_0$ such that

$$E\left[\sum_{k=k_0}^{\infty} a^k\left(x_k^\nu\right)^T x_k^\nu\right] \leq \varepsilon\left(1 + E\left[\left(x_0^\nu\right)^T x_0^\nu\right]\right) \tag{61}$$

$$E\left[\sum_{k=k_0}^{\infty} a^k x_k^T x_k\right] \leq \varepsilon\left(1 + E\left[x_0^T x_0\right]\right) \tag{62}$$

for any $\nu \geq \nu_0$.

*Proof:* The proof is straightforward and uses equation (50), as well as the bound (58) and techniques form [38]. □

### C. Weak Convergence and $\varepsilon$—Optimality

In what follows, we assume that $\bar{K}^\nu \to \bar{K}$ weakly, $\bar{K}$ is Feller continuous and the functions $A^\nu(y,k), A(y,k), A(y)$ are continuous on their $y$ argument. Let us then introduce some notation, needed in order to state the basic results.

*Notation:* Consider the system given by

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

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

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

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

The basic topic of this section is the proof of the following two propositions about the $\varepsilon$—optimality in the finite and infinite horizon LQ control problems respectively.

*Proposition 4:* Assume that $A^\nu(y_k,k) \to A(y_k,k)$ as $\nu \to \infty$. Let us denote by $u_k = L_k(y_k)x_k$ the optimal control law that attains the minimum $J^\star_{\bar{K},A(y_k,k),k_0}$. Then for any $\varepsilon > 0$ there exists a positive integer $\nu_0$ such that

$$J_{\bar{K}^\nu,k_0,A^\nu(y_k,k),u_k=L_k(y_k)x_k}(x_0,y_0) \leq J^\star_{\bar{K}^\nu,k_0,A(y_k,k)} + \varepsilon\left(1 + x_0^T x_0\right) \tag{63}$$

for any $\nu \geq \nu_0$. □

*Proposition 5:* Let us denote by $u_k = L(y_k)x_k$ the feedback strategy that attains the minimum $J^\star_{\bar{K},\infty}$. Then, for any $\varepsilon > 0$ there exists a positive integer $\nu_0$ such that

$$J_{\bar{K}^\nu,\infty,u_k=L(y_k)x_k}(x_0,y_0) \leq J^\star_{\bar{K}^\nu,\infty} + \varepsilon\left(1 + x_0^T x_0\right) \tag{64}$$

for any $\nu \geq \nu_0$. □

The proof of the Propositions 4 and 5 depends on the following lemmas.

*Lemma 4:* Consider a feedback control law $u_k = L_k(y_k)$ which is continuous in $y_k$. Then for any $\varepsilon > 0$, there exists a positive integer $\nu_0$ such that

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

for any $\nu \geq \nu_0$.

*Proof:* The proof is a direct consequence of the properties of the weak convergence [41]. □

*Lemma 5:* Let $f_\nu : D \to \mathbb{R}$ be a sequence of continuous functions and $f$ their pointwise limit. Then, it holds

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

as $\nu \to \infty$.

*Proof:* The proof is a direct consequence of the compactness of $D$ and the properties of the weak convergence. □

*Lemma 6:* For any $\varepsilon > 0$, there exists a positive integer $\nu_0$ such that

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

for any $\nu \geq \nu_0$.

*Proof:* The proof proceeds backwards in time from the step $k_0$ to 0. At each step the Dynamic programming equations, as well as the Lemma 5 are used. □

We then proceed to the proof of the basic results of the current section.

*Proof of Proposition 4:* Lemma 4 implies the existence of a positive integer $\nu_{01}$ such that

$$\begin{aligned} J_{\bar{K}^n u,k_0,u_k=L_k(y_k)x_k}&(x_0,y_0)\\ &< J_{\bar{K},k_0,u_k=L_k(y_k)x_k}(x_0,y_0) + \varepsilon(1 + x_0^T x_0)\\ &= J^\star_{\bar{K},k_0}(x_0,y_0) + \varepsilon(1 + x_0^T x_0) \end{aligned} \tag{65}$$

for any $\nu \geq \nu_{01}$. On the other hand, Lemma 6 implies the existence of a positive integer $\nu_{02}$ such that

$$J^\star_{\bar{K},k_0}(x_0,y_0) < J^\star_{\bar{K}^\nu,k_0}(x_0,y_0) + \varepsilon\left(1 + x_0^T x_0\right) \qquad (66)$$

for any $\nu \geq \nu_{02}$. Inequalities (65) and (65) imply the desired result for $\nu_0 = max\{\nu_{01}, \nu_{02}\}$. □

*Proof of Proposition 5:* In order to complete the proof, the following series of comparisons is made:
$J_{\bar{K}^\nu,\infty,u_k=L(y_k)x_k}$, $J_{\bar{K}^\nu,k_0,u_k=L(y_k)x_k}$, $J_{\bar{K},k_0,u_k=L(y_k)x_k}$, $J^\star_{\bar{K},\infty}$, $J^\star_{\bar{K},k_0'}$, $J^\star_{\bar{K}^\nu,k_0'}$, $J^\star_{\bar{K}^\nu,\infty}$.

Particularly, each of these quantities is compared to the next one. It is shown that each of these quantities is at most slightly larger than the next.

At first let us compare $J_{\bar{K}^\nu,\infty,u_k=L(y_k)x_k}$ with $J_{\bar{K}^\nu,k_0,u_k=L(y_k)x_k}$. Corollary 2 implies that for any $\varepsilon > 0$ there exist integers $k_0$ and $\nu_{01}$ such that

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

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

Lemma 4 implies the existence of a positive integer $\nu_{02}$ such that

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

for any $\nu \geq \nu_{02}$, which serves as the second comparison.

Comparison three holds as an inequality, i.e.,

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

To derive an inequality for the fourth comparison, let us observe that $K_k \to K$ and $c_k \to c$ uniformly as $k \to \infty$, where $K_k, K, c_k$ and $c$ are as in Proposition 3 and Theorem 2 of [38]. It also holds $J^\star_{\bar{K},k} = x_0^T K_k(y_0)x_0 + c_k$ and $J^\star_{\bar{K},\infty} = x_0^T K(y_0)x_0 + c$. Thus, for any $\varepsilon > 0$, there exists a positive integer $k_0'$ such that

$$J_{\bar{K},\infty}(x_0,y_0) \leq J^\star_{\bar{K},k_0'}(x_0,y_0) + \varepsilon\left(1 + x_0^T x_0\right)/4.$$

Lemma 6 implies the existence of a positive integer $\nu_{03}$ such that

$$J_{\bar{K},k_0'}(x_0,y_0) \leq J^\star_{\bar{K}^\nu,k_0'}(x_0,y_0) + \varepsilon\left(1 + x_0^T x_0\right)/4$$

for any $\nu \geq \nu_{03}$. This inequality shows the fifth comparison.

The last comparison holds as an inequality, i.e.,

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

Hence, choosing $\nu_0 = \max\{\nu_{01}, \nu_{02}, \nu_{03}\}$ the desired result is shown. □

### D. *Proof of Theorem 1*

The proof is a direct consequence of the Propositions 4 and 5 and the continuity of the functions involved.

## REFERENCES


[1] A. W. Starr and Y. Ho, "Nonzero sum differential games," *J. Optim. Theory Appl.*, vol. 3, no. 3, pp. 184–206, 1969.

[2] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM, 1999.

[3] P. Samuelson, "An exact consumption loan model of interest with or without the social contrivance of money," *J. Political Economy*, vol. 66, no. 6, pp. 467–482, 1958.

[4] Y. Balasko and K. Shell, "The overlapping-generations model, i: The case of pure exchange without money," *J. Economic Theory*, vol. 23, pp. 281–306, 1980.

[5] S. Jorgensen and D. Yeung, "Inter- and intragenerational renewablere-source extraction," *Annals Operat. Res.*, vol. 88, pp. 275–289, 1999.

[6] R. Myerson, "Population uncertainty and poisson games," *Int. J. Game Theory*, vol. 27, no. 3, pp. 375–392, 1998.

[7] I. Arribas and A. Urbano, "Repeated games with probabilistic horizon," *Math. Social Sci.*, vol. 50, no. 1, pp. 39–60, 2005.

[8] D. W. K. Yeung and L. A. Petrosyan, "Subgame consistent cooperative solution of dynamic games with random horizon," *J. Optim. Theory Appl.*, vol. 150, no. 1, pp. 78–97, 2011.

[9] S. Jorgensen and D. Yeung, "An overlapping generations stochastic deferential game," *Automatica*, vol. 41, pp. 69–74, 2005.

[10] N. Kakogiannis and G. P. Papavassilopoulos, "Games with players having long term and short term horizons," in *Proc. 14th Int. Symp. Dyn. Games*, Banff, AB, Canada, Jul. 19–23, 2010.

[11] G. P. Papavassilopoulos, "University-students game," *Dyn. Games Appl.*, pp. 1–32, 2012.

[12] M. Jungers, G. P. Papavassilopoulos, and H. Abou-Kandil, "Feedback nash strategy for games with player-dependent time horizons," in *Proc. 15th IFAC Workshop Control Appl. Optim.*, Rimini, Italy, Sep. 13–16, 2012, pp. 45–50.

[13] G. P. Papavassilopoulos, H. Abou-Kandil, and M. Jungers, "A nash game with long-term and short-term players," in *Proc. 52nd IEEE Conf. Decision Control*, Firenze, Italy, Dec. 10–13, 2013, pp. 1726–1731.

[14] D. Fudenberg, D. M. Kreps, and E. S. Maskin, "Repeated games with long-run and short-run players," *Rev. Economic Studies*, vol. 57, no. 4, pp. 555–573, 1990.

[15] D. Kreps and R. Wilson, "Repeated games with long-run and short-run players," *J. Economic Theory*, vol. 27, pp. 253–279, 1982.

[16] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.

[17] J. W. Milnor and L. S. Shapley, "Values of large games ii: Oceanic games," *Math. Oper. Res.*, vol. 3, no. 4, pp. 290–307, 1978.

[18] R. J. Aumann, "Markets with a continuum of traders," *Econometrica*, vol. 32, pp. 39–50, 1964.

[19] G. Owen, *Game Theory*. Philadelphia, PA, USA: W.B. Saunders Company, 1969.

[20] B. Jovanovic and R. W. Rosenthal, "Anonymous sequential games," *J. Math. Economics*, vol. 17, no. 1, pp. 77–87, 1988.

[21] J. Bergin and D. Bernhardt, "Anonymous sequential games with aggregate uncertainty," *J. Math. Economics*, vol. 21, no. 6, pp. 543–562, 1992.

[22] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," in *Proc. NIPS'05*, 2005, pp. 1–8.

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

[24] O. Gueant, J. Lasry, and P. Lions, *Mean Field Games and Applications*. Paris, France: Springer, 2010.

[25] M. Huang, R. P. Malhame, and P. E. Caines, "Nash equilibria for large-population linear stochastic systems of weakly coupled agents," in *Analysis, Control and Optimization of Complex Dynamical Systems*. New York: Springer, 2005, ch. 9, pp. 215–252.

[26] M. Huang, R. Malhame, and P. Caines, "Large population stochastic dynamic games: Closed-loop mckean-vlasov systems and the nash certainty equivalence principle," *Commun. Inform. Syst.*, vol. 6, no. 2, pp. 221–251, 2006.

[27] M. Bardi, "Explicit solutions of some linear-quadratic mean field games," *Netw. Heterogeneous Media*, vol. 7, no. 2, pp. 243–261, 2012.

[28] A. Bensoussan, J. Sung, P. Yam, and S. P. Yung, "Linear-quadratic mean field games," *arXiv preprint arXiv:1404.5741*, 2014.

[29] G. P. Papavassilopoulos, "Limiting behavior of lq deterministic infinite horizon nash games with symmetric players as the number of players goes to infinity," *arXiv preprint arXiv:1403.3300*, 2014.

[30] H. Tembine, Q. Zhu, and T. Basar, "Risk-sensitive mean-field games," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 835–850, Apr. 2014.

[31] D. Bauso, H. Tembine, T. Basar *et al.*, "Robust mean field games with application to production of an exhaustible resource," in *Proc. 7th IFAC Symp. Robust Control Design*, Aalborg, Denmark, 2012, pp. 454–459.
[32] V. N. Kolokoltsov, J. Li, and W. Yang, "Mean field games and nonlinear markov processes," *arXiv preprint arXiv:1112.3744*, 2011.
[33] M. Huang, "Large-population l.q.g. games involving a major player: The nash certainty equivalence principle," *SIAM J. Control Optim.*, vol. 48, no. 5, pp. 3318–3353, 2010.
[34] H. Tembine, "Large-scale games in large-scale systems," in *Proc. 5th Int. ICST Conf. Perform. Eval. Methodologies Tools*, 2011, pp. 9–17.
[35] R. Selten, "Reexamination of the perfectness concept for equilibrium points in extensive games," *Int. J. Game Theory*, vol. 4, pp. 25–55, 1975.
[36] T. Basar, "On the uniqueness of the nash solution in linear-quadratic differential games," *Int. J. Game Theory*, vol. 5, pp. 65–90, 1975.
[37] H. Tembine and M. Huang, "Mean field difference games: Mckean-vlasov dynamics," in *Proc. 50th 50th IEEE Conf. Decision Control & Eur. Control Conf. (CDC-ECC)*, 2011, pp. 1006–1011.
[38] I. Kordonis and G. P. Papavassilopoulos, "Stability and lq control of mjls with a markov chain with general state space," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 535–540, Feb. 2014.
[39] 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," *J. Optim. Theory Appl.*, vol. 28, pp. 49–75, 1979.
[40] 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.
[41] A. F. Karr, "Weak convergence of a sequence of markov chains," *Wahrscheinlichkeitstheorie und Verw. Gebiete*, vol. 33, pp. 41–48, 1975.
[42] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. London, U.K.: Springer-Verlag, 1993.
[43] G. Freiling, G. Jank, and H. Abou-Kandil, "On global existence of solutions to coupled matrix riccati equations in closed-loop nash games," *IEEE Trans. Autom. Control*, vol. 41, pp. 264–269, 1996.
[44] W. Arveson, *A Short Course on Spectral Theory*. New York: Springer-Verlag, 2002.
[45] D. Panchenko, *18.175 Theory of Probability fall-2008*. Cambridge, MA, USA: Massachusetts Institute of Technology, 2008, MIT OpenCourseWare License: Creative Commons BY-NC-SA. [Online]. Available: http://ocw.mit.edu

**Ioannis Kordonis** (M'14) received the Diploma in electrical and computer engineering from the National Technical University of Athens, Athens, Greece, in 2009, where he is currently pursuing the Ph.D. degree.

His research interests include Game Theory with an emphasis on Dynamic Games and Stochastic Control theory. He is also interested in applications in the area of Energy—Power Systems Modeling and Control and in Bioengineering.

**George P. Papavassilopoulos** (SM'96) received the Diploma from the Department of Mechanical and Electrical Engineering, National Technical University of Athens, Athens, Greece, in 1975 and the M.S. and Ph.D. degrees from the Department of Electrical Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 1977 and 1979, respectively.

He was a Professor in the Department of EE-Systems, University of Southern California, Los Angeles, CA, USA, from 1979 to 2000 when he joined the Department of Electrical and Computer Engineering, NTUA, as Professor. He has conducted research in dynamic and stochastic game theory, optimization and control theory, parallel algorithms, Markovian learning, BMIs for robust control. He is also interested in applications in biomedical engineering, energy and telecommunication policies.