HAL
archives-ouvertes.fr

Perturbation of Linear Quadratic Systems with Jump Parameters and Hybrid Controls

Rachid El-Azouzi, Mohammed Abbad, Eitan Altman

To cite this version:

Rachid El-Azouzi, Mohammed Abbad, Eitan Altman. Perturbation of Linear Quadratic Systems with Jump Parameters and Hybrid Controls. RR-3816, INRIA. 1999. <inria-00072842>

HAL Id: inria-00072842

https://hal.inria.fr/inria-00072842

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Perturbation of Linear Quadratic Systems with Jump Parameters and Hybrid Controls

Rachid El Azouzi — Mohammed Abbad — Eitan Altman

N° 3816

Novembre 1999

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--3816--FR+ENG

Perturbation of Linear Quadratic Systems with Jump Parameters and Hybrid Controls

Rachid El Azouzi * , Mohammed Abbad † , Eitan Altman ‡

Thème 1 — Réseaux et systèmes
Projet MISTRAL

Rapport de recherche n° 3816 — Novembre 1999 — 21 pages

Abstract: We consider the problem of the perturbation of a class of linear-quadratic differential games with piecewise deterministic dynamics, where the changes from one structure (for the dynamics) to another are governed by a finite-state Markov process. Player 1 controls the continuous dynamics, whereas Player 2 controls the rate of transition for the finite-state Markov process; both have access to the states of both processes. Player 1 wishes to minimize a given quadratic performance index, while player 2 wishes to maximize or minimize the same quantity. The problem above leads to the analysis of some linearly coupled set of quadratic equations (Riccati Equation). We obtain a Taylor expansion in the perturbation for the solution of the equation for a fixed stationary policy of the player 2. This allows us to solve the game or team problem as a function of the perturbation.

Key-words: Sinsular erturbation, hybride stochastic systems, stocastic game, continuous-time Markov chain, small parametre, averaging, lexicographical optimisation, Puiseux serie and infinite horizon

* INRIA B.P 93 2204 Route des Lucioles 06902 Sophia Antipolis Cedex France
† Département de Mathématique et Informatique Faculté des sciences B.P 1014 Rabat Maroc
‡ INRIA B.P 93 2204 Route des Lucioles 06902 Sophia Antipolis Cedex France

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 —Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Systèmes linéaires quadratiques avec sauts rapides

Résumé : Nous considérons un problème de perturbations des systèmes linéaires quadratiques où les paramètres sont des fonctions d'une chaîne de Markov. Le joueur 1 contrôle la dynamique continue, tandis que le joueur 2 contrôle le taux de transition de la chaîne de Markov. Le joueur 1 désire minimiser un coût quadratique et le joueur 2 dèsire minimiser ou maximiser la même quantitè. Nous analysons le comportement asymptotique de ce système quand le taux de transition de la chaîne de Markov contrôllée devient grand, ce qui amène à étudier une équation de Riccati. Nous obtenons une série de Taylor dans la perturbation de la solution de l'équation de Riccati pour une stratégie stationnaire fixée.

Mots-clés : Perturbation singulier, Processus de décision Markovien, Théorie des jeux, Serie de Puiseux, Optimisation lexicographique, Agregation et Horizon infini.

1 Introduction

We consider in this paper the optimal control of a perturbed class of linear-quadratic differential games with Markov jump parameters, when the transition probabilities associated with the underlying Markov chain are also controlled. We study two scenarios: The transition probabilities are controlled cooperatively (which leads to a team problem) or they are controlled antagonistically (which leads to a zero-sum differential game).
Research on control systems with Markovian switching structure was initiated more than thirty years ago by Krassovskii and Liskii [9],[10] and Florentin [11], with follow-up work in the late sixties and early seventies addressing the stochasstic maximum principle [12],[13], dynamic programming [13], and linear-quadratic control [14],[15] in this context. Late eighties and early nineties have witnessed renewed interest in the topic, with concentrated research on theoretical issues like controllability and stabilizability [19],[6],[21],[4],[8] in linear-quadratic systems for continuous time, see also [18],[5],[17] for discrete-time. Perhaps the first theoretical development in differential games context was reported in [3], where a general model was adopted that allows in a multiple player environment the Markov jump process (also controlled by the players) to determine the mode of player, in addition to affecting the system and cost structure. Results in [16] also apply, as a special case, to zero-sum differential games with Markov jump parameters and state feedback information for both players. More recently, this framework has also been used in worst-case (H^∞) control design problems where the underlying system changes stucture at random points of time according to a finite-state Markov process with a constant rate matrix, and with the system dynamics affected by an additional (continuous-valued) disturbance process; some selected publications on this topic are [19], [2], [21], and [7].

In this paper, we restrict to the case in which the linear dynamics is described by a one-dimensional state; it is only for this case that linear-quadratic structure is known to hold when the transition of the jump parameters are controlled as well. We consider perturbation of the non-controlled system and show that the solution of the related coupled Algebraic Riccati Equation (ARE) has a Taylor expansion in the perturbation. This is then extended to the controlled case. We show that there exists a policy for the controller of the Markov jumps which is optimal for all small enough perturbations. Numerical algorithms are proposed to compute the l first terms of the solution of the Coupled Riccati Differential Equation when the transition of the jump parameters are controlled which leads to a good approximation of the optimal policy.

The paper is organized as follows. Section 2 and 3 introduces the general model, and presents the main results in [1]. By using the averaging andagregation techniques, we study computational algorithms to determine solutions of linearly coupled set of quadratic equation (Riccati Equation) and optimal policy in sections 4 and 5, the paper ends with the concluding remarks of section 6.

2 General Model

The continuous dynamics of the game under consideration evolve according to

$$\frac{dx}{dt} = A(\theta)x + B(\theta)u^1, \qquad x(0) = x_0, \tag{1}$$

where $x \in \mathbb{R}$, x_0 is a fixed (known) initial state, u^1 is a control, applied by Player 1, taking values in $U_1 = \mathbb{R}$, and $\theta(t)$ is a controlled, continuous time Markov process, taking values in a finite state space S, of cardinality s. Transitions from state i to j occur at a rate controlled by Player 2, who chooses at time t an action $u^2(t)$ among a finite set $U_2(i)$ of actions available at state i. Let $U_2 := \cup_{i\in S} U_2(i)$. The controlled rate matrix (of transitions within S) is

$$\Lambda = (\lambda_{iaj}), \qquad i, j \in S, \quad a \in U_2(i).$$

The λ_{iaj}'s are real numbers such that for any $i \neq j$, and $a \in U_2(i)$, $\lambda_{iaj} \geq 0$, and for all $a \in U_2(i)$ and $i \in S$, $\lambda_{iai} = -\sum_{j\neq i} \lambda_{iaj}$.
Fix some initial state i_0 of the controlled Markov chain S, consider the class of policies $\gamma^k \in U_k$ for player k (k=1,2), whose elements (taking values in U_k) are of the form

$$u^k(t) = \gamma^k(t, x_{[0,t]}, \theta_{[0,t]}), \qquad t \in [0, +\infty). \tag{2}$$

Here, γ^k is taken to be piecewise continuous in its first argument, and piecewise Lipschitz continuous in its second argument.
Define $\chi = \mathbb{R} \times S$ to be the global state space of the system and $U := U_1 \times U_2$ to be the class of multi-strategies $\gamma := (\gamma^1, \gamma^2)$. Define the running (immediate) cost $L : \chi \times U_1 \to \mathbb{R}$ as:

$$L(x, i, u^1) = \mid x \mid^2_{Q(i)} + \mid u^1 \mid^2_{R(i)}, \tag{3}$$

where $Q(.) \geq 0$ and $R(.) > 0$, and $\mid x \mid_Q$ denotes the Euclidean semi-norm.
To any fixed initial state (x_0, i_0) and a multi-strategy $\gamma \in U$, there corresponds a unique probability measure $P^\gamma_{x_0,i_0}$ on the canonical probability space Ω of the states and actions of the players, equipped with the standard Borel σ-algebra. Denote by $E^\gamma_{x_0,i_o}$ the expectation

operator corresponding to $P^{\gamma}_{x_0,i_0}$. We denote by $(x(t),\theta(t)), u(t), t \in [0,+\infty)$, the stochastic processes corresponding to the states and actions, respectively. In terms of this notation and convention, and for each fixed initial state (x_0, i_0) and a multy-strategy $\gamma \in U$, we introduce the discounted (expected) cost function:

$$J_\beta(x_0,i_0,\gamma) = E^{\gamma}_{x_0,i_0}\left(\int_0^{+\infty} e^{-\beta t} L(x(t),\theta(t),u^1(t))dt\right), \tag{4}$$

where $\beta \geq 0$ is a discount factor.

Condition 1:

The pair $(A(\theta(t)), B(\theta(t)))$ is stochastically stabilizable for all stationary policies of Player 2, and $(A(i), Q(i))$ is observable for each $i \in S$.

The problem is the derivation of a solution to

$$\hat{J}_\beta(x,i) = val_{\gamma\in U} J_\beta(x,i,\gamma), \tag{5}$$

where $val_{\gamma\in U}$ stands for the

$$\inf_{\gamma^1\in U_1}\sup_{\gamma^2\in U_2} \quad and \quad \inf_{\gamma^1\in U_1}\inf_{\gamma^2\in U_2}.$$

3 The unperturbed solution process

We briefly mention results for the unperturbed case Altman et al.[1]
Introduce a backward controlled Markov operator G^γ associated with (1)-(4) (for n=1) as follow: For each $\varphi(x,i)$ for which $\varphi(.,i) \in C^1$ for all $i \in S$ (where C^1 is the space of all functions with continuous partial derivatives), and for each $u = (u^1, a) = (u^1, u^2) \in U(i)$,

$$G^u\varphi(x,i) = \frac{\partial\varphi}{\partial x} f(x,u^1,i) + \sum_{j\in S}\lambda_{iaj}\varphi(x,i),$$

where $f(x,u^1,i) = A(i)x + Bu^1$ and $x(t) \in \mathbb{R}$.
Next we introduce the dynamic programming equations associated with the team and game problems above , where the continuous-differentiability of the value functions will be replaced simply by continuity, and their solutions interpreted in the generalized (viscosity) sense if differentiability fails. The Hamilton-Jacobi-Bellman equation is:

$$\beta\varphi(x,i) = val_{u^1,u^2}[G^{(u^1,u^2)}\varphi(x,i) + L(x,i,u^1)]. \tag{6}$$

Associate with (6) the corresponding set C^1, of functions φ in x. We now stipulate a structure for $\varphi(x,i)$ that is quadratic in x:

$$\varphi(x,i) = x^2 P(i), \quad i \in S, \tag{7}$$

where $P(i)$ is a real number for each $i \in S$. Substituting this structural form into (6), we obtain:

$$\begin{aligned} \beta x^2 P(i) &= \min_{u^1} \left[2(A(i)x + B(i)u^1)P(i)x + |u^1|^2_{R(i)} \right] + x^2 Q(i) \\ &\quad + x^2 opt_a \sum_{j \in S} \lambda_{iaj} P(j), \end{aligned} \tag{8}$$

where "opt" stands for minimization in the team problem, and maximization in the game problem. The unique minimizing control u^1 in (8) is $\gamma^1_{opt}(x,i) = -R^{-1}(i)B(i)P(i)x$, whose substituting into (8) leads to:

$$\beta P(i) = Q(i) + 2A(i)P(i) - N(i)[P(i)]^2 + opt_{a \in U_2(i)} \sum_{j \in S} \lambda_{iaj} P(j), \tag{9}$$

where $N(i) = [B(i)]^2/R(i)$. We now provide a precise statement of this result in the following theorem.

Theorem 3.1 *Let the condition 1 be satisfied, with $B(i) \neq 0$ and $Q(i) > 0$ $\forall i \in S$. Then:*
(i) The coupled set of equation (9) admits a unique positive solution, P^*.
(ii) There exists a unique stationary optimal policy for Player 1, given by

$$\gamma^{1^*}(x,i) = -[B(i)P^*(i)/R(i)]x, \quad i \in S.$$

(iii) There is an optimal (minimizing in the team case, and maximizing in the game case) stationary policy for Player 2.
(iv) The optimal value $\hat{J}_\beta(x,i) = x^2 P^*(i)$

Proof: See Theorem 3.5 in Altman et al.[1]. ∎

4 The Perturbed Solution Process for a fixed policy of player 2.

We now introduce the perturbation for the transition probability rate, i.e

$$\lambda^\epsilon_{iaj} = \lambda_{iaj} + \frac{1}{\epsilon}\mu_{iaj}, \tag{10}$$

where fore each $(\lambda_{iaj})_{s\times s}$ and $(\mu_{iaj})_{s\times s}$ are the transition probability rate matrices .
By using standard theory of unperturbed case, the resolution of pertubed problem leads to the analysis of some perturbed linearly Coupled Riccati Differential Equation given by:

$$\beta P_\epsilon(i) = 2A(i)P_\epsilon(i) - [P_\epsilon(i)]^2 N(i) + Q(i) + opt_{a\in U_2(i)} \sum_{j\in S} \lambda^\epsilon_{iaj} P_\epsilon(j), \quad i \in S, \tag{11}$$

with $Q(i) > 0$ and $N(i) \neq 0, \quad i \in S$.
The problem addressed in this paper is to find a solution to the equation (11) for any $\epsilon \in (0, h]$ (where h is small enough).
Let F denote the set of stationary policies for player 2 that depend only on the current state of the Markov chain. With a fixed $f \in F$, we denote the solution of (11) by $P^f_\epsilon(i), i \in S$. For each fixed $f \in F$, (11) is precisely the set of coupled algebraic Riccati equation that arises in the quadratique optimal control, we can conclude that the linearly coupled set of quadratique equations:

$$\beta P^f_\epsilon(i) = Q(i) + 2A(i)P^f_\epsilon(i) - N(i)[P^f_\epsilon(i)]^2 + \sum_{j\in S}(\lambda_{if_ij} + \frac{1}{\epsilon}\mu_{if_ij})P^f_\epsilon(j), \qquad i \in S. \tag{12}$$

admits a unique positive solution set $P^f_\epsilon(i)$ provided that $N(i) \neq 0$ and $Q(i) > 0$, $\forall i \in S$. Since f is fixed, let denote P^f_ϵ by P_ϵ.

Theorem 4.1 *The solution of equation (12) is a Puiseux series i.e, there exists a positive integer M such that*

$$P_\epsilon(i) = \sum_{n=-M}^{+\infty} P_n(i)\epsilon^{\frac{n}{M}}.$$

Proof
>From Puiseux's Theorem [20], we have

$$P_\epsilon(i) = \sum_{n=-K}^{+\infty} P_n(i)\epsilon^{\frac{n}{M}}.$$

>From (12), if -K<-M then for $-K \leq n < -M$, we obtain: $N(i)[P_n(i)]^2 = 0 \ \forall i \in S$. Since $N(i) \neq 0$, then $P_{-n}(i) = 0 \ \forall i \in S$. ∎

>From Theorem 4.1 and (12) it follows that:

$$(\beta - 2A(i)) \sum_{n=-M}^{+\infty} P_n(i)\epsilon^{\frac{n}{M}} \quad = \quad Q(i) - N(i) \sum_{n=-2M}^{+\infty} \sum_{k=-2M}^{n} P_{k+M}(i)P_{n-(k+M)}(i)\epsilon^{\frac{n}{M}}$$

$$+\sum_{n=-2M}^{+\infty}\sum_{j\in S}\mu_{if_ij}P_{n+M}(j)\epsilon^{\frac{n}{M}}+\sum_{n=-M}^{+\infty}\sum_{j\in S}\lambda_{if_ij}P_n(j)\epsilon^{\frac{n}{M}},\ i\in S.$$

Then we obtain the following set of equations:
If $-2M\leq n<-M$, then

$$0=-N(i)\sum_{k=-2M}^{n}P_{k+M}(i)P_{n-(k+M)}(i)+\sum_{j\in S}\mu_{if_ij}P_{n+M}(j),\quad i\in S.$$

If $-M\leq n<0$, then

$$(\beta-2A(i))P_n(i)=-N(i)\sum_{k=-2M}^{n}P_{M+k}(i)P_{n-(k+M)}(i)+\sum_{j\in S}\lambda_{if_ij}P_n(j)+\sum_{j\in S}\mu_{if_ij}P_{n+M}(j),\quad i\in S.$$

If $n=0$, then

$$(\beta-2A(i))P_0(i)=Q(i)-N(i)\sum_{k=-2M}^{0}P_{M+k}(i)P_{-(k+M)}(i)+\sum_{j\in S}\lambda_{if_ij}P_0(j)+\sum_{j\in S}\mu_{if_ij}P_M(j),\ i\in S.$$

If $n\geq 0$, then

$$(\beta-2A(i))P_n(i)=-N(i)\sum_{K=-2M}^{n}P_{M+K}(i)P_{n+(K+M)}(i)+\sum_{j\in S}\lambda_{if_ij}P_n(j)+\sum_{j\in S}\mu_{if_ij}P_{n+M}(j),\ i\in S.$$

We consider the Markov chain associated with the transition probability rates $(\mu_{if_ij})_{i,j\in S}$. There exists a partition of S into a family of m recurrent classes and, a transient class T, $\bar{\xi}$ will denote the class of the recurrent state:

$$S=(\sqcup_{n=1}^{m}\bar{\xi}_n)\sqcup T\quad with\quad \bar{\xi}_n\sqcap\bar{\xi}_{n'}=\emptyset\quad if\quad n\neq n'.$$

Hence

$$\mu_{if_ij}=0\quad if\quad i\in\bar{\xi}_n\quad and\quad j\in\bar{\xi}_{n'},\quad n\neq n'.$$

To each class $\bar{\xi}$ is associated the invariant measure (row vector) $m_{\bar{\xi}}$ of the recurrent subchain defined on the class $\bar{\xi}$. Let:

$$C=(C_{ij})_{1\leq i,j\leq s}\quad where\quad C_{ij}=\mu_{if_ij}.$$

We shall denote by $q_{\bar{\xi}}(i)$ the probability to end in the class $\bar{\xi}\in\bar{S}=\{\bar{\xi_1},..,\bar{\xi_m}\}$ starting from i. The m functions (column vectors) $q_{\bar{\xi}}(.)$ are the solutions to $Cv=0$, and form a basis of the m-dimensional subspace Ker(C).

Remark 1 *If v is a solution to $Cv = 0$, then*

$$\begin{aligned} v(\bar{\xi}) &= v(i), \qquad \forall i \in \bar{\xi} \\ v(i) &= \sum_{\bar{\xi} \in \bar{S}} q_{\bar{\xi}}(i) v(\bar{\xi}), \qquad i \in T, \end{aligned}$$

where $(v(\bar{\xi}))_{\bar{\xi}}$ are some real numbers.

>From the above set of equations, it follows that, if $n = -2M$, then

$$\sum_{j \in S} \mu_{if_i j} P_{-M}(j) = N(i)[P_{-M}(i)]^2 \geq 0, \quad i \in S. \tag{13}$$

Since the equation (12) admits a unique positive solution for $\epsilon \in (0, h]$, then $P_{-M}(i) \geq 0$, $i \in S$. Let $i^* = \arg\max_{i \in S} P_{-M}(i)$ then $-\mu_{i^* f_{i^*} i^*} P_{-M}(i^*) \geq \sum_{\substack{j \in S \\ j \neq i^*}} \mu_{i^* f_{i^*} j} P_{-M}(j)$. Thus $\sum_{j \in S} \mu_{i^* f_{i^*} j} P_{-M}(j) \leq 0$. >From (13), it follows that: $N(i^*)[P_{-M}(i^*)]^2 = 0$ and then $P_{-M}(i) = 0$ for all $i \in S$.
If $n = -2M + 1 < -M$, we get (since $P_{-M}(i) = 0$),

$$\sum_{j \in S} \mu_{if_i j} P_{-M+1}(j) = 0. \tag{14}$$

If $n = -2M + 2 < -M$, then

$$\sum_{j \in S} \mu_{if_i j} P_{-M+2}(j) = N(i)[P_{-M+1}(i)]^2. \tag{15}$$

>From (14) and Remark 1, we obtain:
$\bar{P}_{-M+1}(\bar{\xi}) = P_{-M+1}(i)$ for $i \in \bar{\xi}$ and $P_{-M+1}(i) = \sum_{\bar{\xi} \in \bar{S}} q_{\bar{\xi}}(i) \bar{P}_{-M+1}(\bar{\xi})$ for $i \in T$. Now, equation (15) can be written as: $\sum_{j \in S} \mu_{if_i j} P_{-M+2}(j) = N(i)[\bar{P}_{-M+1}(\bar{\xi})]^2$, $i \in \bar{\xi}$. Multiplying the last equation by $m_{\bar{\xi}}(i)$ for each $i \in \bar{\xi}$ and summing over $\bar{\xi}$, to obtain:

$$\sum_{j \in S} \left(\underbrace{\sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) \mu_{if_i j}}_{=0} \right) P_{-M+2}(j) = \bar{N}(\bar{\xi})[\bar{P}_{-M+1}(\bar{\xi})]^2,$$

where $\bar{N}(\bar{\xi}) = \sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) N(i) > 0$. Thus, $\bar{N}(\bar{\xi})[\bar{P}_{-M+1}(\bar{\xi})]^2 = 0$ and $\bar{P}_{-M+1}(\bar{\xi}) = 0$. Therefore $P_{-M+1}(i) = 0$, $i \in S$ and $\sum_{j \in S} \mu_{if_i j} P_{-M+2}(j) = 0$, $i \in S$. By using the same procedure for the case where $-2M \leq n < -M$, we obtain:

$$P_{-M}(i) = ... = P_{E(\frac{-M-1}{2})}(i) = 0 \quad i \in S, \tag{16}$$

and

$$\sum_{j\in S}\mu_{if_ij}P_k(j)=0,\quad k=E(\frac{-M+1}{2}),....,-1, \tag{17}$$

where $E(y)$ stand for the greatest integer wich smaller than or egal to y.
In the sequel, we consider two cases for $-M\leq n<0$.
Case 1: M is even.
If $n=-M$, then

$$\sum_{j\in S}\mu_{if_ij}P_0(j)=N(i)[P_{\frac{-M}{2}}(i)]^2,\qquad i\in S. \tag{18}$$

By using the same technique to derive equations (14) and (15), we obtain

$$P_{-\frac{M}{2}}(i)=0\quad and\quad \sum_{j\in S}\mu_{if_ij}P_0(j)=0. \tag{19}$$

If $n=-M+1$, then $\sum_{j\in S}\mu_{if_ij}P_1(j)=0$, $i\in S$.
If $n=-M+2<0$, then $\sum_{j\in S}\mu_{if_ij}P_2(j)=N(i)[P_{\frac{-M+2}{2}}(i)]^2$, $i\in S$. It follows that $P_{\frac{-M+2}{2}}(i)=0\ \ and\ \ \sum_{j\in S}\mu_{if_ij}P_2(j)=0$.
If we continue with the same procedure for $-M\leq n<0$, it follows that:

$$P_{\frac{-M}{2}}(i)=.....=P_{-1}(i)=0,\quad i\in S, \tag{20}$$

and

$$\sum_{j\in S}\mu_{if_ij}P_k(j)=0\quad for\quad k=0,1,....,M-1. \tag{21}$$

Case 2: M is odd.
If $n=-M$, then

$$\sum_{j\in S}\mu_{if_ij}P_0(j)=0,\quad i\in S. \tag{22}$$

If $n=-M+1<0$, then $\sum_{j\in S}\mu_{if_ij}P_1(j)=N(i)\left[P_{\frac{-M+1}{2}}(i)\right]^2$, $i\in S$. Therefore $P_{\frac{-M+1}{2}}(i)=0$ and $\sum_{j\in S}\mu_{if_ij}P_1(j)=0$, $i\in S$. By using the same procedure as in the case where M is even, we obtain analogous results.
If $n=0$, we get:

$$(\beta-2A(i))P_0(i)=Q(i)-N(i)[P_0(i)]^2+\sum_{j\in S}\lambda_{if_ij}P_0(j)+\sum_{j\in S}\mu_{if_ij}P_M(j),\quad i\in S. \tag{23}$$

>From (21) or (22), we have that:

$$\sum_{j\in S}\mu_{if_ij}P_0(j)=0,\quad i\in S. \tag{24}$$

Using the equation (24) and remark 2, we obtain

$$\begin{cases} \bar{P}_0(\bar{\xi}) = P_0(i), & i \in \bar{\xi}, \quad \forall \bar{\xi} \in \bar{S}. \\ P_0(i) = \sum_{\bar{\xi} \in \bar{S}} q_{\bar{\xi}}(i) \bar{P}_0(\bar{\xi}), & i \in T. \end{cases} \tag{25}$$

Substituting the first expression of (25) in (23) to get:

$$(\beta - 2A(i))\bar{P}_0(\bar{\xi}) = Q(i) - N(i)[\bar{P}_0(\bar{\xi})]^2 + \sum_{j \in S} \mu_{if_ij} P_M(j) + \sum_{j \in S} \lambda_{if_ij} P_0(j), \quad i \in \bar{\xi}. \tag{26}$$

Multiplying (26) by $m_{\bar{\xi}}(i)$, $i \in \bar{\xi}$ and summing up over $\bar{\xi}$:

$$\begin{aligned} \sum_{i \in \bar{\xi}} \left[m_{\bar{\xi}}(i)(\beta - 2A(i)) \right] \bar{P}_0(\bar{\xi}) &= \sum_{i \in \bar{S}} Q(i) m_{\bar{\xi}}(i) - \sum_{i \in \bar{\xi}} N(i) m_{\bar{\xi}}(i) [\bar{P}_0(\bar{\xi})]^2 \\ &+ \sum_{j \in S} \left(\underbrace{\sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) \mu_{if_ij}}_{=0} \right) P_M(j) + \sum_{\bar{\xi}' \in \bar{S}} \left(\sum_{\substack{i \in \bar{\xi} \\ j \in \bar{\xi}'}} m_{\bar{\xi}}(i) \lambda_{if_ij} \right) \bar{P}_0(\bar{\xi}') \\ &+ \sum_{j \in T} \sum_{i \in \bar{\xi}} \sum_{\bar{\xi}' \in \bar{S}} m_{\bar{\xi}}(i) \lambda_{if_ij} q_{\bar{\xi}'}(j) \bar{P}_0(\bar{\xi}'). \end{aligned} \tag{27}$$

We introduce the following notations: $\bar{A}(\bar{\xi}) = \sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) A(i)$, $\bar{Q}(\bar{\xi}) = \sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) Q(i)$, $\bar{B}(\bar{\xi}) = (\bar{N}(\bar{\xi}))^{\frac{1}{2}} = (\sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) N(i))^{\frac{1}{2}}$ and $\lambda_{\bar{\xi}\bar{\xi}'} = \sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i) \left(\sum_{j \in \bar{\xi}'} \lambda_{if_ij} + \sum_{j \in T} q_{\bar{\xi}'}(j) \lambda_{if_ij} \right)$. Thus, the equation (27) becomes:

$$(\beta - 2\bar{A}(\bar{\xi}))\bar{P}_0(\bar{\xi}) = \bar{Q}(\bar{\xi}) - \bar{N}(\bar{\xi})[\bar{P}_0(\bar{\xi})]^2 + \sum_{\bar{\xi}' \in \bar{S}} \lambda_{\bar{\xi}\bar{\xi}'} \bar{P}_0(\bar{\xi}'), \quad \bar{\xi} \in \bar{S}. \tag{28}$$

Note that $\bar{\theta}(t)$ defines a aggregated Markov chain defined on the state space $\bar{S}$ with the infinitesimal generator matrix $\bar{\Gamma} = (\lambda_{\bar{\xi}\bar{\xi}'})_{m \times m}$

Condition 2

The pair $(\bar{A}(\bar{\theta}), \bar{B}(\bar{\theta}))$ is stochastically stabilizable and $(\bar{A}(\bar{\xi}), \bar{Q}(\bar{\xi}))$ is observable for each $\bar{\xi} \in \bar{S}$.

Under condition 2, $\bar{Q}(\bar{\xi}) > 0$, the equation (28) admits a unique positive solution (see Altman et al.[1]), hence P_0 is uniquely determined by (23) and (24). Furthermore, the following jump linear system:

$$\dot{x} = D(\bar{\theta})x, \tag{29}$$

where $D(\bar{\xi}) = (\bar{A}(\bar{\xi}) - \bar{N}(\bar{\xi})\bar{P}_0(\bar{\xi}))$, is mean-square stable (Costa et al[4]).
If $n = 1$, then

$$(\beta - 2A(i))P_1(i) = -2N(i)[P_0(i)P_1(i)] + \sum_{j\in S} \lambda_{if_ij} P_1(j) + \sum_{j\in S} \mu_{if_ij} P_{M+1}(j), \quad i \in S. \quad (30)$$

>From (21), we have:

$$\sum_{j\in S} \mu_{if_ij} P_1(j) = 0. \quad (31)$$

By using the same procedure from (30) and (31), we obtain:

$$[\beta - 2D(\bar{\xi})]\bar{P}_1(\bar{\xi}) = \sum_{\bar{\xi}'\in\bar{S}} \lambda_{\bar{\xi}\bar{\xi}'} \bar{P}_1(\bar{\xi}'), \quad \bar{\xi} \in \bar{S} \quad (32)$$

Since the model (29) is mean-square stable, one applies proposition 2 [4] to derive that equation (32) has a unique solution. Since 0 is solution to (32), then 0 is the unique solution to (32). It follows that the matirice $(\beta I - \tilde{A})$ is regular, where $\tilde{A}_{\bar{\xi}\bar{\xi}} = 2\bar{A}(\bar{\xi}) - \bar{N}(\bar{\xi})\bar{P}_0(\bar{\xi}) + \lambda_{\bar{\xi}\bar{\xi}}$ and $\tilde{A}_{\bar{\xi}\bar{\xi}'} = \lambda_{\bar{\xi}\bar{\xi}'}$ and (30) yields: $\sum_{j\in S} \mu_{if_ij} P_{M+1}(j) = 0, \ \ i \in S$.
If $0 < n < M$, then

$$(\beta - 2A(i))P_n(i) = -2N(i)P_0(i)P_n(i) + \sum_{j\in S} \mu_{if_ij} P_{n+M}(j) + \sum_{j\in S} \lambda_{if_ij} P_n(j).$$

>From (21), it follows that $\sum_{j\in S} \mu_{if_ij} P_n(j) = 0$. Then we can derive $P_1(i) = .. = P_{M-1}(i) = 0$, and

$$\sum_{j\in S} \mu_{if_ij} P_{n+M}(j) = 0 \qquad n = 0, 1, ..., M-1. \quad (33)$$

If $n = M$, then

$$(\beta - 2A(i))P_M(i) = -2N(i)P_0(i)P_M(i) + \sum_{j\in S} \mu_{if_ij} P_{2M}(j) + \sum_{j\in S} \lambda_{if_ij} P_M(j). \quad (34)$$

>From (23), we get

$$\sum_{j\in S} \mu_{if_ij} P_M(j) = \alpha(i), \qquad i \in S, \quad (35)$$

where $\alpha(i) = (\beta - 2A(i))P_0(i) + 2N(i)[P_0(i]^2 - Q(i) - \sum_{j\in S} \lambda_{if_ij} P_0(j)$.
If $M < n < 2M$, then $(\beta - 2A(i))P_n(i) = -2N(i)P_0(i)P_n(i) + \sum_{j\in S} \mu_{if_ij} P_{n+M}(j) + \sum_{j\in S} \lambda_{if_ij} P_n(j)$. From(33), we have that $\sum_{j\in S} \mu_{if_ij} P_n(j) = 0$, and therefore $P_n(j) =$

$0, \ \forall j \in S$ *and* $n = M+1, .., 2M-1$.
Similarly, we have that:

$$\begin{aligned} P_n(i) &= 0, \quad n < 0 \\ P_o(i) &> 0, \quad i \in S \\ P_n(i) &= 0, \quad n \neq kM, \quad k = 1, 2, ... \end{aligned}$$

Now, we can state the following result.

Proposition 4.1 *Let condition 1 and 2 holds, with $B(i) \neq 0$ and $Q(i) > 0 \ \forall i \in S$. Then the coupled algebraic Riccati equation (12) admits a unique positive solution, which admits Taylor serie.*

The next theorem is about the computation of the terms in the expansion of P_ϵ and moreover the proof of the theorem is constructive in the sense that it provide an alrorithm for calculation.

Notation

Let $Ker(C)$ denotes the Kernel i.e $Ker(C) = \{y \in \mathbb{R}^s / Cy = 0\}$ and $Im(C)$ the range of the operator $C(f)$ i.e $Im(C) = \{y \in \mathbb{R}^s / \exists x \in \mathbb{R}^s, y = Cx\}$

Theorem 4.2 *P_ϵ has the expansion $\sum_{n=0}^{+\infty} P_n(i)\epsilon^n$, with $P_n = \tilde{P}_n + \bar{P}_n$ where $\tilde{P}_n \in Im(C)$ and $\bar{P}_n \in Ker(C)$. The sequence $(\tilde{P}_n, \bar{P}_n)$ is uniquely detemined by:*

$*$ *P_0 is the unique solution of*

$$(\beta - 2\bar{A}(\bar{\xi}))P_0(\bar{\xi}) = \bar{Q}(\bar{\xi}) - \bar{N}(\bar{\xi})[P_0(\bar{\xi})]^2 + \sum_{\bar{\xi} \in \bar{S}} \lambda_{\bar{\xi}\bar{\xi}'} P_0(\bar{\xi}'). \tag{36}$$

$*$ *$\tilde{P}_n$, $n > 0$ the unique solution of the linear system*

$$C\tilde{P}_n = \alpha_{n-1}, \tag{37}$$

where

$$\alpha_{n-1}(i) = (\beta - 2A(i))P_{n-1}(i) + N(i)P_0(i)P_{n-1}(i) - \sum_{j \in S} \lambda_{if_ij} P_{n-1}(j) - h_{n-1}(i).$$

with $h_n(i) = N(i) \sum_{k=1}^{n-1} P_k(i) P_{n-1-k}(i)$

$*$ *$\bar{P}_n$ $n > 0$ is a solution to:*

$$\left\{ \begin{array}{l} \bar{P}_n(i) = \bar{P}_n(\bar{\xi}), \quad where, \quad \bar{P} = (\beta I - \tilde{A})^{-1} T_n, \qquad i \in \bar{\xi}, \quad \bar{\xi} \in \bar{S} \\ \bar{P}_n(i) = \sum_{\bar{\xi} \in \bar{S}} q_{\bar{\xi}'}(i) \bar{P}_n(\bar{\xi}'), \qquad i \in T, \end{array} \right. \tag{38}$$

where $T_n(\bar{\xi}) = \sum_{i \in \bar{\xi}} m_{\bar{\xi}}(i)\{[\beta - 2A(i) - N(i)P_0(i)]\tilde{P}_n(i) + \sum_{j \in S} \tilde{P}_n(j) - h_{n-1}(i)\}$

Proof Let $P_\epsilon(i) = \sum_{n=0}^{+\infty} P_n(i)\epsilon^n$ be the expansion of P_ϵ, where P_n satisfies the equations

$$\begin{cases} \sum_{j\in S} \mu_{if_ij} P_0(j) = 0 \\ (\beta - 2A(i))P_0(i) = Q(i) - N(i)[P_0(i)]^2 + \sum_{j\in S} \lambda_{if_ij} P_0(j) + \sum_{if_ij} \mu_{if_ij} P_1(i) \\ (\beta - 2A(i))P_n(i) = -N(i)P_0(i)P_n(i) + \sum_{j\in S} \lambda_{if_ij} P_n(j) + \sum_{j\in S} \mu_{if_ij} P_{n+1}(j) \\ \qquad\qquad -N(i)\sum_{k=1}^{n-1} P_k(i)P_{n-k}(i), \qquad n > 0 \end{cases} \tag{39}$$

>From the preceding results, P_0 is the unique solution of the equation (36).
$C(f)$ admits the eigenvalue 0 because $\sum_{j\in S} \mu_{if_ij} = 0$. This eigenvalue is semi simple (the eigenspace associated to the eigenvalue 1 admits a base of eigenvectors) this can be shown easily by noting that $q_{\bar\xi}$, $\bar\xi \in \bar S$ defined below, form a base of m-dimensional suspace $Ker(C)$. From this property we have the decomposition: $\mathbb{R}^s = Ker(C) \oplus Im(C)$, . Therefore, $P_n = \bar P_n + \tilde P_n$, where $\bar P_n \in Ker(C)$ and $\tilde P_n \in Im(C)$. Substituting this structur in (37), we get

$$\begin{aligned} (\beta - 2A(i) + 2N(i)P_0(i))\bar P_n(i) &= \sum_{j\in S} \lambda_{if_ij}\bar P_n(j) - (\beta - 2A(i) + 2N(i)P_0(i))\tilde P_n(i) - f_{n-1}(i) \\ &+ \sum_{j\in S} \lambda_{if_ij}\tilde P_n(j) + \sum_{j\in S} \mu_{if_ij} P_{n+1}(j). \end{aligned} \tag{40}$$

We shall show that (37) (respectively (40)) admits a unique solution. We have $\alpha_{n-1} \in Im(C)$, since $\tilde P_n \in Im(C)$, then (37) admits a unique solution.
On the other hand, $\bar P_n \in Ker(C)$, then

$$\begin{cases} \bar P_n(i) = \bar P_n(\bar\xi) \qquad i \in S \\ \bar P_n(i) = \sum_{\bar\xi' \in \bar S} q_{\bar\xi'}(i)\bar P_n(\bar\xi'). \end{cases} \tag{41}$$

Substituting this structur (41) in (40), multiplying each equation by $m_{\bar\xi}$ such that $i \in \bar\xi$ and summing up over $\bar\xi$, we obtain:

$$[\beta - 2\bar A(\bar\xi) + 2\bar N(\bar\xi)P_0(\bar\xi)]\bar P_n(\bar\xi) - \sum_{\bar\xi' \in \bar S} \lambda_{\bar\xi\bar\xi'} \bar P_n(\bar\xi') = T_n(\bar\xi). \tag{42}$$

Recall the definition of $\tilde A$ below (32). As we showed there, $(\beta I - \tilde A)$ is nonsingular, hence the unique the solution of (42) is $\bar P_n = (\beta I - \tilde A)^{-1} T_n$. ∎

Iterative algorithm to compute P_0:

1- Set $P_0^0(i) = 0$, $i \in S$.

2-Compute $P_0^l(i), i \in S, l \geq 1$, iteratively by solving the following set of quadratic equations:

$$(\beta + \lambda_{\bar\xi}^f - 2\bar A(\bar\xi))P_0^{l+1}(\bar\xi) = -\bar N(\bar\xi)[P_0^{l+1}(\bar\xi)]^2 + \tilde Q(f, P_0^l, \bar\xi), \tag{43}$$

where $\tilde{Q}(f, P_0^l, \bar{\xi}) = \bar{Q}(\bar{\xi}) + \sum_{\bar{\xi} \in \bar{S}} \lambda^f_{\bar{\xi}\bar{\xi}'} P_0^l(\bar{\xi}) + \lambda^f_{\bar{\xi}} P_0^l(\bar{\xi})$ and $\lambda_{\bar{\xi}} =\mid \lambda_{\bar{\xi}\bar{\xi}} \mid$
>From the definition of $\tilde{Q}(f, P_0^l, \bar{\xi})$, if $P_0^{l+1}(\bar{\xi}) \geq P_0^l(\bar{\xi})$ then

$$\tilde{Q}(f, P_0^{l+1}, \bar{\xi}) \geq \tilde{Q}(f, P_0^l, \bar{\xi}). \tag{44}$$

Note that (43) is a standard Riccati equation that corresponds to a system that always remains in the state $\bar{\xi}$, and where the weighting on the quadratic cost for the state is $\tilde{Q}(f, P_0^l, \bar{\xi})$. Therefore, its solution $P_0^{l+1}(i)$ will be increasing. It follows that,

$$if \quad \tilde{Q}(f, P_0^{l+1}, \bar{\xi}) \geq \tilde{Q}(f, P_0^l, \bar{\xi}) \quad then \quad P_0^{l+2}(\bar{\xi}) \geq P_0^{l+1}(\bar{\xi}). \tag{45}$$

In view of the fact that $P_1(\bar{\xi}) \geq P_0(\bar{\xi}) = 0$ (44)-(45) establish by induction the desired result that the sequences P_0^l is nondecreasing for each $i \in S$ and therefore have respective limits, with the former satisfying(36).

Remark 2 *The Taylors expansion of P_ϵ can be computed recursively in n:*
P_0 is computed by the iterative algorithm,
$\tilde{P}_1$ is solution of $(37)_1$ (this is a linear system)
$\bar{P}$ is computed by $(38)_1$,
$\tilde{P}_2$ is solution of $(37)_2$,
$\bar{P}_2$ is computed by $(38)_2$,
and so on .....

5 Optimal and uniformly good strategies

We now describe a policy iteration algorithm that will lead to a policy which is ϵ^k-optimal for all ϵ sufficiently small, where the optimization in $\mathbb{R}$ can be done by using lexicographical optimization.

Definition 5.1 *A sequence (a_n) $n = 0, ..., \infty$ is said to be lexicographical positive if there exists K such that $a_n = 0$ if $n < K$ and $a_K > 0$. The sequence (a_n) is said to be lexicographical greater than the sequence (b_n) if the sequence $(a_n - b_n)$ is lexicographical positive.*

Notation
Let $\overrightarrow{Opt}$ denote the lexicographical optimization, $V_\epsilon(i) = \sum_{n=-1}^{\infty} V_n(i)\epsilon^n$,
$\overrightarrow{V_\epsilon(i)} = (V_{-1}(i), V_0(i), ...)$, $\overrightarrow{V_\epsilon(i)}^l = (V_{-1}(i), ..., V_l(i))$ and $V_\epsilon^l(i) = \sum_{n=-1}^{l} V_n(i)\epsilon^n$.

Lemma 5.1
1- $\overrightarrow{P_\epsilon^f(i)} \succeq 0 \quad \forall i \in S$ is equivalent to $\exists \rho : \forall \epsilon \leq \rho \quad , P_\epsilon^f(i) \geq 0, \forall i \in S$
2- $\overrightarrow{P_\epsilon^f(i)}^l \succeq 0, \quad \forall i \in S$ is equivalent to $\exists \rho : \forall \epsilon \leq \rho \quad , P_\epsilon^{f,l}(i) \geq 0 \quad \forall i \in S.$

The proof is trivial ∎

Let g be the strategy defined by:

$$g_i = \arg_{a \in U_2(i)} opt \sum_{j \in S} \lambda^{\epsilon}_{iaj} P^f_{\epsilon}(j), \quad i \in S. \tag{46}$$

Lemma 5.2

If g is the strategy defined by (46), and P^g_{ϵ} is the solution to:

$$(\beta - 2A(i))P^g_{\epsilon}(i) = -N(i)[P^g_{\epsilon}(i)]^2 + Q(i) + \sum_{j \in S} \lambda^{\epsilon}_{ig_i j} P^g_{\epsilon}(j), \quad i \in S,$$

then $P^g_{\epsilon}(i) \leq P^f_{\epsilon}(i)$ for team problem (resp $P^g_{\epsilon}(i) \geq P^f_{\epsilon}(i)$ for game problem).

Proof

This result is given by the outer-loop algorithm in Altman et al[1]. ∎

Theorem 5.1 *Let g be the strategy defined by:*

$$g_i = \arg_{a \in U_2(i)} \overrightarrow{opt} \overrightarrow{\sum_{j \in S} \lambda^{\epsilon}_{iaj} P^f_{\epsilon}(j)}, \quad i \in S. \tag{47}$$

If $\overrightarrow{P^g_{\epsilon}}$ is the solution to:

$$(\beta - 2A(i))\overrightarrow{P^g_{\epsilon}(i)} = -N(i)\overrightarrow{[P^g_{\epsilon}(i)]^2} + \overrightarrow{Q(i)} + \overrightarrow{\sum_{j \in S} \lambda^{\epsilon}_{ig_i j} P^g_{\epsilon}(j)}, \quad i \in S, \tag{48}$$

then $\overrightarrow{P^g_{\epsilon}(i)} \preceq \overrightarrow{P^f_{\epsilon}(i)}$ for team problem (resp $\overrightarrow{P^g_{\epsilon}(i)} \geq \overrightarrow{P^f_{\epsilon}(i)}$ for game problem).

Proof

To see fron the team problem ("opt"="min") that

$$\overrightarrow{\sum_{j \in S} \lambda^{\epsilon}_{if_i j} P^f_{\epsilon}(j)} \succeq \overrightarrow{\sum_{j \in S} \lambda^{\epsilon}_{ig_i j} P^f_{\epsilon}(j)}$$

then $\exists \rho$, s.t $\forall \epsilon \leq \rho$ $\sum_{j \in S} \lambda^{\epsilon}_{if_i j} P^f_{\epsilon}(j) \geq \sum_{j \in S} \lambda^{\epsilon}_{ig_i j} P^f_{\epsilon}(j)$. From lemme 5.2 we have that $\forall \epsilon \leq \rho$ $P^f_{\epsilon}(i) \geq P^g_{\epsilon}(i)$ $\forall i \in S$ and is equivalent to (lemme 5.1) $\overrightarrow{P^g_{\epsilon}(i)} \preceq \overrightarrow{P^f_{\epsilon}(i)}$ $\forall i \in S$. ∎

Theorem 5.2 *Let $(\overrightarrow{P^{f^k}})_{k \in \mathbb{N}}$ be the sequence defined by :*

$$f^k_i = \arg_{a \in U_2(i)} \overrightarrow{opt} \overrightarrow{\sum_{j \in S} \lambda^{\epsilon}_{if^{k-1}_i j} P^{f^{k-1}}_{\epsilon}(j)},$$

where $P_\epsilon^{f^k}$ is the solution to (48) ($g = f^k$), then there exists f^* such that $\lim_{K\to+\infty} \overrightarrow{P_\epsilon^{f^K}} = \overrightarrow{P_\epsilon^{f^*}}$. Furthermore there exists n_0 such that $\overrightarrow{P_\epsilon^{f^k}} = \overrightarrow{P_\epsilon^{f^*}}$ for all $k \geq n_0$ and f^* is the optimal stationary policy for player 2.

Proof

>From Theorem 5.1, the sequence $\overrightarrow{P_\epsilon^{f^k}}$ is monotonous and lexicographical positive, hence it converge. The other part of theorem follows from the fact that the action space is finite.
We now provide a precise statement of this result in the following theorem:

Theorem 5.3

Let condition 1 and 2 holds, with $B(i) \neq 0$ and $Q(i) > 0\ \forall \in S$. Then:
(i)The coupled set of equation (11) admits a unique positive solution, P_ϵ^*, which admits Taylor serie.
(ii) There exists a unique stationary optimal policy for Player 1, given by

$$\gamma_\epsilon^{1*}(x,i) = -[B(i)P_\epsilon^*(i)/R(i)]x(i), i \in S.$$

(ii) There is an optimal (minimising in the team case, and maximising in the game case) stationary policy for Player 2 (f^*), independently of variation of ϵ in the vicinity of zero.

Theorem 5.4

1- Let g be the strategy such that $l \geq 0$

$$\overrightarrow{\sum_{j\in S} \lambda_{if_ij}^\epsilon P_\epsilon^f(j)}^l \succeq \overrightarrow{\sum_{j\in S} \lambda_{ig_ij}^\epsilon P_\epsilon^f(j)}^l \tag{49}$$

then $\overrightarrow{P_\epsilon^f(i)}^l \succeq \overrightarrow{P_\epsilon^g(i)}^l$ and $\exists\rho$ s.t $\forall\epsilon \leq \rho$, $P_\epsilon^{f,l}(i) \geq P_\epsilon^{g,l}(i)$.
2- $\overrightarrow{P_\epsilon^*(i)}^l = (0, P_0^*(i), .., P_l^*(i))$ where $P_\epsilon^*(i)$ is solution to (11), is uniquely determined by the equation

$$(\beta - 2A(i))\overrightarrow{P_\epsilon(i)}^l = \overrightarrow{Q(i)}^l - N(i)\overrightarrow{[P_\epsilon(i)]^2}^l + \overrightarrow{opt}_{a\in U2(i)} \overrightarrow{\sum_{j\in S} \lambda_{iaj}^\epsilon P_\epsilon(j)}^l. \tag{50}$$

Proof

1- From (49), $\exists\rho\ \exists\rho'$ s.t $\forall\epsilon \leq \rho$, we have $\sum_{j\in S} \lambda_{if_ij}^\epsilon P_\epsilon^f(j) + \rho'\epsilon^{l+1} \geq \sum_{j\in S} \lambda_{ig_ij}^\epsilon P_\epsilon^f(j)$.
Consider the iteration (inner-loop):

$$\begin{aligned}(\beta + \lambda_i^\epsilon - 2A(i))P_\epsilon^{(k+1)}(i) &= -N(i)[P_\epsilon^{(k+1)}(i)]^2 + Q(i) - \rho'\epsilon^{l+1} + \sum_{j\in S} \lambda_{ig_ij}^\epsilon P_\epsilon^{(k)}(j) \\ &+ \lambda_i^\epsilon P_\epsilon^{(k)}(i), k = 0, 1, ..; \quad P_\epsilon^{(0)}(i) = P_\epsilon^f(i),\end{aligned}$$

where $\lambda_i^\epsilon = \max_{a\in U_2(i)} \mid \lambda_{iaj}^\epsilon \mid = \lambda_i + \frac{1}{\epsilon}\mu_i$.
Note that

$$Q(i) - \rho'\epsilon^{l+1} + \sum_{j\in S} \lambda_{ig_ij} P_\epsilon^f(j) + \bar{\lambda}_i P_\epsilon^f(i) \leq Q(i) + \sum_{j\in S} \lambda_{if_ij} P_\epsilon^f(j) + \bar{\lambda}_i P_\epsilon^f(i)$$

Then we have $P_\epsilon^{(1)}(i) \leq P_\epsilon^f(i)$.
By using same procedure in Altman et al[1], we can conclude that the inner-loop iteration generates monotonicaly nondecreasing positive sequence, wich has to converge to the solution of :

$$(\beta - 2A(i))P_\epsilon'^g(i) = -N(i)[P'^g_\epsilon(i)]^2 + Q(i) - \rho'\epsilon^{l+1} + \sum_{j\in S} \lambda_{ig_ij}^\epsilon P'^g_\epsilon(j), \quad i \in S,$$

By using same procedure for as in the resolution for equation (12), we find that $\overrightarrow{P_\epsilon'^g(i)}^l = \overrightarrow{P_\epsilon^g(i)}^l$, then $\exists \bar{\rho}, \bar{\rho'}$ s.t $\forall \epsilon \leq \bar{\rho}$, $P_\epsilon^g(i) \leq P_\epsilon^f(i) + \bar{\rho'}\epsilon^{l+1}$, this equivalent to $\overrightarrow{P_\epsilon^g(i)}^l \preceq \overrightarrow{P_\epsilon^f(i)}^l$.
2- By using first party of this theorem, we can construct a monotonicaly positive sequence $\overrightarrow{P_\epsilon^{f^K}}^l$ which has converge ie $\lim_{K\to+\infty} \overrightarrow{P_\epsilon^{f^K}}^l = \overrightarrow{P_\epsilon^{f^*}}^l$ and $\overrightarrow{P_\epsilon^{f^*}}^l$ solution to (50). Therefore $\overrightarrow{P_\epsilon^{f^*}(i)}^l \preceq \overrightarrow{P_\epsilon^f(i)}^l$ $\forall i \in S$ $\forall f \in F$ then $\exists \rho$ s.t $\forall \epsilon \leq \rho$ $P_\epsilon^{f^*,l}(i) \leq P_\epsilon^{f,l}(i)$ $\forall i \in S$ $\forall f \in F$, hence we have $P_\epsilon^{f^*,l}(i) = P_\epsilon^{*,l}(i)$. ∎

Remark 3
$P^{f^*,l}$ (resp f^*) is an ϵ^l-optimal solution of (11) (resp ϵ^l-optimal strategy for Player 2).

Policy Iteration Algorithm
Now, we present an algorithm to compute the ϵ^l-optimal solution of (11) and ϵ^{l+1}-optimal stationary policy for player 2.
Step 1:
Choose some strategy f_1 in F, $k := 1$.
Step 2:
Compute $l+1$ first terms of Taylor serie $P_\epsilon^{f_k}$ solution to (12) (Theorem 4.2).
Step 3 :
Let $n := 1$ and U_2^n be defined by :

$$\begin{cases} U_2^{-1}(i) = \arg opt_{a\in U_2(i)} \sum_{j\in S} \mu_{iaj} P_0^{f_k}(j) \\ U_2^n(i) = \arg opt_{a\in U_2^{n-1}} \sum_{j\in S} (\lambda_{iaj} P_n^{f_k}(j) + \mu_{iaj} P_{n+1}^{f_k}(j)) \end{cases}$$

Let $F_n^k = \prod_{i \in S} U_2^n$. If $|F_n^k| = 1$ set $F_l^k = F_n^k$ and go to step 4.
If $|F_n^k| \neq 1$ and $n \leq l$ then $n \to n+1$ and go to step 3. Otherwise, go to step 4.
Step 4:
If F_l^k is identical to F_l^{k-1}, stop: F_l^k is the set of all ϵ^l-optimal strategies. Otherwise, choose some strategy f_{k+1} in F_l^k and go to step 2.

6 Conclusions and further research

We studied in this paper a linear quadratic control problem with Markov jumps, in which not only the original state is controlled but also the transitions of the Markov parameters. The values and optimal policies for both controllers are given using the solution of a corresponding set of CARE (coupled algebraic Riccatti equations). We showed that this solution has a Taylor expansion in the perturbation and computed the different terms in this expansion. We showed that there exists a policy for the controller of the Markov jumps which is optimal for all small enough perturbations. Numerical methods were proposed to compute or apprximate this policy.
We restricted to the case in which the state is one dimensional, since the quadratic structure of the optimal value (and the linear structure of the optimal control) does not extend to the multi-dimensional case. We shall carry on future research to identify special cases in which the multidimensional problem has still a quadratic value. The singular perturbation problem will then be handled by extending our results for the noncontrolled case to the multidimensional setting.

References

[1] E. Altman, T. Basar, and Zigang Pan, "Piecewise-Deterministic Game with Hybrid Contols" Annals of Dynamic Games, Vol.5, J. Filar 1998.

[2] T. Basar, "Minimax contol of switching systems under sampling," System Control Letters, 25(5): 315-325, 1995.

[3] T. Basar and A. Haurie, Feedback equilibria in differential games with structral and modal uncertainties, vol. 1 of Advances in Large Scale Systems, pp. 163-201. Connecticut: JAI Press Inc., May 1984. (J. B. Cruz, Jr., ed.).

[4] O. L. V. Costa, J. B. Do Val, J. C. Geromel, "Contonuous-time state-feedback H_2-control of Markovian jump linear system via convex analysis," Automatica 35, pp. 259-268, 1999.

[5] A. Czornick "On discrete-time linear quadratic control" System & cintrol Letters, pp. 101-107, 1999

[6] Y. Ji and H. J. Chizeck, "Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control," IEEE Transaction on Automatic Control, vol. AC-35, pp. 777-778, July 1990.

[7] Z. Pan and T. basar, "H^∞-Contol of Markovian jump systems and solution to associated piecewise-deterministic differntial game," in G.J. Olsder, editor, Annals of Dynamic Games, vol. 3, pp. 61-94, Birkhauser, 1995.

[8] X. Feng, K.A. Loparo, and H. J Chizeck, " Stochastic stability properties of jump linear systems" IEEE Trans. Automat. Control, vol. 37 NO. 1. January 1992.

[9] N. N. Krassovskii and E. A. Lidskii, "Analytical design of controllers in system with random attributes I, II, III," Automation Remote Contr., vol. 22, pp. 1021-1025, 1141-1146, 1289-1294, 1961 .

[10] E. A. Lidskii, "Optimal control of systems with random properties," Appl. Math. Mech., vol. 27, pp. 33-45, 1961.

[11] J. J Floretin, "Optimal contol of continuous-time Markov, stochastic system," J. Electron. Control, vol. 10, 1961.

[12] H. J Kushner, "On the stochastic maximum principle: Fixed time of control," J. Math. Appl., vol. 11, pp. 78-92, 1965.

[13] R. Rishel, "Dynamic programming and minimum principle for systems with jump Markov diturbances," SIAM Jounal on control and Optimization, vol. 13, pp. 338-371, Ferbuary 1975.

[14] D. D. Swoder, "Feedback control of a class of linear system with jump parameters," IEEE Trans. Automat. Control, vol. 14, pp. 9-14, 1969.

[15] W. M. Wonham, "Random differential equation in control theory," in Probabilistic Methods in Applied Mathematic (A. T. Bharucha-Ried, ed.), pp. 191-199, New York, NY: Academic Press, 1970.

[16] M. Mariton, "On controllability of linear systems with stochastic jump parameter," IEEE Trans. Automat. Control, vol. AC-31, pp. 680-683, 1986.

[17] J. B. Moore, X. Y Zhou, and A. E.B. Lim " Discrete time LQG controls with control dependent noise" System & Control Letters 36 (1999) 199-206, 1999

[18] P. Shi, R. K. Agarwal, and E. Boukas, "Robust stabilization of discrete time-delay Markovian jump systems" 14th World Congress of IFAC, 1999.

[19] C. E. de Souza and M. Fragoso, "H^∞ control of linear systems with Markovian jumping parameters," Control Theory and Technology, vol. 9, no. 2, pp. 457-466, 1993.

[20] W. W. Szczechila, S. A. Connell, J.A. Filar, And O. J. Vrieze, "On the puiseux series expansion of the limit discount equation of stochastic games", Siam J, control optimal, vol. 35, No. 3, pp. 860-875, May 1997.

[21] Y. Yan, and J. Lam, " Stochastic stabilizable and H^∞ control for Markovian jump time-delay systems" Proceding of the 14th IFAC, 1999.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399