The averaging principle for perturbations of continuous time control problems with fast controlled jump parameters

Rachid El Azouzi
Département de Math & Info
Faculté des sciences
B.P 1014 Rabat
Maroc
relazouz@sophia.inria.fr

Eitan Altman
INRIA B.P 93
2004 Route des Lucioles
06902 Sophia Antipolis
France
altman@sophia.inria.fr

Vladimir Gaitsgory
School of Mathematics
University of South Australia
The Levels Campus
SA 5095, Australia
V.Gaitsgory@unisa.edu.au

Abstract

We consider a class of singularly perturbed zero-sum 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 quantity. For player 2, we consider two possible scenarios: one in which it wishes to minimize the same quantity (team framework), and one in which it wishes to maximize it (zero sum game). The transition rates of the Markov process are fast, of the order of $1/\epsilon$. To solve the above problem, we use the dynamic programming approach. In particular, we study the asymptotic properties of the underlying system for sufficiently small ϵ. The viscosity solution method is employed to verify the convergence of the value function, which allows us to obtain the convergence in a general setting and helps us to characterize the structure of the limit system. We apply this to the special case of linear quadratic games with jump parameters, which allows us to obtain an explicit solution for the limiting problem.

1 Introduction

We consider in this paper a system with a continuous time control. The parameters of the systems may change according to a jump Markov process whose transitions are further controlled by a second controller. Our purpose is to study the behavior of the system in the case that the controlled Markov jump parameters change, in some sense, much faster than the continuous time part of the state and the continuous control.

The averaging approach has already been used successfully in deterministic controlled systems with two time scales, see e.g. [10]; the basic idea is to solve separately the control of the fast part (assuming a "frozen" state for the slow part), and then to solve the other control component. For the case of hybrid systems with both a continuous part and a jump Markov chain, the averaging approach has been successfully used in the case that either only the continuous part is controlled [16] or in the case that the Markov jump process is controlled, see [18, 15] and references therein.

We consider both cooperation (a team problem) between the controllers of the continuous part (called agent 1 or player 1) and that of the Markov jump process (called agent 2 or player 2) as well as antagonistic objectives (a zero-sum game) between these players. The objective of player 1 in both cases is to minimize an expected nonlinear cost of the state trajectory over a finite horizon problem. Using an averaging procedure, we show that the above minimization problem can be approximated by the solution of some deterministic optimal control problem. The value of the latter is obtained as the viscosity solution of some dynamic programming equation.

Our result can be regarded, as an extension of results on uncontrolled motions establishing that the solution of the original stochastic system is approximated by the solution of some deterministic system obtained via averaging over the fast random dynamics [12] to the case when this random dynamics is defined by the controlled Markov chain. Note that the problems dealt with here are closely connected with the ones of stochastic singularly perturbed control, see, e.g. [3, 4, 6, 7, 13, 17], which have been extensively investigated by singular perturbations or averaging techniques to Hamilton-Jacobi-Bellman equation for the problems in continuous time [3, 4], or to the dynamic programming equation for the singularly perturbed MDP [5, 7, 17]. This paper generalizes the approach in [15, 18] not only in considering another controller for the continuous part, but also in allowing for an unbounded control set and in dropping the local Lipschitz continuity conditions on the costs and dynamics. This allows us to handle linear quadratic control problems with jump parameters, in which the transitions of the jump process are controlled as well. Such models have been used in [1] for modeling and solving hybrid problems in telecommunications in which controller 1 represented a flow control and controller 2 the admission control. For such linear quadratic models, exact solutions are only known for the case that the continuous part of the state is one-dimensional [1] and the procedure we propose in this paper can be used to obtain asymptotically tight approximations through a simpler deterministic limiting control problem. Note that, although we consider very short expected time between the control transitions of player 2, the fact that the control set for player 1 is unbounded makes it possible for it to produce rapid changes in the continuous part of the state. This makes the application of the averaging principle harder in our framework. We make assumptions on the cost function that discourage player 1 to cause fast changes in the dynamics. The structure of this paper is as follows. We begin by introducing the general framework. Then in Section 3 we present the general results and general ideas of the proof by solving a simpler model. The general model is then solved in Section 4. Finally, we apply our results in Section 5 to the linear quadratic framework.

2 General Model

The continuous dynamics of the game under consideration evolve according to

$$\frac{dx}{dt} = g(x(t), u^1(t), \theta(t)), \qquad x(t_0) = x, \tag{1}$$

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

$$\Lambda = (\frac{1}{\epsilon}\nu_{iaj}), \qquad i, j \in \Theta, \quad a \in U_2(i).$$

The ν_{iaj}'s are real numbers such that for any $i \neq j$, and $a \in U_2(i)$, $\nu_{iaj} \geq 0$, and for all $a \in U_2(i)$ and $i \in \Theta$, $\nu_{iai} = -\sum_{j\neq i} \nu_{iaj}$. The scaler ϵ is small positive number. Fix some initial state i_0 of the controlled Markov chain Θ. 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

$$u^k(t) = \gamma^k(t, x_t, \theta_t), \qquad t \in [t_0, T). \tag{2}$$

Here, γ^k is taken to be piecewise continuous in its first argument, and piecewise Lipschitz continuous in its second argument.
Define $\chi = \mathbb{R}^p \times \Theta$ to be the global state space of the system and $U := U_1 \times U_2$ to be the class of multy-strategies $\gamma := (\gamma^1, \gamma^2)$. 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 $E^\gamma_{x_0,i_0}$ the expectation operator corresponding to $P^\gamma_{x_0,i_0}$. We denote by $(x(t), \theta(t)), u(t), t \in [0, T)$, 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 (expected) cost function:

$$J_\epsilon(t_0, x_0, i_0, \gamma) = E^\gamma_{x_0,i_0}\{\varphi(x(T)) + \int_{t_0}^T L(x(t), u^1(t), \theta(t))dt\}, \tag{3}$$

where L is the running cost function and φ is the terminal cost function. We denote the cost to go from any time state pair (t, x, i) by

$$J_\epsilon(t, x, i, \gamma) = E^\gamma_{x,i}\{\varphi(x(T)) + \int_t^T L(x(\sigma), u^1(\sigma), \theta(\sigma))d\sigma\}.$$

We consider two problems, in both the objective is to compute

$$V_\epsilon(t, x, i) = \overline{opt}_{\gamma\in U} J_\epsilon(t, x, i, \gamma), \tag{4}$$

where $\overline{opt}_{\gamma\in U}$ will stand for

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

in the first case, called the game problem, or for

$$\inf_{\gamma^1\in U_1} \inf_{\gamma^2\in U_2}$$

in the second case, called the team problem.

Associated problem for the fast dynamic: Next we introduce another control problem, which will be associated with the limit behavior of (1) and (4). We define the following continuous time MDP with two controllers, where controller 1 controls only the immediate cost and controller 2 controls only the transitions. The MDP is parameterized by some vectors x and λ both in $\mathbb{R}^p$.

- The state space is Θ, as the second component in the state space of our original control problem.
- The action spaces are again U_1 and U_2, respectively.
- The rate of the controlled transitions is given by $\{\nu_{iaj}\}$.
- The instantaneous cost is given by $L(x, u^1, \theta) + \lambda^T g(x, u^1, \theta)$.

We denote the stochastic process induced by a policy $\hat{\gamma}$ over the state Θ by $\{\hat{\theta}(t)\}$. Roughly speaking, the problem will be related to the control of our original problem when "zooming" over the fast dynamics; the first component of the state is then frozen.

"Frozen" Problem: For given x, and λ and initial state $i \in \Theta$, define

$$\hat{Q}(x, \lambda, i) := \inf_{\hat{u}^1} \Big(L(x, \hat{u}^1, i) + \lambda^T g(x, \hat{u}^1, i) \Big).$$

Define now

$$H(x, i, \lambda, S) = \frac{1}{S}\int_0^S \hat{Q}(x, \lambda, \hat{\theta}(\sigma))d\sigma. \tag{5}$$

(The distribution of $H(x, i, \lambda, S)$ clearly depends on u^2). Define now

$$H(x, \lambda) := opt_{\hat{u}^2} \limsup_{S\to+\infty} E_i^{\hat{u}^2} H(x, i, \lambda, S) \tag{6}$$

Note that due to the ergodicity condition defined before (Assumption A1), the infinite horizon problem does not depend on the initial state i.

Limit Hamilton-Jacobi-Issacs equations for the stochastic hybrid control: Consider the Hamilton-Jacobi-Bellman (HJB, in short) (or Issacs) equation:

$$\frac{\partial v(t, x)}{\partial t} + H\left(x, \frac{\partial v(t, x)}{\partial x}\right) = 0 \quad (t, x) \in [0, T] \times \mathbb{R}^p, \tag{7}$$

with the boundary condition

$$v(T, x) = \varphi(x), \quad \forall x \in \mathbb{R}^p. \tag{8}$$

with Hamiltonian $H(x, \lambda)$ being equal to the limit of $H(x, \lambda, S)$ as S tends to $+\infty$.

In the following sections, it will be established under different sets of conditions that the limit $V_0(t, x)$ of the optimal value of our problem is given by the unique viscosity solution of Hamilton-Jacobi equations (7). We shall further give conditions for that limit to exist:

$$\lim_{\epsilon \to 0} V_\epsilon(t, x, i) = V_0(t, x), \quad \forall t, x, i. \tag{9}$$

3 Simplified model and convergence results

In order to present the main ideas for the derivation of the results, we begin by presenting a simplified model for which the exposition of some of the main results and their derivations are relatively simple. The simplicity is obtained at the cost of imposing strong assumptions that are difficult to verify. Then verifiable sufficient conditions for these assumptions, or for other weaker assumptions (that will be shown to guarantee the same results) will be presented in later sections.

We make the following assumption in this section.

Assumption A0:

- (i) U_1 is compact subset of $\mathbb{R}^r$,
- (ii) For any compact set D, there exists a compact set D_1 such that if the initial value $(x, i) \in D \times \Theta$ at time $t \geq 0$, then all the solutions of (1) obtained with different control functions do not leave $D_1 \times \Theta$ during $[t, T]$ where $T < \infty$ is the finite horizon that is considered.
- (iii) g is Lipschitz in $D_1 \times U_1$ and L is Lipschitz in D_1, with some Lipschitz constant C_2.
- (iv) For any sequence ϵ_n converging to 0, we can find a subsequence which we denote by $\epsilon(k)$, for which the following limit (that may depend on the subsequence) exists

  $$\lim_{k \to +\infty} V_{\epsilon(k)}(t, x, i) = V_0(t, x). \quad \forall t, x, i. \tag{10}$$

  We further assume that the convergence in (10) is uniform with respect to (t, x, i) from any compact subset of $[t_0, T] \times \mathbb{R}^p \times \Theta$, then there exists function $\hat{\mu}(\epsilon)$ tending to zero as ϵ tend to zero such that

  $$|V_{\epsilon(k)}(t, x, i) - V_0(t, x)| \leq \hat{\mu}(\epsilon(k)) \tag{11}$$
  $$, (t, x, i) \in [t_0, T] \times D_1 \times \Theta$$

Lemma 3.1 *Under A0, the (HJB) equation (7) has a unique continuous viscosity solution satisfying the boundary condition(8).*

Proof: Under A0, there exists some real constant C_1 such that

$$\begin{aligned}
&\hat{Q}(x, \lambda_1, \theta(\tau)) - \hat{Q}(x, \lambda_2, \theta(\tau)) \\
&\quad \leq (\lambda_1 - \lambda_2)^T g(x, \hat{u}^1(\tau), \hat{\theta}(\tau)) \\
&\quad \leq C_2 |\lambda_1 - \lambda_2| (1 + |x| + |\hat{u}^1(\tau)|) \\
&\quad \leq C_2 (C_1 + 1)(1 + |x|) |\lambda_1 - \lambda_2|
\end{aligned}$$

where $|\hat{u}^1(\tau)| \leq C_1 \; \forall \tau \in [0, +\infty)$ (since U_1 is compact). This implies that

$$H(x, \lambda_1) - H(x, \lambda_2) \leq C_2 (C_1 + 1)(1 + |x|) |\lambda_1 - \lambda_2|. \tag{12}$$

On the other hand, we have

$$\begin{aligned}
&\hat{Q}(x, \lambda, \theta(\tau)) - \hat{Q}(y, \lambda, \theta(\tau)) \\
&\quad \leq L(x, \hat{u}^1(\tau), \hat{\theta}(\tau)) - L(y, \hat{u}^1(\tau), \hat{\theta}(\tau)) \\
&\qquad + \lambda^T (g(x, \hat{u}^1(\tau), \hat{\theta}(\tau)) - g(y, \hat{u}^1(\tau), \hat{\theta}(\tau))) \\
&\quad \leq C_2 (|x - y| + |\lambda| |x - y|)
\end{aligned}$$

where C_2 is Lipschitz constant of L and g.

Then, as $S \to +\infty$, we obtain

$$H(x, \lambda) - H(y, \lambda) \leq C_2 (|x - y| + |\lambda| |x - y|) \tag{13}$$

Hence, by using Theorem 3.15 in [14, p.158], we conclude that the (HJB) equation has a unique continuous viscosity solution satisfying boundary condition(8). $\blacksquare$

Theorem 3.1 *Assume that A0 holds. Then (9) holds where V_0 is given by the unique viscosity solution of Hamilton-Jacobi equations (7).*

Proof: The optimality principle implies the following:

$$\begin{aligned}
V_\epsilon(t, x, i) \;=\; & \overline{opt}_\gamma E^u_{x,i} \{ \int_t^{t+\Delta} L(x(\tau), u^1(\tau), \theta(\tau)) d\tau \\
& + V_\epsilon(t + \Delta, x(t + \Delta), \theta(t + \Delta)) \}
\end{aligned} \tag{14}$$

Let $x \in D$. Choose some sequence ϵ_n converging to 0, and let $\epsilon(k)$ be a further subsequence along which a limit $V_0(t, x)$ exists (as in (10)). The equations and inequalities below are thus restricted to ϵ within this subsequence. Using (12), we obtain

$$\begin{aligned}
&\Big| V_0(t, x) - \overline{opt}_\gamma E^\gamma_{x,i} \{ \int_t^{t+\Delta} L(x(\tau), u(\tau), \theta(\tau)) d\tau \\
&\quad + V_0(t + \Delta, x(t + \Delta)) \} \Big| \leq 2\hat{\mu}(\epsilon)).
\end{aligned} \tag{15}$$

Let now $v(t', x')$ have continuous partial derivatives and satisfy the condition $v(t, x) = V_0(t, x)$ and $v(t', x') \geq V_0(t', x')$ for (t', x') in some neighborhood of (t, x). One obtains for Δ small:

$$\begin{aligned}
v(t, x) \;\leq\; & \overline{opt}_\gamma E^\gamma_{x,i} \Big\{ \int_t^{t+\Delta} L(x(\tau), u^1(\tau), \theta(\tau)) d\tau \\
& + v(t + \Delta, x(t + \Delta)) \Big\} + 2\hat{\mu}(\epsilon)
\end{aligned} \tag{16}$$

Taking into account that $v(t, x)$ has partial derivatives we obtain $v(t + \Delta, x(t + \Delta))$ as a Taylor series:

$$\begin{aligned}
v(t + \Delta, x(t + \Delta)) \;=\; & v(t, x) + \frac{\partial v(t, x)}{\partial x} (x(t + \Delta) - x) \\
& + \frac{\partial v(t, x)}{\partial t} \Delta + o(\Delta).
\end{aligned} \tag{17}$$

We used above the fact that

$$\begin{aligned}
|x(t + \Delta) - x| &= \left| \int_t^{t+\Delta} g(x(\tau), u^1(\tau), \theta(\tau)) d\tau \right| \\
&= \left| \int_t^{t+\Delta} g(x, u^1(\tau), \theta(\tau)) d\tau + o(\Delta) \right| \leq \tilde{\xi} \Delta,
\end{aligned}$$

for some constant $\tilde{\xi}$, since g is Lipschitz in $D_1 \times U_1$. Substituting (17) in (16) and dividing by Δ, one obtains

$$\frac{\partial v(t,x)}{\partial t} + \overline{opt}_\gamma E^\gamma_{x,i} \frac{1}{\Delta}\Big\{ \int_t^{t+\Delta} (L(x, u^1(\tau), \theta(\tau)) \\ + \frac{\partial v(t,x)}{\partial x} g(x, u^1(\tau), \theta(\tau)) d\tau \Big\} + \frac{2\hat{\mu}(\epsilon)}{\Delta} \\ + \frac{o(\Delta)}{\Delta} + O(\Delta) \geq 0. \tag{18}$$

In changing the arguments of L and g from $x(\tau)$ to x, we used the fact that L and g are Lipschitz in D_1. Define now $\Delta(\epsilon)$ as follows

$$\Delta(\epsilon) = \max\{(\hat{\mu}(\epsilon))^{1/2}, \epsilon^{1/2}\}$$

Let $\sigma = (\tau - t)\epsilon^{-1}$ and $S = \frac{\Delta(\epsilon)}{\epsilon}$,we have

$$\frac{\partial v}{\partial t}(t,x) + \overline{opt}_\gamma E^\gamma_{x,i} \frac{1}{S}\Big\{ \int_0^S (L(x, \hat{u}^1(\sigma), \hat{\theta}(\sigma)) \\ + \frac{\partial v(t,x)}{\partial x} g(x, \hat{u}^1(\sigma), \hat{\theta}(\sigma)) \Big\} + \frac{O(\hat{\mu}(\epsilon))}{\Delta(\epsilon)} + \frac{o(\Delta(\epsilon))}{\Delta(\epsilon)} \\ + O(\Delta(\epsilon)) \geq 0 \tag{19}$$

hence, passing to the limit in (19) as $\epsilon(k)$ tends to zero, then from (6), one obtains

$$\frac{\partial v(t,x)}{\partial t} + H\left(x, \frac{\partial v(t,x)}{\partial x}\right) \geq 0$$

which means that $V_0(t,x)$ is a viscosity super-solution of (7) on $[t_0, T] \times \mathbb{R}^p$.
Similarly, taking $v(t,z)$ having continuous partial derivatives and satisfying the conditions: $v(t,x) = V_0(t,x)$ and $v(t',x') \leq V_0(t',x')$ in some neighborhood of (t,x), one can obtain that

$$\frac{\partial v(t,x)}{\partial t} + H\left(x, \frac{\partial v(t,x)}{\partial x}\right) \leq 0$$

which means that $V_0(t,x)$ is a viscosity super-solution of (7). Thus, $V_0(t,x)$ is a viscosity solution of (7) on $[t_0, T] \times \mathbb{R}^p$. According to Lemma 3.1, (7) has a unique solution, and therefore $V_0(t,x)$ does not depend on the subsequence ϵ_n. This concludes the proof. ∎

Remark 1 The method of asymptotic decomposition of singular perturbed problem into "auxiliary" fast problem with fixed slow parameters x and λ and limit HJB equations as well as the idea of the proof of theorem 3.1 were exposed in [11] in a differential game setting. Other related papers are [2] and [9].

4 Extension of the simple model

The assumptions required in the previous section turn out to be quite restrictive, and often hard to verify. They exclude the linear quadratic model. This motivates us to study new weaker assumptions instead of the previous ones. In particular, (12) need not hold in this section.

Basic Assumptions Introduce the following assumptions.
Assumption A1: The continuous Markov chain $\theta(t)$ has a unichain structure under any stationary policy for the player 2 which does not depend on the state $x(t)$.

Assumption A2: We make the following assumption for the team problem: There exists a subset $\Re$ of policies such that

$$\inf_{\gamma \in \Re} J_\epsilon(t,x,i,\gamma) = V_\epsilon(t,x,i)$$

for all t, x, i and all ϵ small enough; moreover, for any compact D, there exists $C > 0$ such that for all initial values $x \in D$ for all ϵ small enough

$$\forall \gamma \in \Re, \quad \forall t \in [t_0, T] \quad E^\gamma_{x.i}[(|x(t)|^{\alpha+1})] \leq C,$$

for some $\alpha > 0$, where $x(t)$ is a solution of system (1) obtained with the control γ.

We make the following assumption for the game problem: There exists a subset $\Re^1$ such that

$$\inf_{\gamma^1 \in \Re^1} \sup_{\gamma^2 \in U_2} J_\epsilon(t,x,i,\gamma) = V_\epsilon(t,x,i)$$

for all t, x, i and all ϵ small enough; moreover, for any compact D, there exists $C > 0$ such that for all initial values $x \in D$ for all ϵ small enough

$$\forall \gamma^1 \in \Re_1, \forall \gamma^2 \in U_2, \quad \forall t \in [t_0, T] \quad E^\gamma_{x.i}[(|x(t)|^{\alpha+1})] \leq C,$$

for some $\alpha > 0$, where $x(t)$ is a solution of system (1) obtained with the control γ.

Assumption A3: There exists M such that for all $(x_l, u^1_l, i) \in \mathbb{R}^p \times \mathbb{R}^r \times \Theta$, $l = 1, 2$,

$$|g(x_1, u^1_1, i) - g(x_2, u^2_2, i)| \leq M(|x_1 - x_2| + |u^1_1 - u^2_1|). \tag{20}$$

and for all $(x_l, u^1_l, i) \in \mathbb{R}^p \times \mathbb{R}^r \times \Theta$, $l = 1, 2$,

$$|L(x_1, u^1, i) - L(x_2, u^1, i)| \leq M(1 + |x_1|^\alpha + |x_2|^\alpha)|x_1 - x_2|$$

$$|L(x, u^1, i)| \leq M(1 + |x|^{\alpha+1} + |u^1|^{\alpha+1}) \tag{21}$$

and $|\varphi(x_1) - \varphi(x_2)| \leq M(1 + |x_1|^\alpha + |x_2|^\alpha)|x_1 - x_2|$

Assumption A4 For each $(x, i) \in \mathbb{R}^p \times \Theta$, there exists $\kappa > 0$ and $c_0 > 0$ such that

$$\kappa |u^1|^{\alpha+1} \leq L(x, u^1, i) + c_0$$

We begin by stating the main results.

Theorem 4.1 *Let Assumption A1-A4 be true and assume that the Hamilton-Jacobi equations (7) have a unique value V_0 that satisfies the boundary condition (8). Then the limit (9) exists and equals to this solution.*

The proof of the Theorem is quite technical [8]. It is based on the following Lemmas whose proofs too can be found in [8].

Lemma 4.1 *Under Assumption A2 and A3, for any compact D, there exists $\bar{M}_2 > 0$ such that for all $x \in D$, and $y \in \mathbb{R}^p$*

$$V_\epsilon(t,x,i) - V_\epsilon(t,y,i) \leq \bar{M}_2(|x-y| + |x-y|^{\alpha+1}). \quad (22)$$

Lemma 4.2 *Under A2 and A3, for any compact D, for all $(t_k, x, i)) \in ([t_0,T] \times D \times \Theta)^2$ $k = 1,2$, there exists M_1^* such that*

$$|V_\epsilon(t_1,x,i) - V_\epsilon(t_2,x,i)| \leq M_1^*(|t_1 - t_2|) \quad (23)$$

Lemma 4.3 *Under A1-A4, for any $(t,x) \in [t_0,T] \times D$, and for any $i, i^* \in \Theta$ there exists some $\bar{\mu}(\epsilon)$ converging to 0 as $\epsilon \to 0$, such that*

$$|V_\epsilon(t,x,i) - V_\epsilon(t,x,i^*)| \leq \bar{\mu}(\epsilon)$$

Lemma 4.4 *Given any sequence ϵ_l tending to zero, one can find a subsequence $\epsilon_{l_k} = \epsilon_k$ of this sequence such that there exists the limit*

$$V_0(t,x) \overset{\text{def}}{=} \lim_{\epsilon_k \to 0} V_{\epsilon_k}(t,x,i), \quad (24)$$

with the convergence being uniform on compact set $[0,T] \times D_1 \times \Theta$.

We then proceed in three steps. We first show that one may restrict to controls in $\Re$ such that for any compact D, and Δ small enough $\forall x \in D$ and $i \in \Theta$

$$E^\gamma_{x,i} \int_t^{t+\Delta} |u^1(\tau)|^{\alpha+1} d\tau \leq \bar{M}_4 \Delta + \bar{\mu}_1(\epsilon) \quad (25)$$

where $\bar{M}_4$ is a constant independent of x, t and Δ, $\bar{\mu}_1(\epsilon)$ tending to zero as ϵ tend to zero, and α is defined in Assumption A3.

Denote by $\Re'$ the subspace of controls satisfying (25). Let $\gamma \in \Re'$. We show in a second step of the proof that for every $x \in D$, there exists $\bar{M}_5 > 0$ and $\tilde{\mu}(\epsilon)$ tending to zero as ϵ tend to zero it satisfies

$$E^\gamma_{x,i}(|x(t+\Delta) - x|^{\alpha+1}) \leq \bar{M}_5 \Delta^{\alpha+1} + \tilde{\mu}(\epsilon). \quad (26)$$

Using the previous Lemmas and steps, the remainder of the proof follows by adapting the proof of Theorem 3.1 to our new setting [8].

5 Application to linear quadratic control

We consider in this section the optimal control of linear jump parameter systems under a quadratic cost criterion. Now we introduce the continuous dynamic

$$\frac{\partial x(t)}{\partial t} = A(\theta(t))x(t) + B(\theta(t))u^1(t)$$

The quadratic cost is

$$L(x,u^1,i) = |x|^2_{Q(i)} + |u^1|^2_{R(i)} \quad (27)$$

$$\varphi(x,i) = |x|^2_{Q_f(i)} \quad (28)$$

where $Q(i) > 0$ (resp $Q_f(i) > 0$ and $R(i) > 0$), and $|x|_{Q(.)}$ (resp $|x|_{Q_f(.)}$ and $|x|_{R(.)}$) denotes the Euclidean norm. In order to define the HJB, we first compute

$$\hat{Q}(x,\lambda,i) := \inf_{u^1} \Big(L(x,\hat{u}^1,i) + \lambda^T g(x,\hat{u}^1,i) \Big).$$

It can be easily seen that the unique minimizing control u^1 is

$$u^1 = -\frac{1}{2} R^{-1}(i) B^T(i) \lambda,$$

whose substituting leads $\hat{Q}(x,\lambda,i) =$

$$x^T Q(i) x + \lambda^T A(i) x - \frac{1}{4} \lambda^T B(i) R^{-1}(i) B^T(i) \lambda.$$

We obtain a family of infinite horizon stochastic games (or team problems), all with the same state and action spaces as above and the controlled rate matrix (of transition within Θ') which is $\Lambda' = (\mu_{iaj}),\ i,j \in \Theta,\ a \in U_2(i)$, and with the immediate cost defined as $\hat{Q}(x,\lambda,i)$, whose value is $H(x,\lambda)$ (it does not depend on i).

We assume the following on the terminal cost:

Assumption A5 The terminal weighting matrices $Q_f(i)$, $i \in \Theta$ are of the following structure:

$$Q_f(i) = Q_0 + \epsilon Q'_f(i) \quad \forall i \in \Theta,$$

where Q_0, and $Q'_f(.)$ are symmetric.

Theorem 5.1 *Let Assumption A1 and A5 be true, and assume that the following Hamilton-Jacobi equation has a unique solution V_0.*

$$\frac{\partial V_0(t,x)}{\partial t} + H\left(x, \frac{\partial V_0(t,x)}{\partial t}\right) = 0. \quad (29)$$

and satisfies the boundary condition

$$V_0(T,x) = |x|^2_{Q_0}, \quad \forall x \in \mathbb{R}^p. \quad (30)$$

Then $$\lim_{\epsilon \to 0} V_\epsilon(t,x,i) = V_0(t,x). \quad (31)$$

Proof: From Theorem 4.1, it is sufficient to show that special case satisfied Assumption A1-A4. It is clear that our system verifies Assumptions A3 and A4.
We now prove that assumption A2 holds.
For some positive definite $p \times p$-matrices $L'(t_0,i)$, $i \in \Theta$, we say that the policy is admissible ($\gamma \in \bar{\Re}_Q$) if for all ϵ small enough and $\forall (t,x) \in [t_0,T] \times \mathbb{R}^p$

$$E^\gamma_{x,i}[x^T(t)x(t)] \leq x^T L'(t_0,i) x \ ,$$

where $x(t)$ is s solution to (1) obtained with control γ and initial state (x,i).
We show that any policy in $\bar{\Re}$ satisfies A2, and that $\bar{\Re}$ contains δ-optimal policies for every $\delta > 0$.

From [1] we know that there are positive matrices $\bar{P}_\epsilon(t,i)$ and $\tilde{P}_\epsilon(t,i)$, $i \in \Theta$, $t \in [t_0,T]$, such that

$$x^T \bar{P}_\epsilon(t,i) x \leq V_\epsilon(t,x,i) \leq x^T \tilde{P}_\epsilon(t,i) x. \quad (32)$$

The lower and upper bounds in (32) are obtained by using some fixed stationary policies for player 2 (one of these bounds is obtained, however, for a policy defined for some modified controlled transitions). From Theorem 2 in [16], we know that the optimal value when the continuous time Markov process is not controlled (finite horizon) satisfies

$$V_\epsilon(t,x,i) = x^T(P(t,i)+O(\epsilon))x. \qquad (33)$$

This implies that the bounds in (32) can be written, as functions of ϵ, as

$$\begin{aligned} x^T\bar{P}(t,i)x + \mu_1(\epsilon)x^Tx &\leq V_\epsilon(t,x,i) \\ &\leq x^T\tilde{P}(t,i)x + \mu_2(\epsilon)x^Tx \end{aligned}$$

where $\bar{P}(t,i)$ and $\tilde{P}(t,i)$ are an $p \times p$ some positive symmetric matrix for each $i \in \Theta$, $t \in [t_0,T]$, and $\mu_k(\epsilon)$ tending to zero as ϵ tend to zero, $k = 1,2$.
Let $L'(t_0,i) = \beta(\tilde{P}(t_0,i) + 1)$ where $\beta = \sup_{t\in[t_0,T],i\in\Theta}\beta(t,i)$, where

$$\|x\|^2 = x^Tx \leq \beta(t,i)\|x\|^2_{\bar{P}(t,i)}.$$

Since the matrices $Q(i)$ and $Q_f(i)$ are positive definite, then $\bar{P}(t,i)$, $\forall(t,i) \in [t_0,T]\times\Theta$, is positive matrix and continuous in t, hence $\beta(t,i)$ is continuous in t $\forall i \in \Theta$ then $\beta(t,i) < \infty$, $\forall(t,i) \in [t_0,T]\times\Theta$.
If we take a policy $\tilde{\gamma} \notin \Re_Q$, we have

$$\exists t^* \in [t_0,T] \quad s.t \quad E^\gamma_{x,i}[x^T(t^*)x(t^*)] > x^TL'(t_0,i)x$$

then we show that for all ϵ small enough

$$\begin{aligned} J_\epsilon(t_0,x,i,\tilde{\gamma}) \quad &\geq \quad V_\epsilon(t_0,x,i) + (1+\beta\mu_1(\epsilon)+\mu_2)x^Tx \\ &\quad +\beta\mu_1(\epsilon)x^T\tilde{P}(t_0,i)x \end{aligned}$$

then the policy $\tilde{\gamma}$ is not ϵ-optimal policy. Hence for any compact set D, there exists $N_1 > 0$ such that for all ϵ small enough we have $\forall(\gamma,t,x,i)\Re_Q \times [t_0,T]\times D\times\Theta$

$$E^\gamma_{x,i}[x^T(t)x(t)] \leq N_1 \qquad (34)$$

Moreover, if we take one policy $\hat{\gamma} = (\hat{\gamma}^1,\hat{\gamma}^2)$ such that $\hat{\gamma}^1 = -K(\theta(t))x(t)$, then the system (1) (we note $\tilde{\Re}_Q$ subset of these strategies) under $\hat{\gamma} \in \tilde{\Re}_Q$ yields

$$\frac{dx}{dt} = (A(\theta(t)) - K(\theta(t)))x(t), \qquad x(t_0) = x. \qquad (35)$$

Since $t \in [t_0,T]$, then for any compact set D, there exists $N_2 > 0$ such that for all ϵ small enough

$$\forall\hat{\gamma} \in \tilde{\Re}_Q, \qquad |x(t)| \leq N_2$$

hence by using the same procedure to derive general case and the theorem is proven. ∎

References

[1] E. Altman, T. Başar and Z. Pan, "Piecewise-deterministic differential games and dynamic teams with hybrid controls," *Anals of Dynamic Games*, vol. **6**, Birkhäuser, V. Gaitsgory, J. Filar and K. Mizukami, editors, 1999.

[2] Z. Artstein and V. Gaitsgory, "Linear-quadratic tracking of coupled slow and fast targets," *Mathematics of Control Systems Signals*, vol. **10**, pp.1-30, 1997.

[3] A. Bensoussann, and G. L. Blankenship, "Singular perturbations in stochastic control," in *Singular Perturbations and Asymptotic Analysis in Control Systems*, P. V. Kokotovic, A. Bensoussann and G. L. Blankenship (Eds.), Lecture Notes in Control and Information Sciences. Springer-Verlag, New York, pp. 171–260, 1987.

[4] A. Bensoussan, *"Perturbation methods in optimal control problems,"* John Wiley, New York, 1989.

[5] D. Bielecki and J. A. Filar, "Singular perturbed Markov control problem: limiting average cost," *Annals of Operations Research*, vol. **28**(1), pp. 153–168, 1991.

[6] E. K. Boukas, Q. Zhang and G. Yin, "Robust production and maintenance planning in stochastic manufacturing systems," *IEEE Trans. Automat. Control*, vol. **40**(6), pp. 1098–1102, 1995.

[7] F. Delebecque and J. Quadrat, "Contribution of stochastic control singular perturbation averaging and team theories to an example of large scale systems: management and hydropwer production," *IEEE Trans. Automat. Control*, vol. **23**(2), pp. 209–222, 1978.

[8] R. El Azouzi, E. Altman and V. Gaitsgory, "The averaging principle for perturbations of continuous time control problems with fast controlled jump parameters," manuscript.

[9] J. Filar, V.Gaitsgory and A.Haurie, "Control of Singularly Perturbed Hybrid Stochastic Systems," To appear in *IEEE Trans. on Automatic Control*, 2000.

[10] V. Gaitsgory, "Use of the averaging method in control problems," *Differential Equations* (translated from Russian), vol. **22**, pp. 1290-1299, 1986.

[11] V.Gaitsgory, "Limit Hamilton-Jacobi-Isaacs Equations for Singularly Perturbed Zero-Sum Differential Games," *J. Math. Anal. Appl*, vol. **202**, pp.862-899, 1996.

[12] I. I. Gikhman, "Po povodu odnoi teoremi n.i. bogolubova," *Ukrain. Mat. Zh.*, vol. **4**(2), pp. 215–218, 1956.

[13] H. Kushner, *"Weak convergence and singularly perturbed stochastic control and filtering problems,"* Birkhauser, Boston-Basel, Berlin, 1990.

[14] Martino Bardi and Italo Capuzzo-Dolcetta, *"Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman Equation,"* Birkhauser, Boston-Basel, Berlin, 1997.

[15] M. T. Nguyen, E. Altman and V. Gaitsgory, "On stochastic hybrid zero-sum games with non-linear slow dynamic," to appear in *the Annals of Dynamic Games*, 2000.

[16] Z. Pan and T. Basar, "H^∞ contol of large scale jump linear systems via averaging and agregation," *Proc. 34th IEEE Conference on Decision and Control*, New Orleans, Louisiana, pp. 2574-2579, December (13-15)1995.

[17] A. A. Pervozvansky and V. Gaitsgory, (1988). *"Theory of suboptimal decisions,"* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

[18] P. Shi, V. Gaitsgory and E. Altman, "On asymptotic optimization of a class of nonlinear stochastic hybrid systems," *ZOR - Methods and Models in Operations Research*, vol. **47**(2), 1998.