



HAL
open science

Characterization of the optimal trajectories for the averaged dynamics associated to singularly perturbed control systems

Oana Silvia Serea

► **To cite this version:**

Oana Silvia Serea. Characterization of the optimal trajectories for the averaged dynamics associated to singularly perturbed control systems. 2014. hal-00933680

HAL Id: hal-00933680

<https://hal.science/hal-00933680>

Submitted on 20 Jan 2014

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.

Characterization of the optimal trajectories for the averaged dynamics associated to singularly perturbed control systems*

Oana-Silvia Serea

Université de Perpignan Via Domitia,
Laboratoire de Mathématiques et de Physique, EA 4217,
52 avenue Paul Alduy, 66860 Perpignan Cedex, France.
Email: oana-silvia.serea@univ-perp.fr

January 20, 2014

Abstract

The aim of this paper is to study singularly perturbed control systems. Firstly, we provide linearized formulation version for the calculus of the value function associated with the averaged dynamics. Secondly, we obtain necessary and sufficient conditions in order to identify the optimal trajectory of the averaged system.

Key words: Optimal control, Singular perturbations, occupational measures.

AMS Classification: 34A60, 34K26, 49J24, 49J45, 49L25, 93C15.

1 Introduction

The aim of this paper is to study singularly perturbed control systems. Firstly, we provide linearized formulation version for the calculus of the value function associated with the averaged dynamics. Secondly, we obtain necessary and sufficient conditions in order to identify the optimal trajectory of the averaged system.

Linear programming techniques have proved to be very useful in dealing with deterministic and stochastic control problems. A wide literature is available on the subject both in the deterministic and in the stochastic setting ([7, 3, 11, 12, 10, 14, 4, 6]).

One of the advantages of transforming a nonlinear control problem into a linear optimization problem consists in the possibility of obtaining approximation results for the value function. Following the methods presented in [6] and [13], one can approximate the occupational measures by Dirac measures and construct an optimal feedback control. Moreover, when considering the ergodic control problem (see, e.g. [2]), the study of the behavior of the value function is simplified whenever this value is expressed by a linear problem.

Recently linearized versions of the standard continuous infinite horizon discounted control problems are provided in [13, 5]. When the perturbed system is fully nonlinear it is very difficult to characterize the optimal trajectories using the classical methods given by the Pontryagin maximum principle because we do not know exactly the form of the averaged dynamics. For works in this direction we refer the reader for instance to [8, 20, 9] and references therein.

Fortunately, using our approach we can characterize optimal trajectories. Our method consists in

*This version improves the paper with the same title published in Journal of Differential Equations (Volume 255, Issue 11, 2013) by fixing a small gap in the proof of Theorem 14 and correcting some misprints.

embedding the trajectories into a space of probability measures satisfying a convenient condition. This condition is given in terms of the coefficient functions. The results allow to characterize the set of constraints as the closed convex hull of occupational measures associated to controls. We first consider a general Mayer control problem for explaining the approach. Secondly, using linearization techniques and the dual formulation we characterize the optimal occupational measures. Note that as far as we know our method was not used before in the singularly perturbed setting. Moreover, it's advantage consists in the fact that we do not require to "explicitly" calculate the averaged dynamics.

This paper is organized as follows: In the second section we aim at presenting singularly perturbed control systems and the averaging method and some important results concerning the singularly perturbed systems and the value functions associated with this problem. The third section deals with the linear formulation associated with classical control problems. The fourth section concerns with the characterization of optimal occupational measures for Mayer type control problems. In the fifth section we provide linearization techniques for the averaged system. In the sixth section we describe optimal occupational measures/trajectories for the averaged system. Finally, we provide an appendix for the convenience of the reader.

2 Singularly perturbed control systems

In the following we shortly present our problem. We consider the following dynamics:

$$(1) \quad \begin{cases} dx_s^{t,x,y,u} = f(x_s^{t,x,y,u}, y_s^{t,x,y,u}, u_s) ds, \\ \varepsilon dy_s^{t,x,y,u} = g(x_s^{t,x,y,u}, y_s^{t,x,y,u}, u_s) ds, \end{cases}$$

for all $s \geq t$, where $(t, x, y) \in [0, \infty) \times \mathbb{R}^M \times \mathbb{R}^N$ and ε is a small real parameter. The evolutions of the two state variables x and y of the system are of different scale. We call x the "slow" variable and y the "fast" variable. We recall that a control $(u_t)_{t_0 \leq t < \infty}$ is said to be admissible if it is Lebesgue-measurable on $[t_0, \infty)$ and u takes its values in a compact, metric space U . We let \mathcal{U} denote the family of all admissible controls on $[0, \infty)$. The functions $f : \mathbb{R}^M \times \mathbb{R}^N \times U \rightarrow \mathbb{R}^M$ and $g : \mathbb{R}^M \times \mathbb{R}^N \times U \rightarrow \mathbb{R}^N$ are assumed to be continuous on $[0, \infty) \times \mathbb{R}^M \times \mathbb{R}^N$ and Lipschitz-continuous in (x, y) , uniformly with respect to the control parameter $u \in U$.

We denote by $(x_{(\cdot)}^{t,x,y,u;\varepsilon}, y_{(\cdot)}^{t,x,y,u;\varepsilon})$ the solution of (1) starting from $(t, x, y) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$, for some $u \in \mathcal{U}$.

We let $h : \mathbb{R}^M \rightarrow \mathbb{R}$ be a given bounded function and define the following payoff:

$$(2) \quad C_{t,x,y;\varepsilon}(u) = h(x_T^{t,x,y,u;\varepsilon}),$$

for all $(t, x, y) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$ and all $u \in \mathcal{U}$. The value function associated with (1) and (2) is:

$$(3) \quad W_{\varepsilon,h}(t, x, y) = \inf_{u \in \mathcal{U}} C_{t,x,y;\varepsilon}(u)$$

for all $(t, x, y) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$.

The asymptotic behavior of the value function (3) when $\varepsilon \rightarrow 0$ is a very interesting problem. Whenever the control system (1) has some stability property, it is possible to prove that the trajectories $(x_{(\cdot)}^{t,x,y,u;\varepsilon}, y_{(\cdot)}^{t,x,y,u;\varepsilon})$ of (1) converge towards some solution of some system obtained by formally replacing ε by 0 in (1). This is the so called Tikhonov approach which has been successfully developed in [23, 24] for instance.

When $dy_s^{t,x,y,u} = g(x_s^{t,x,y,u}, y_s^{t,x,y,u}, u_s) ds$ is not stable, another approach consists in investigating relationships between the system (1) and a new differential inclusion

$$(4) \quad dx_s^{t,x} \in F(x_s^{t,x})$$

obtained by an averaging method, that will be described later on. We emphasize that, in general, the averaged system is set-valued. We refer the reader to [9, 16] for averaging methods. It is important to notice that only the behavior of the "slow" variable $x_{(\cdot)}^{t,x,y,u;\varepsilon}$ is concerned by this approach.

Let us also recall that the sequence of functions $(W_{\varepsilon,h}(\cdot, \cdot, y))_{\varepsilon>0}$ converges to the value function $W_{F,h} : [0, T] \times \mathbb{R}^M \rightarrow \mathbb{R}$ defined by

$$(5) \quad W_{F,h}(t, x) = \inf_{x_{(\cdot)}^{t,x} \in S_F(t,x)} h(x_T^{t,x}),$$

for all $(t, x) \in [0, T] \times \mathbb{R}^M$ if h is uniformly continuous and bounded.

Here, $S_F(t, x)$ stands for the set of solutions of (4) starting from $(t, x) \in [0, T] \times \mathbb{R}^M$ and $S_\varepsilon(t, x, y)$ for the set of all the trajectories of (1) $(x_{(\cdot)}^{t,x,y,u;\varepsilon}, y_{(\cdot)}^{t,x,y,u;\varepsilon})$ starting from $(t, x, y) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$. Moreover, we define $S_F(0, x) =: S_F(x)$ and by $S_\varepsilon(0, x, y) =: S_\varepsilon(x, y)$, for all $(x, y) \in \mathbb{R}^{M+N}$.

2.1 The averaging method

Let us shortly explain the behavior of the perturbed system (1) as $\varepsilon \rightarrow 0$. If one makes the change of variables $\tau = \frac{s}{\varepsilon}$ in the system (1) with $t = 0$, and sets $(X_\tau, Y_\tau, U_\tau) = (x_{\varepsilon\tau}, y_{\varepsilon\tau}, u_{\varepsilon\tau})$ for $\tau \in [0, \frac{T}{\varepsilon}]$, one gets

$$\begin{cases} dX_\tau^{t,x,y,u} = \varepsilon f(X_\tau^{t,x,y,u}, Y_\tau^{t,x,y,u}, U_\tau) d\tau, \\ dY_s^{t,x,y,u} = g(X_\tau^{t,x,y,u}, Y_\tau^{t,x,y,u}, U_\tau) d\tau. \end{cases}$$

When ε tends to 0, we are led to consider the following associated system:

$$(6) \quad dy_s^{t,y,u} = g(x, y_s^{t,y,u}, u_s) ds,$$

for $s \in [0, +\infty)$, where x is fixed in \mathbb{R}^M . For every fixed $x \in \mathbb{R}^M$, we denote by $y_{(\cdot)}^{t,y,u;x}$ the unique solution of (6) corresponding to the control u and to the initial value y .

We follow an averaging method (cf. for instance [9], [17]): we set, for $(x, y) \in \mathbb{R}^M \times \mathbb{R}^N$, $S > 0$, and $u \in \mathcal{U}$,

$$\begin{aligned} A(x, y, S, u) &= \frac{1}{S} \int_0^S f(x, y_s^{t,y,u;x}, u_s) ds, \\ F(x, y, S) &= \{A(x, y, S, u); u \in \mathcal{U}\}. \end{aligned}$$

We shall make the following assumption on the systems:

$$(7) \quad \begin{cases} \forall R > 0, \text{ there exist nonempty bounded subsets } N_R \text{ and } \Omega_R \text{ of } \mathbb{R}^N \text{ such that:} \\ \mathbf{1} \forall (x, y) \in B(0, R) \times N_R, \forall (x_{(\cdot)}^{0,x,y,u;\varepsilon}, y_{(\cdot)}^{0,x,y,u;\varepsilon}) \in S_\varepsilon(0, x, y), \\ y_s^{0,x,y,u;\varepsilon} \in \Omega_R, \text{ for all } s \in [0, T] \text{ and all } \varepsilon > 0. \\ \mathbf{2} \forall (x, y) \in B(0, R) \times N_R, \forall u \in \mathcal{U}, y_s^{0,y,u;x} \in \Omega_R, \text{ for all } s \geq 0. \end{cases}$$

Note that the previous assumption says in fact that all solutions of (1) starting from $B(0, R) \times N_R$ and all solutions of (6) will belong to $\mathbb{R}^N \times \Omega_R$ and, respectively, to Ω_R . Under an assumption of either total controllability or stability of the associated system (6), the set $F(x, y, S)$ converges, in the sense of the Hausdorff metric, towards a compact convex set $F(x)$ of \mathbb{R}^M .

If the set-valued map F is Lipschitz, then the set of slow solutions $x_{(\cdot)}^{t,x,y,u;\varepsilon}$ converges towards the set of solutions of the differential inclusion in the sense that

$$\Pi_M^\varepsilon(t, x, y) := \left\{ x_{(\cdot)}^{t,x,y,u;\varepsilon} : \left(x_{(\cdot)}^{t,x,y,u;\varepsilon}, y_{(\cdot)}^{t,x,y,u;\varepsilon} \right) \in S_\varepsilon(t, x, y) \right\}$$

converges to $S_F(t, x)$ for the Hausdorff metric.

Let us describe the following property of total controllability: every two points of Ω_R can be joined in bounded time by a trajectory of the associated system for any x in $B(0, R)$. Namely:

$$(8) \quad \begin{cases} \forall R > 0, \exists t_R \geq 0 \text{ such that } : \forall x \in B(0, R), \forall y_1, y_2 \in \Omega_R, \\ \text{there exists } u \in \mathcal{U} \text{ and } s \leq t_R \text{ such that } y_s^{0, y_2, u; x} = y_1. \end{cases}$$

We notice that, the previous assumption is in fact a controllability assumption. The controllability of systems is a very important property and a large literature studying it is available (see for instance [22], or Chapter 2 and 3 from [19]). Moreover, if (8) holds true, then, for every x in $B(0, R)$, Ω_R is actually invariant for the associated system (6). Indeed, if y is in Ω_R , then it can be joined to a y_0 in N_R , so that any point of a trajectory with initial value y is also a point of a trajectory with initial value y_0 . Hence, it belongs to Ω_R by the assumption (7). Conditions for invariance (or in particular, for having (7) and (8)) are presented for instance in [1] (Chapter 5).

We also need the following stability property:

$$(9) \quad \begin{cases} \forall R > 0, \text{ there exists } \xi_R \in L^1([0, +\infty[, \mathbb{R}^+) \text{ such that } : \\ \forall x \in B(0, R), \forall y_1, y_2 \in \Omega_R, \forall u \in \mathcal{U}, \\ \left| y_s^{t, y_1, u; x} - y_s^{t, y_2, u; x} \right| \leq \xi_R(s) |y_1 - y_2| \text{ for every } s \geq 0. \end{cases}$$

Remark 1 *If one assumes the existence of some constant C_R such that:*

$$(y_2 - y_1, g(x, y_2, u) - g(x, y_1, u)) \leq -C_R |y_2 - y_1|^2,$$

for every $u \in U$, then (9) is satisfied with $\xi_R(\tau) = e^{-C_R \tau}$. This is a classical assumption in order to obtain Tychonoff's Theorem,

Under the assumption (9), the map F is Lipschitz (cf. [9]), while under the assumption (8) alone, the map F is only continuous, so one needs further conditions to get F Lipschitz-continuous. Whenever the system is weakly coupled (i.e. if the trajectories $y_{(\cdot)}^{t, y, u; x}$ of the associated system (6) do not actually depend on x), then F is Lipschitz (cf. [17]). Another possibility to get F Lipschitz (or, at least, locally Lipschitz (cf. [21])) is to strengthen the assumption (8) into an assumption (10) of Lipschitz controllability. This condition should guarantee that every two points of Ω_R can be joined in a time less or equal to some constant C_R multiplied by the distance between the initial points. We say that the system (1) satisfies (10) if:

$$(10) \quad \begin{aligned} \forall R > 0, \exists c_R > 0, \forall y_1, y_2 \in \Omega_R, \forall x \in B(0, R), \\ \exists u \in \mathcal{U}, \exists s \leq c_R |y_1 - y_2|, \text{ such that } y_s^{0, y_2, u; x} = y_1. \end{aligned}$$

When F is locally Lipschitz, one has the following result:

Proposition 2 (cf. [21]) *We assume that (7) holds true, either the assumption (8) or the assumption (9) is true, f is bounded and F is locally Lipschitz. Then, for every $R_0 > 0$, every $R \geq R_0 + T \|f\|_\infty$, there exists a function $\mu_R : (0, +\infty) \rightarrow \mathbb{R}$, such that $\lim_{\varepsilon \rightarrow 0} \mu_R(\varepsilon) = 0$ and, for every $\varepsilon > 0$ and every $(x, y) \in B(0, R_0) \times N_R$:*

-for every $\left(x_{(\cdot)}^{t, x, y, u_\varepsilon; \varepsilon}, y_{(\cdot)}^{t, x, y, u_\varepsilon; \varepsilon} \right) \in S_\varepsilon(t, x, y)$, there exists $x_{(\cdot)}^{t, x} \in S_F(t, x)$ and:

$$(11) \quad \sup_{s \in [t, T]} \left| x_s^{t, x, y, u_\varepsilon; \varepsilon} - x_s^{t, x} \right| \leq \mu_R(\varepsilon);$$

-for any $x_{(\cdot)}^{t, x} \in S_F(t, x)$, there exists $(x_\varepsilon(\cdot), y_\varepsilon(\cdot)) \in \left(x_{(\cdot)}^{t, x, y, u_\varepsilon; \varepsilon}, y_{(\cdot)}^{t, x, y, u_\varepsilon; \varepsilon} \right) \in S_\varepsilon(t, x, y)$, such that the inequality (11) holds.

Notice that the previous result can be reformulated using the Hausdorff distance for sets. More precisely, we have that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} d_H(S_\varepsilon(t, x, y), S_F(t, x) \times \{0\}) &= 0 \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} d_H(\text{cl}S_\varepsilon(t, x, y), \text{cl}S_F(t, x) \times \{0\}) &= 0 \Leftrightarrow \\ \lim_{\varepsilon \rightarrow 0} d_H(\text{cl}S_\varepsilon(t, x, y), S_F(t, x) \times \{0\}) &= 0 \end{aligned}$$

because of the regularity of F which implies that $S_F(t, x)$ is closed with respect to the uniform convergence topology.

If the cost function h is bounded and uniformly continuous, the convergence of the value functions is a direct consequence of the convergence of trajectories. More precisely, we have $W_{\varepsilon, h} \rightarrow W_{F, h}$ with respect with the uniform convergence.

3 Linear formulation for classical control problems

Throughout this section, we will be dealing with the following control system

$$(12) \quad \begin{cases} dx_t^{t_0, x_0, u} = b(x_t^{t_0, x_0, u}, u_t) dt, & t_0 \leq t \leq T, \\ x_{t_0}^{t_0, x_0, u} = x_0 \in \mathbb{R}^N, \end{cases}$$

where $T > 0$ is a finite time horizon, $t_0 \in [0, T]$ and the control u is a measurable function which takes its values in a compact, metric space U . We let \mathcal{U} denote the family of all admissible controls on $[0, T]$. We assume that the dynamics $b: \mathbb{R}^N \times U \rightarrow \mathbb{R}^N$ satisfies

$$(13) \quad \begin{cases} b \text{ is bounded and uniformly continuous on } \mathbb{R}^N \times U, \\ |b(x, u) - b(y, u)| \leq c|x - y|, \end{cases}$$

for all $(x, y, u) \in \mathbb{R}^N \times \mathbb{R}^N \times U$, for some positive real constant $c > 0$.

3.1 Occupational measures

We begin by giving some properties of the linear formulation associated to control problems. Let us suppose that $T > 0$ is a fixed time horizon. We fix $t \geq 0$ and $x_0 \in \mathbb{R}^N$. To every $r > t$ and $u \in \mathcal{U}$, one can associate a couple of occupational measures $\gamma^{t, r, x_0, u} = (\gamma_1^{t, r, x_0, u}, \gamma_2^{t, r, x_0, u}) \in \mathcal{P}([t, r] \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^N)$ defined by

$$\begin{cases} \gamma_1^{t, r, x_0, u}(A \times B \times C) = \frac{1}{r-t} \int_t^r 1_{A \times B \times C}(s, x_s^{t, x_0, u}, u_s) ds, \\ \gamma_2^{t, r, x_0, u} = \delta_{x_r^{t, x_0, u}}, \end{cases}$$

for all Borel sets $A \subset [t, r]$, $B \subset \mathbb{R}^N$ and $C \subset U$. When $x \in \mathbb{R}^N$, δ_x stands for the Dirac measure. One can also define $(\gamma_1^{t, t, x_0, u}, \gamma_2^{t, t, x_0, u}) \in \mathcal{P}(\{t\} \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^N)$ by setting

$$\gamma_1^{t, t, x_0, u} = \delta_{t, x_0, u_t}, \gamma_2^{t, t, x_0, u} = \delta_{x_0}.$$

For every $r \geq t$, the family of occupational measures

$$(14) \quad \Gamma(t, r, x_0) = \left\{ \left(\gamma_1^{t, r, x_0, u}, \gamma_2^{t, r, x_0, u} \right), \text{ for all } u \in \mathcal{U} \right\}$$

can be embedded into a larger set

$$\Theta(t, r, x_0) = \left\{ \begin{array}{l} (\gamma_1, \gamma_2) \in \mathcal{P}([t, r] \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^N), \forall \phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N), \\ \int_{[t,r] \times \mathbb{R}^N \times U \times \mathbb{R}^N} (\phi(t, x_0) + (r-t) \mathcal{L}^u \phi(s, y) - \phi(r, z)) \gamma_1(ds dy du) \gamma_2(dz) = 0 \end{array} \right\},$$

where

$$\mathcal{L}^u \phi(s, y) = \langle b(y, u), D\phi(s, y) \rangle + \partial_t \phi(s, y),$$

for all $\phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N)$ and all $s \geq 0, y \in \mathbb{R}^N$. Here $C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N)$ is the class of real bounded, differentiable functions defined on $\mathbb{R}_+ \times \mathbb{R}^N$ and having Lipschitz gradient.

Remark 3 1. For every finite time horizon $T > 0$, there exists a constant $c > 0$ such that, whenever $u \in \mathcal{U}$, $|x_s^{t,x_0,u}| \leq c(1 + |x_0|)$, for all $0 \leq t \leq s \leq T$ and all $x_0 \in \mathbb{R}^N$. Therefore, whenever $\gamma \in \Gamma(t, r, x_0)$,

$$\text{Supp}(\gamma_1) \subset [t, r] \times K_{x_0,c} \times U, \text{ Supp}(\gamma_2) \subset K_{x_0,c},$$

where $K_{x_0,c} = \{x \in \mathbb{R}^N : |x| \leq c(1 + |x_0|)\}$. We emphasize that the constant c can be chosen independently of t, x_0, u .

Remark 4 The set $\Theta(t, r, x_0)$ contains all occupational measures $\gamma^{t,r,x_0,u(\cdot)}$ issued from x_0 at time t . This follows from the following equality

$$\begin{aligned} -\phi(t, x_0) + \int_{\mathbb{R}^N} \phi(r, z) \gamma_2^{t,r,x_0,u}(dz) &= -\phi(t, x_0) + \phi(r, x_r^{t,x_0,u}) \\ &= \int_t^r (\partial_t \phi(s, x_s^{t,x_0,u}) + \langle b(x_s^{t,x_0,u}, u_s), D\phi(s, x_s^{t,x_0,u}) \rangle) ds \\ &= \int_{[t,r] \times \mathbb{R}^N \times U} (r-t) \mathcal{L}^u \phi(s, y) \gamma_1^{t,r,x_0,u}(ds dy du), \end{aligned}$$

for regular test functions $\phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N)$. One notices that the equality constraint in the definition of $\Theta(t, r, x_0)$ can alternatively be written

$$\phi(t, x_0) + \int_{[t,r] \times \mathbb{R}^N \times U} (r-t) \mathcal{L}^u \phi(s, y) \gamma_1(ds dy du) - \int_{\mathbb{R}^N} \phi(r, z) \gamma_2(dz) = 0.$$

It follows that $\Theta(t, r, x_0)$ is convex.

3.2 Linearized formulation for Mayer type control problems

We consider $g' : \mathbb{R}^N \rightarrow \mathbb{R}$ assumed to satisfy ¹

$$(15) \quad \begin{cases} \text{(i) the function } g' \text{ is bounded,} \\ \text{(ii) there exist real constants } c, c_1 > 0 \text{ such that} \\ |g'(x) - g'(y)| \leq c|x - y|, \quad g'(x) \geq c_1 \end{cases}$$

for all $(x, y) \in \mathbb{R}^N \times \mathbb{R}^N$. To every $(t, x) \in [0, T] \times \mathbb{R}^N$ and $u \in \mathcal{U}$, one associates the cost

$$C_{g'}(t, x, u) = g'(x_T^{t,x,u})$$

¹Note that we make the assumption $g'(x) \geq c_1 > 0$ for simplifying the calculus. Indeed, we can always add a positive constant sufficiently large to the function g' in order to obtain the previous inequality because g' is bounded.

and the corresponding value function

$$(16) \quad V_{g'}(t, x) = \inf_{u \in \mathcal{U}} C_{g'}(t, x, u).$$

Under the assumptions (13) and (15), the value function $V_{g'}$ is the unique bounded, uniformly continuous viscosity solution of the Hamilton Jacobi (HJ) equation

$$(17) \quad \partial_t V_{g'}(t, x) + H(x, DV_{g'}(t, x)) = 0,$$

for all $(t, x) \in [0, T] \times \mathbb{R}^N$, and $V_{g'}(T, \cdot) = g'(\cdot)$ on \mathbb{R}^N , where the Hamiltonian is given by

$$(18) \quad H(x, p) = \inf_{u \in U} \langle b(x, u), p \rangle,$$

for all $(t, x, p) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N$. For proofs of the connection between $V_{g'}$ and (17), the reader is referred to [2] and the references therein.

We also consider the linearized problems

$$\Lambda_{g'}(t, x) = \inf_{\gamma = (\gamma_1, \gamma_2) \in \Theta(t, T, x)} \int_{\mathbb{R}^N} g'(z) \gamma_2(dz)$$

and its dual

$$(19) \quad \eta_{g'}(t, x) = \sup \left\{ \begin{array}{l} \eta \in \mathbb{R} : \exists \phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N) \text{ s.t. } \forall (s, y, v, z) \in [t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N, \\ \eta \leq (T - t) \mathcal{L}^v \phi(s, y) + g'(z) - \phi(T, z) + \phi(t, x) \end{array} \right\},$$

for all $(t, x) \in [0, T] \times \mathbb{R}^N$. The following result links the three quantities. Its proof follows the ideas in [5] and [15]. For reader's convenience, we give the proof in the Appendix.

Proposition 5 *If (13) and (15) hold true, then, for every $(t, x) \in [0, T] \times \mathbb{R}^N$, one has*

$$V_{g'}(t, x) = \Lambda_{g'}(t, x) = \eta_{g'}(t, x).$$

We have the following characterization of the set of constraints $\Theta(t, T, x_0)$:

Corollary 6 *The set $\Theta(t, T, x_0)$ is the closed convex hull of the family of occupational couples $\Gamma(t, T, x_0)$*

$$\Theta(t, T, x_0) = cl(\text{co}(\Gamma(t, T, x_0))),$$

for all $T \geq t \geq 0$. The operator cl designates the closure with respect to the topology induced by the weak convergence of probability measures.

For further details, the reader is referred to [13].

Remark 7 *Consequently, if $\gamma = (\gamma_1, \gamma_2) \in \Theta(t, T, x_0)$, there exists a sequence of convex combinations $\left(\sum_{i=1}^{k_n} \alpha_i^n \gamma^{t, T, x_0, u_i^n} \right)_n$ converging to γ . Thus,*

$$\text{Supp}(\gamma_1) \subset [t, T] \times K_{x_0, c} \times U, \quad \text{Supp}(\gamma_2) \subset K_{x_0, c}.$$

In particular, $\Theta(t, T, x_0)$ is compact.

4 Characterization of optimal measures

In this section we study necessary and sufficient conditions for characterizing optimal occupational measures. In the sequel we denote $\Theta(t, T, x) := \Theta(t, x)$ for simplicity. Recall that

$$V_{g'}(t, x) = \Lambda_{g'}(t, x) = \eta_{g'}(t, x)$$

and

$$(20) \quad \eta_{g'}(t, x) = \sup \left\{ \begin{array}{l} \eta \in \mathbb{R} : \exists \phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N) \text{ s.t. } \forall (s, y, v, z) \in [t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N, \\ \eta \leq (T-t) \mathcal{L}^v \phi(s, y) + g'(z) - \phi(T, z) + \phi(t, x) \end{array} \right\},$$

for all $(t, x) \in [0, T] \times \mathbb{R}^N$. We denote by

$$(21) \quad D_{g'}(t, x) = \left\{ \begin{array}{l} (\eta, \phi) \in \mathbb{R} \times C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N) \text{ s.t.} \\ \eta = \inf_{(s, y, v, z) \in [t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N} \{(T-t) \mathcal{L}^v \phi(s, y) + g'(z) - \phi(T, z) + \phi(t, x)\} \end{array} \right\},$$

for all $(t, x) \in [0, T] \times \mathbb{R}^N$.

Remark 8 We note that $\Theta(t, x)$ contains measures with compact support (see Remark 7). Consequently, we may consider that the test functions have also compact support. Moreover, in the previous formulation, we can replace the infimum over $(s, y, v, z) \in [t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N$ with the minimum over a compact set which contains the support of the occupational measures.

Consequently, the dual formulation becomes:

$$(22) \quad V_{g'}(t, x) = \sup \{ \eta, (\eta, \phi) \in D_{g'}(t, x) \}.$$

Definition 9 We say that $(\bar{\eta}, \bar{\phi}) \in D_{g'}(t, x)$ is an optimal pair whenever we have $V_{g'}(t, x) = \bar{\eta}$.

Remark 10 It is not always sure that optimal pairs exist. Fortunately, in this case we can define the set Ω using the sets $\Omega_{\frac{1}{n}}$ associated to a sequence of pairs $(\eta_{\frac{1}{n}}, \phi_{\frac{1}{n}}) \in D_{g'}(t, x)$ having the property that $\eta_{\frac{1}{n}} \nearrow \eta(t, x) = V(t, x)$. More precisely, 1_Ω coincides with the $\liminf_{n \rightarrow \infty} 1_{\Omega_{\frac{1}{n}}}$. Note that such a sequence always exists because of (22).

Definition 11 We denote by

$$(23) \quad \Omega_{g'}(t, x) = \left\{ \begin{array}{l} (s, y, v, z) \in [t, r] \times \mathbb{R}^N \times U \times \mathbb{R}^N, \text{ s.t.} \\ \bar{\eta} = (T-t) \mathcal{L}^v \bar{\phi}(s, y) + g'(z) - \bar{\phi}(T, z) + \bar{\phi}(t, x), \\ \text{for every optimal pair } (\bar{\eta}, \bar{\phi}) \in D_{g'}(t, x). \end{array} \right\}$$

Proposition 12 Let $(t, x) \in [0, T] \times \mathbb{R}^N$ be fixed. Then, $\gamma \in \Theta(t, x)$ is optimal for $\Lambda_{g'}(t, x)$ iff $\gamma(\Omega_{g'}(t, x)) = 1$.

Proof.

Suppose that $\gamma \in \Theta(t, x)$ is such that $\gamma(\Omega_{g'}(t, x)) = 1$, i.e. the support of γ is included in $\Omega_{g'}(t, x)$. Then, by definition, we have the following equality

$$\bar{\eta} = (T-t) \mathcal{L}^v \bar{\phi}(s, y) + g'(z) - \bar{\phi}(T, z) + \bar{\phi}(t, x).$$

on $\Omega_{g'}(t, x)$, for all optimal pairs $(\bar{\eta}, \bar{\phi}) \in D_{g'}(t, x)$. Consequently,

$$\int_{\Omega_{g'}(t, x)} \bar{\eta} \gamma(ds dx dudz) = \int_{\Omega_{g'}(t, x)} g'(z) \gamma(ds dx dudz) = \int_{[t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N} g'(z) \gamma(ds dx dudz)$$

because $\gamma \in \Theta(t, x)$. It follows that

$$\gamma(\Omega(t, x)) \bar{\eta} = V(t, x) = \int_{[t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N} g'(z) \gamma(ds dx du dz)$$

and that $\gamma \in \Theta(t, x)$ is optimal. Conversely, suppose that $\gamma \in \Theta(t, x)$ is optimal. Then, we have:

$$\begin{aligned} V(t, x) &= \int_{[t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N} g'(z) \gamma(ds dx du dz) = \\ &\int_{[t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N} [(T-t) \mathcal{L}^v \bar{\phi}(s, y) + g'(z) - \bar{\phi}(T, z) + \bar{\phi}(t, x)] \gamma(ds dx du dz) \geq \\ &\int_{\Omega_{g'}(t, x)} [(T-t) \mathcal{L}^v \bar{\phi}(s, y) + g'(z) - \bar{\phi}(T, z) + \bar{\phi}(t, x)] \gamma(ds dx du dz) = \\ &\int_{\Omega_{g'}(t, x)} \bar{\eta} \gamma(ds dx du dz) = \bar{\eta} \gamma(\Omega_{g'}(t, x)) = V(t, x) \gamma(\Omega_{g'}(t, x)). \end{aligned}$$

for all optimal pairs $(\bar{\eta}, \bar{\phi}) \in D_{g'}(t, x)$. Consequently, $\gamma \in \Theta(t, x)$ is such that $\gamma(\Omega_{g'}(t, x)) = 1$. Indeed, if $\gamma(\Omega_{g'}(t, x)) < 1$ we get

$$V(t, x) = \int_{[t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N} g'(z) \gamma(ds dx du dz) > V(t, x).$$

because $V(t, x) > 0$. The last inequality leads to a contradiction. ■

5 Dual formulation for the averaged system

In order to characterize the optimal controls/measures for the averaged system, we first state its dual formulation. In the sequel we assume that (7) holds true, either the assumption (8) or the assumption (9) is true, and F is locally Lipschitz. As previously, let us suppose that $T > 0$ is a fixed time horizon. We fix $\varepsilon > 0$, $t \geq 0$ and $(x_0, y_0) \in \mathbb{R}^M \times \mathbb{R}^N$. To every $u \in \mathcal{U}$, one can associate a couple of occupational measures $\gamma^{t, T, x_0, y_0, u; \varepsilon} = \left(\gamma_1^{t, T, x_0, y_0, u; \varepsilon}, \gamma_2^{t, T, x_0, y_0, u; \varepsilon} \right) \in \mathcal{P}([t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^M \times \mathbb{R}^N)$ defined by

$$\begin{cases} \gamma_1^{t, T, x_0, y_0, u; \varepsilon}(A \times B \times C \times D) = \frac{1}{T-t} \int_t^T 1_{A \times B \times C \times D}(s, x_s^{t, x_0, y_0, u; \varepsilon}, y_s^{t, x_0, y_0, u; \varepsilon}, u_s) ds, \\ \gamma_2^{t, T, x_0, y_0, u; \varepsilon} = \delta_{x_T^{t, x_0, y_0, u; \varepsilon}, y_T^{t, x_0, y_0, u; \varepsilon}}, \end{cases}$$

for all Borel sets $A \subset [t, T]$, $B \subset \mathbb{R}^M$, $C \subset \mathbb{R}^N$ and $D \subset U$. When $(x, y) \in \mathbb{R}^M \times \mathbb{R}^N$, $\delta_{x, y}$ stands for the Dirac measure. One can also define $\gamma^{t, t, x_0, y_0, u; \varepsilon} = \left(\gamma_1^{t, t, x_0, y_0, u; \varepsilon}, \gamma_2^{t, t, x_0, y_0, u; \varepsilon} \right) \in \mathcal{P}(\{t\} \times \mathbb{R}^M \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^M \times \mathbb{R}^N)$ by setting

$$\gamma_1^{t, t, x_0, y_0, u; \varepsilon} = \delta_{t, x_0, y_0, u_t}, \gamma_2^{t, t, x, y, u; \varepsilon} = \delta_{x, y}.$$

The family of occupational measures

$$(24) \quad \Gamma(t, x_0, y_0; \varepsilon) = \left\{ \left(\gamma_1^{t, T, x_0, y_0, u; \varepsilon}, \gamma_2^{t, T, x_0, y_0, u; \varepsilon} \right), \text{ for all } u \in \mathcal{U} \right\}.$$

can be embedded into a larger set

$$\Theta(t, x_0, y_0; \varepsilon) = \left\{ \begin{array}{l} (\gamma_1, \gamma_2) \in \mathcal{P}([t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^M \times \mathbb{R}^N) \\ \forall \phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N), \\ \int_{[t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N} (\phi(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \varepsilon} \phi(s, x, y) \\ - \phi(T, z, w)) \gamma_1(ds dx dy du) \gamma_2(dz dw) = 0, \end{array} \right\}$$

where

$$\mathcal{L}^{u;\varepsilon}\phi(s, x, y) = \left\langle \left(f(x, y, u), \frac{1}{\varepsilon}g(x, y, u) \right), D\phi(s, x, y) \right\rangle + \partial_t\phi(s, x, y),$$

for all $\phi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N)$ and all $s \geq 0, (x, y) \in \mathbb{R}^M \times \mathbb{R}^N$.

Remark 13 *Using similar arguments as in the previous sections, the set $\Theta(t, T, x_0, y_0; \varepsilon)$ contains all occupational measures $\gamma^{t, T, x_0, y_0, u; \varepsilon}$ issued from (x_0, y_0) at time t . Moreover, it is also convex and relatively compact with respect to the weak convergence of probability measures (due to Prohorov's Theorem).*

We suppose that h satisfies the hypotheses (15). As previously, the linearized problem is

$$\Lambda_{\varepsilon, h}(t, x_0, y_0) = \inf_{\gamma=(\gamma_1, \gamma_2) \in \Theta(t, x_0, y_0; \varepsilon)} \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma_2(dzdw)$$

and its dual is

$$(25) \quad \eta_{\varepsilon, h}(t, x_0, y_0) = \sup \left\{ \begin{array}{l} \eta \in \mathbb{R} : \exists \phi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N) \text{ s.t.} \\ \forall (s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R} \times \mathbb{R}^N, \\ \eta \leq (T-t) \mathcal{L}^{v; \varepsilon}\phi(s, x, y) + h(z) - \phi(T, z, w) + \phi(t, x_0, y_0). \end{array} \right\},$$

for all $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$. Consequently, for every $\varepsilon > 0$ we have

$$W_{\varepsilon, h}(t, x_0, y_0) = \Lambda_{\varepsilon, h}(t, x_0, y_0) = \eta_{\varepsilon, h}(t, x_0, y_0).$$

We continue with the dual formulation for the averaged problem. We denote by

$$\Theta(t, x_0, y_0) = \left\{ \begin{array}{l} (\gamma_1, \gamma_2) \in \mathcal{P}([t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U) \times \mathcal{P}(\mathbb{R}^M \times \mathbb{R}^N) \\ \forall \psi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M) \text{ and } \forall \phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N), \\ \int_{[t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N} (\psi(t, x_0) + (T-t) \mathcal{L}^{u, f}\psi(s, x) \\ - \psi(T, z)) \gamma_1(dsdx dy du) \gamma_2(dzdw) = 0 \text{ and} \\ \int_{[t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N} \mathcal{L}^{u, g}\phi(s, x, y) \gamma_1(dsdx dy du) \gamma_2(dzdw) = 0 \end{array} \right\}$$

where

$$\mathcal{L}^{u, f}\psi(s, x, y) = \langle f(x, y, u), D_x\psi(s, x) \rangle + \partial_t\psi(s, x)$$

and

$$\mathcal{L}^{u, g}\phi(s, x, y) = \langle g(x, y, u), D_y\phi(s, x, y) \rangle,$$

for all $\phi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N)$ and all $s \geq 0, (x, y) \in \mathbb{R}^M \times \mathbb{R}^N$.

We define the following linearized problem

$$\Lambda_h(t, x_0, y_0) = \inf_{\gamma=(\gamma_1, \gamma_2) \in \Theta(t, x_0, y_0)} \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma_2(dzdw).$$

We denote by

$$(26) \quad \eta_h(t, x_0, y_0) = \sup \left\{ \begin{array}{l} \eta \in \mathbb{R} : \exists \alpha \in C(\mathbb{R}_+) \text{ with } \lim_{\varepsilon \rightarrow 0} \alpha(\varepsilon) = 0 \text{ s.t. } \forall \varepsilon > 0, \\ \exists \phi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N) \text{ and } \psi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M) \text{ s.t.} \\ \|\phi - \psi\|_{\infty} + \|\nabla_x\phi - \nabla_x\psi\|_{\infty} \leq \alpha(\varepsilon) \text{ and s.t.} \\ \forall (s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R} \times \mathbb{R}^N, \\ \eta \leq (T-t) \mathcal{L}^{v; \varepsilon}\phi(s, x, y) + h(z) - \phi(T, z, w) + \phi(t, x_0, y_0) \end{array} \right\},$$

for all $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$. Consequently, we can formulate the main result of this section:

Theorem 14 *We have that*

$$W_h(t, x_0) = \Lambda_h(t, x_0, y_0) = \eta_h(t, x_0, y_0).$$

for all $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$.

Proof. In a first step we recall that there exist an optimal measure $\bar{\gamma}^{t, T, x_0, y_0; \varepsilon} = \left(\bar{\gamma}_1^{t, T, x_0, y_0; \varepsilon}, \bar{\gamma}_2^{t, T, x_0, y_0; \varepsilon} \right) \in \Theta(t, x_0, y_0; \varepsilon)$ such that

$$\Lambda_{\varepsilon, h}(t, x_0, y_0) = \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \bar{\gamma}_2^{t, T, x_0, y_0; \varepsilon}(dzdw)$$

for all $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$ and $\varepsilon > 0$. Moreover, we can find a subsequence and a probability measure $\bar{\gamma}$ such that $\bar{\gamma}^{t, T, x_0, y_0; \varepsilon} \rightharpoonup \bar{\gamma}$ (see Corollary 6 and Remark 7). Using the definitions of $\Theta(t, x_0, y_0; \varepsilon)$ and $\Theta(t, x_0, y_0)$, it is easy to see that $\bar{\gamma} \in \Theta(t, x_0, y_0)$. Consequently,

$$(27) \quad \Lambda_h(t, x_0, y_0) \leq \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma_2^{t, T, x_0, y_0}(dzdw) = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma_2^{t, T, x_0, y_0; \varepsilon}(dzdw) = \\ \lim_{\varepsilon \rightarrow 0} \Lambda_{\varepsilon; h}(t, x_0, y_0) = \lim_{\varepsilon \rightarrow 0} W_{\varepsilon, h}(t, x_0, y_0) = W_h(t, x_0)$$

for all $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$.

We continue by considering $\gamma \in \Theta(t, x_0, y_0)$ and $\eta \in \mathbb{R}$ such that

$$\begin{aligned} & \exists \alpha \in C(\mathbb{R}_+) \text{ with } \lim_{\varepsilon \rightarrow 0} \alpha(\varepsilon) = 0 \text{ s.t. } \forall \varepsilon > 0, \\ & \exists \phi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N) \text{ and } \psi \in C^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M) \text{ s.t.} \\ & \|\phi - \psi\|_\infty + \|\nabla_x \phi - \nabla_x \psi\|_\infty \leq \alpha(\varepsilon) \text{ and s.t.} \\ & \forall (s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R} \times \mathbb{R}^N, \\ & \eta \leq (T - t) \mathcal{L}^{v; \varepsilon} \phi(s, x, y) + h(z) - \phi(T, z, w) + \phi(t, x_0, y_0) \end{aligned}$$

By integrating with respect to γ we obtain that

$$\eta \leq \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma_2(dzdw)$$

and consequently,

$$\eta_h(t, x_0, y_0) \leq \int_{\mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma_2(dzdw)$$

for all $\gamma \in \Theta(t, x_0, y_0)$. We have that

$$(28) \quad \eta_h(t, x_0, y_0) \leq \Lambda_h(t, x_0, y_0).$$

Using Proposition 19 there exist two families $V_{\varepsilon, h}^{\delta_\varepsilon} \in C_b^{1,1}([0, T + \delta_\varepsilon] \times \mathbb{R}^M \times \mathbb{R}^N)$ and $V_h^{\delta_\varepsilon} \in C_b^{1,1}([0, T + \delta] \times \mathbb{R}^N)$ as in the definition of $\eta(t, x_0, y_0)$ s. t.

$$\begin{aligned} & \forall (s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N, \\ & V_{\varepsilon, h}^{\delta_\varepsilon}(t, x_0, y_0) \leq (T - t) \mathcal{L}^{v; \varepsilon} V_{\varepsilon, h}^{\delta_\varepsilon}(s, x, y) + h(z) - V_{\varepsilon, h}^{\delta_\varepsilon}(T, z, w) + V_{\varepsilon, h}^{\delta_\varepsilon}(t, x_0, y_0). \end{aligned}$$

Moreover $\lim_{\varepsilon \rightarrow 0} V_{\varepsilon, h}^{\delta_\varepsilon}(t, x_0, y_0) = W_h(t, x_0)$. Using the definition of $\eta_h(t, x_0, y_0)$, we obtain that $V_{\varepsilon, h}^{\delta_\varepsilon}(t, x_0, y_0) \leq \eta(t, x_0, y_0)$. Consequently,

$$(29) \quad W_h(t, x_0) = \lim_{\varepsilon \rightarrow 0} V_{\varepsilon, h}^{\delta_\varepsilon}(t, x_0, y_0) \leq \eta_h(t, x_0, y_0)$$

By combining the inequalities (27), (28) and (29) we complete the proof. \blacksquare

6 Characterization of optimal trajectories for the averaged system

As already mentioned in the introduction, when the perturbed system is fully nonlinear it is very difficult to characterize the optimal trajectories using the Pontryagin maximum principle because we do not know exactly the form of the averaged dynamics. An alternative to this method is to look at the support of the occupational measures contained in the set $\Theta(t, x_0, y_0)$ in order to obtain optimal trajectories from every $(t, x_0) \in [0, T] \times \mathbb{R}^M$. We denote by

$$(30) \quad D_{\varepsilon, h}(t, x_0, y_0) = \left\{ \begin{array}{l} (\eta, \phi) \in \mathbb{R} \times C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N) \text{ s.t.} \\ \eta = \inf_{(s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N} \{(T-t) \mathcal{L}^{v; \varepsilon} \phi(s, x, y) \\ + h(z) - \phi(T, z, w) + \phi(t, x_0, y_0)\} \end{array} \right\},$$

for all $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$.

Remark 15 *We note that $\Theta(t, x_0, y_0; \varepsilon)$ contains measures with compact support. Consequently, we may consider that the test functions have also compact support. Moreover, in the previous formulation, we can replace the infimum over $(s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N$ with the minimum over a compact set which contains the support of the occupational measures.*

Consequently, we can reformulate the dual formulation as follows:

$$W_{\varepsilon, h}(t, x_0, y_0) = \sup \{ \eta, (\eta, \phi) \in D_{\varepsilon, h}(t, x_0, y_0) \}.$$

Definition 16 *We say that $(\bar{\eta}_\varepsilon, \bar{\phi}_\varepsilon) \in D_{\varepsilon, h}(t, x_0, y_0)$ is an optimal pair whenever we have $W_{\varepsilon, h}(t, x_0, y_0) = \bar{\eta}_\varepsilon$.*

Definition 17 *Let $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$ be fixed. We denote by*

$$(31) \quad \Omega_{\varepsilon, h}(t, x_0, y_0) = \left\{ \begin{array}{l} (s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N \text{ s.t.} \\ \bar{\eta}_\varepsilon = \{(T-t) \mathcal{L}^{v; \varepsilon} \phi(s, x, y) + h(z) - \phi(T, z, w) + \phi(t, x_0, y_0)\} \\ \text{for every optimal pair } (\bar{\eta}_\varepsilon, \bar{\phi}_\varepsilon) \in D_{\varepsilon, h}(t, x_0, y_0) \end{array} \right\},$$

$$(32) \quad \Omega_{\frac{1}{n}, h}(t, x_0, y_0) = \left\{ \begin{array}{l} (s, x, y, v, z, w) \in [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N \text{ s.t.} \\ \bar{\eta}_{\frac{1}{n}} = \{(T-t) \mathcal{L}^{v; \frac{1}{n}} \phi(s, x, y) + h(z) - \phi(T, z, w) + \phi(t, x_0, y_0)\} \\ \text{for every optimal pair } (\bar{\eta}_{\frac{1}{n}}, \bar{\phi}_{\frac{1}{n}}) \in D_{\frac{1}{n}, h}(t, x_0, y_0) \\ \text{and with } \bar{\phi}_{\frac{1}{n}} \text{ as in the definition of } \eta_h(t, x_0, y_0) \end{array} \right\}$$

and by

$$(33) \quad 1_{\Omega_h} = \liminf_{n \rightarrow \infty} 1_{\Omega_{\frac{1}{n}, h}}.$$

Note that, as in Section 4, $\gamma_\varepsilon \in \Theta(t, x_0, y_0; \varepsilon)$ is optimal iff $\gamma_\varepsilon(\Omega_{\varepsilon, h}(t, x_0, y_0)) = 1$ for all $\varepsilon > 0$.

Theorem 18 *Let $(t, x_0, y_0) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^N$ be fixed. Then, $\gamma \in \Theta(t, x_0, y_0)$ is optimal iff $\gamma(\Omega_h(t, x_0, y_0)) = 1$.*

Proof. First, suppose that $\gamma \in \Theta(t, x_0, y_0)$ is such that $\gamma(\Omega_h(t, x_0, y_0)) = 1$. In order to shorten the formulas, we denote by $\Delta := [t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N$. We have the following

$$\begin{aligned} \int_{\Delta} h(z) \gamma(dsdxdydudzdw) &= \int_{\Omega_h(t, x_0, y_0)} h(z) \gamma(dsdxdydudzdw) \text{ and} \\ \int_{\Omega_h(t, x_0, y_0)} \left(\phi(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \phi(s, x, y) - \phi(T, z, w) + h(z) \right) \gamma_1(dsdxdydu) \gamma_2(dzdw) &= \\ \int_{\Delta} \mathbf{1}_{\Omega_h(t, x_0, y_0)} \left(\phi(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \phi(s, x, y) - \phi(T, z, w) + h(z) \right) \gamma_1(dsdxdydu) \gamma_2(dzdw) &\leq \\ \int_{\Delta} \liminf_{n \rightarrow \infty} \mathbf{1}_{\Omega_{\frac{1}{n}, h}(t, x_0, y_0)} \left(\phi(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \phi(s, x, y) - \phi(T, z, w) + h(z) \right) \gamma_1(dsdxdydu) \gamma_2(dzdw) & \\ \liminf_{n \rightarrow \infty} \int_{\Delta} \mathbf{1}_{\Omega_{\frac{1}{n}, h}(t, x_0, y_0)} \left(\phi(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \phi(s, x, y) - \phi(T, z, w) + h(z) \right) \gamma_1(dsdxdydu) \gamma_2(dzdw) & \\ \forall \phi \in C_b^{1,1}(\mathbb{R}_+ \times \mathbb{R}^M \times \mathbb{R}^N). & \end{aligned}$$

Consider now an optima pair $(\bar{\eta}_{\frac{1}{n}}, \bar{\phi}_{\frac{1}{n}}) \in D_{\frac{1}{n}, h}(t, x_0, y_0)$ and with $\bar{\phi}_{\frac{1}{n}}$ as in the definition of $\eta_h(t, x_0, y_0)$. Using the previous estimation, we obtain

$$\begin{aligned} \int_{\Delta} h(z) \gamma(dsdxdydudzdw) &\leq \\ \liminf_{n \rightarrow \infty} \int_{\Delta} \mathbf{1}_{\Omega_{\frac{1}{n}, h}(t, x_0, y_0)} \left(\bar{\phi}_{\frac{1}{n}}(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \bar{\phi}_{\frac{1}{n}}(s, x, y) - \bar{\phi}_{\frac{1}{n}}(T, z, w) \right) \gamma(dsdxdydudzdw) &= \\ \liminf_{n \rightarrow \infty} \int_{\Delta} \mathbf{1}_{\Omega_{\frac{1}{n}, h}(t, x_0, y_0)} \bar{\eta}_{\frac{1}{n}} \gamma(dsdxdydudzdw) &\leq W_h(t, x_0). \end{aligned}$$

Consequently, $\gamma \in \Theta(t, x_0, y_0)$ is optimal.

Secondly, let $\gamma \in \Theta(t, x_0, y_0)$ be optimal. For an optimal pair $(\bar{\eta}_{\frac{1}{n}}, \bar{\phi}_{\frac{1}{n}}) \in D_{\frac{1}{n}, h}(t, x_0, y_0)$ as in the definition of $\eta(t, x_0, y_0)$ we obtain

$$\begin{aligned} W_h(t, x_0) + \alpha\left(\frac{1}{n}\right) &= \int_{\Delta} h(z) \gamma(dsdxdydudzdw) + \alpha\left(\frac{1}{n}\right) \geq \\ \int_{\Delta} \left(\bar{\phi}_{\frac{1}{n}}(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \bar{\phi}_{\frac{1}{n}}(s, x, y) - \bar{\phi}_{\frac{1}{n}}(T, z, w) \right) \gamma_1(dsdxdydu) \gamma_2(dzdw) &\geq \\ \int_{\Omega_{\frac{1}{n}, h}(t, x_0, y_0)} \left(\bar{\phi}_{\frac{1}{n}}(t, x_0, y_0) + (T-t) \mathcal{L}^{u; \frac{1}{n}} \bar{\phi}_{\frac{1}{n}}(s, x, y) - \bar{\phi}_{\frac{1}{n}}(T, z, w) \right) \gamma_1(dsdxdydu) \gamma_2(dzdw) &= \\ \int_{\Omega_{\frac{1}{n}, h}(t, x_0, y_0)} \bar{\eta}_{\frac{1}{n}} \gamma_1(dsdxdydu) \gamma_2(dzdw) - \alpha\left(\frac{1}{n}\right) &= \gamma\left(\Omega_{\frac{1}{n}, h}(t, x_0, y_0)\right) \bar{\eta}_{\frac{1}{n}}. \end{aligned}$$

and

$$\begin{aligned} W_h(t, x_0) &= \int_{\Delta} h(z) \gamma(dsdxdydudzdw) \geq \liminf_{n \rightarrow \infty} \gamma\left(\Omega_{\frac{1}{n}, h}(t, x_0, y_0)\right) \bar{\eta}_{\frac{1}{n}} \geq \\ W_h(t, x_0) \gamma\left(\liminf_{n \rightarrow \infty} \Omega_{\frac{1}{n}, h}(t, x_0, y_0)\right) &\geq W_h(t, x_0) \gamma(\Omega_h(t, x_0, y_0)). \end{aligned}$$

Consequently, $\gamma \in \Theta(t, x_0, y_0)$ is such that $\gamma(\Omega_h(t, x_0, y_0)) = 1$ Indeed, if $\gamma(\Omega_h(t, x_0, y_0)) < 1$ we get

$$W_h(t, x_0, y_0) = \int_{[t, T] \times \mathbb{R}^M \times \mathbb{R}^N \times U \times \mathbb{R}^M \times \mathbb{R}^N} h(z) \gamma(dsdxdudz) > W_h(t, x_0, y_0)$$

which leads to a contradiction. ■

7 Appendix

We prove the Proposition 5. We will make use of the following result due to N. V. Krylov (cf. [18], Theorem 2.1):

Proposition 19 *There exists a constant $C > 0$ such that, for every $\delta \in (0, 1]$, there exists a function $V^\delta \in C_b^{1,1}([0, T + \delta] \times \mathbb{R}^N)$ (classical) subsolution of (17) defined on $[0, T + \delta] \times \mathbb{R}^N$ satisfying*

$$(i) \left| V^\delta(t, \cdot) - g'(\cdot) \right| \leq C\delta, \text{ for } t \in [T, T + \delta], \text{ and}$$

$$(ii) \left| V^\delta(\cdot) - V_{g'}(\cdot) \right| \leq C\delta, \text{ on } [0, T] \times \mathbb{R}^N.$$

Proof. (of Proposition 5) We have seen that $\gamma^{t,T,x,u} \in \Theta(t, T, x)$, for all $(t, x) \in [0, T] \times \mathbb{R}^N$ and all $u \in \mathcal{U}$. It follows that

$$(34) \quad V_{g'}(t, x) \geq \Lambda_{g'}(t, x).$$

For any $\gamma \in \Theta(t, T, x)$, whenever $(\eta, \phi) \in \mathbb{R} \times C_2^{1,1}(\mathbb{R}_+ \times \mathbb{R}^N)$ satisfies

$$\eta \leq (T - t) \mathcal{L}^v \phi(s, y) + g'(z) - \phi(T, z) + \phi(t, x),$$

for all $(s, y, v, z) \in [t, T] \times \mathbb{R}^N \times U \times \mathbb{R}^N$, we have

$$\eta \leq \int_{\mathbb{R}^N} g'(z) \gamma_2(dz).$$

Hence,

$$(35) \quad \eta_{g'}(t, x) \leq \Lambda_{g'}(t, x) < \infty.$$

To complete the proof, one only needs to show that

$$(36) \quad V_{g'}(t, x) \leq \eta_{g'}(t, x).$$

To this purpose, we apply the previous Proposition and get, for every $\delta > 0$, the existence of some regular V^δ such that

$$\partial_t V^\delta(t, x) + H(t, x, DV^\delta(t, x)) \geq 0,$$

for all $(t, x) \in [0, T] \times \mathbb{R}^N$. Thus, choosing C as in Proposition 19, for every $(t, s, x, y, z) \in [0, T]^2 \times \mathbb{R}^{3N}$ and every $v \in U$, one has

$$V^\delta(t, x) - C\delta \leq (T - t) \mathcal{L}^v V^\delta(s, y) + g'(z) - V^\delta(T, z) + V^\delta(t, x).$$

Hence,

$$V^\delta(t, x) - C\delta \leq \eta_{g'}(t, x).$$

The inequality (36) follows by passing to the limit as $\delta \rightarrow 0$ and recalling that Proposition 19 (ii) holds true. The proof of our Theorem is now complete. ■

Acknowledgments The author would like to thank the referee for valuable and very constructive comments which helped to improve the paper. The author was supported in part by the European project SADCO - Sensitivity Analysis for Deterministic Controller Design, (Marie Curie ITN-264735) and by the ANR Jeudy: ANR-10-BLAN 0112.

References

- [1] J.P. Aubin. *Viability theory*. Birkhäuser, Boston, 1991.
- [2] M. Bardi and I. Capuzzo Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1997.
- [3] A.G. Bhatt and V.S. Borkar. Occupation measures for controlled markov processes: Characterization and optimality. *Ann. of Probability*, 24:1531–1562, 1996.
- [4] V. Borkar and V. Gaitsgory. Averaging of singularly perturbed controlled stochastic differential equations. *Appl. Math. Optimization*, 56(2):169–209, 2007.
- [5] R. Buckdahn, D. Goreac, and M. Quincampoix. Stochastic optimal control and linear programming approach. *Appl. Math. Optimization*, 63(2):257–276, 2011.
- [6] L. Finlay, V. Gaitsgory, and I. Lebedev. Linear programming solutions of periodic optimization problems: approximation of the optimal control. *J. Ind. Manag. Optim.*, 3(2):399–413, 2007.
- [7] W. Fleming and D. Vermes. Convex duality approach to the optimal control of diffusions. *SIAM J. Control Optimization*, 36(2):1136–1155, 1989.
- [8] E. Friedman. Exact slow-fast decomposition of a class of non-linear singularly perturbed optimal control problems via invariant manifolds. *International Journal of Control*, 72:16091618, 1999.
- [9] V. Gaitsgory. Suboptimization of singularly perturbed control systems. *SIAM J. Control and Optim.*, 30(5):1228–1249., 1992.
- [10] V. Gaitsgory. On a representation of the limit occupational measures set of a control system with applications to singularly perturbed control systems. *SIAM J. Control Optim.*, 43(1):325–340, 2004.
- [11] V. Gaitsgory and A. Leizarowitz. Limit occupational measures set for a control system and averaging of singularity perturbed control systems. *J. Math. Anal. Appl.*, 233(2):461–475, 1999.
- [12] V. Gaitsgory and M.-T. Nguyen. Multiscale singularly perturbed control systems: Limit occupational measures sets and averaging. *SIAM J. Control Optimization*, 41(3):954–974, 2002.
- [13] V. Gaitsgory and M. Quincampoix. Linear programming approach to deterministic infinite horizon optimal control problems with discounting. *SIAM J. Control Optimization*, 48(4):2480–2512, 2009.
- [14] V. Gaitsgory and S. Rossomakhine. Linear programming approach to deterministic long run average problems of optimal control. *SIAM J. Control Optim.*, 44(6):2006–2037, 2006.
- [15] D. Goreac and O.S. Serea. Mayer and optimal stopping stochastic control problems with discontinuous cost. *Journal of Mathematical Analysis and Applications*, 380(1):327–342, 2011.
- [16] G. Grammel. Singularly perturbed differential inclusions. an averaging approach. *Set-valued Analysis*, 4(4):361–374, 1996.
- [17] G. Grammel. Averaging of singularly perturbed systems. *Averaging of singularly perturbed systems*, 28(11):1851–1865, 1997.

- [18] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probab. Theory Related Fields*, 117(1):1–16, 2000.
- [19] J. Macki and A. Strauss. *Introduction to optimal control theory*. Springer, New York, 1982.
- [20] D. Naidu. Singular perturbations and time scales in control theory and applications: an overview. *Dynamics of Continuous, Discrete and Impulsive Systems, Series B*, 24:233–278, 2002.
- [21] M. Quincampoix and F. Watbled. Averaging method for discontinuous mayer's problem of singularly perturbed control systems. *Nonlinear Anal., Theory Methods Appl.*, 54(5(A)):819–837, 2003.
- [22] W.J. Terrell. Some fundamental control theory i: Controllability, observability, and duality. *The American Mathematical Monthly*, 106:705–719, 1999.
- [23] A.N. Tichonov. Systems of differential equations containing small parameter near derivatives. *Mat. Sbornik.*, 31:575–586, 1952.
- [24] V. Veliov. A generalization of the tikhonov theorem for singularly perturbed differential inclusions. *J. Dynam. Control Systems*, 3(3):291–319, 1997.