

MULTISCALE SINGULARLY PERTURBED CONTROL SYSTEMS: LIMIT OCCUPATIONAL MEASURES SETS AND AVERAGING*

VLADIMIR GAITSGORY[†] AND MINH-TUAN NGUYEN^{†‡}

Abstract. An averaging technique for nonlinear multiscale singularly perturbed control systems is developed. Issues concerning the existence and structure of limit occupational measures sets generated by such systems are discussed. General results are illustrated with special cases.

Key words. multiscale singularly perturbed control systems, occupational measures, averaging method, limit occupational measures sets, nonlinear control, approximation of slow motions

AMS subject classifications. 34E15, 34C29, 34A60, 93C70, 34A4

PII. S0363012901393055

1. Introduction. In this paper we consider a singularly perturbed control system containing several small parameters $\epsilon_1, \dots, \epsilon_m$ ($m \geq 1$). The parameters are introduced in such a way that the state variables of the system are decomposed into a group of “slow” variables which change their values with the rates of the order $O(1)$ and m groups of “fast” variables which change their values with the rates of the orders $O(\epsilon_1^{-1}), O(\epsilon_1^{-1}\epsilon_2^{-1}), \dots, O(\epsilon_1^{-1}\epsilon_2^{-1} \dots \epsilon_m^{-1})$, respectively.

The main contribution of the paper is the description of the structure of the limit control system, the solutions of which allow us to approximate the slow variables when the parameters ϵ_i , $i = 1, \dots, m$, tend to zero. The role of controls in the limit system is played by probability measures defined on the product of the original control set and a subset of the state space containing all the fast trajectories (both are assumed to be compact). These probability measures are chosen from a limit set of occupational measures generated by the admissible controls and trajectories of an associated system which describes the dynamics of the fast variables if the slow ones are “frozen” (see exact definitions below). The existence of such a set (called limit occupational measures set (LOMS)) and its structure are the central issues discussed in the paper.

Singularly perturbed control systems (SPCS) with one small parameter ($m = 1$) have been intensively studied in the literature, the most common approaches being related either to Tikhonov-type theorems justifying the equating of the small parameter to zero with further application of the boundary layer method (see [24], [30]) to asymptotically describe the fast dynamics (see, e.g., [13], [21], [22], [25], [28], [31]) or to different types of averaging techniques (see [1], [2], [3], [4], [5], [8], [11], [14], [15], [16], [17], [18], [19], [20], [27], [32]) which allow us to deal with the situation when the equating of the parameter to zero may not lead to a right approximation.

The literature on multiscale SPCS ($m > 1$) is much less intensive. Most available references concern linear control systems (see, e.g., [12], [26], and references therein).

*Received by the editors July 30, 2001; accepted for publication (in revised form) February 25, 2002; published electronically September 19, 2002. This work was supported by Australian Research Council grant A49906132.

<http://www.siam.org/journals/sicon/41-3/39305.html>

[†]University of South Australia, School of Mathematics, The Mawson Lakes Campus, Mawson Lakes SA 5095, Australia (v.gaitsgory@unisa.edu.au).

[‡]Present address: Joint Systems Branch, DSTO, P.O. Box 1500, Edinburgh SA 5111, Australia (minh@linus.levels.unisa.edu.au).

A technique of averaging type applicable to nonlinear control systems having a triangular structure (weakly coupled) was proposed in [20].

In [18] an averaging technique allowing us to deal with a general form of SPCS containing two small parameters ($m = 2$) was developed. The extension of the technique to the case $m > 2$ is, however, hardly possible since it involves a multiple averaging over time and leads to really complex expressions which are difficult to comprehend. In this paper, an averaging over time is replaced by averaging over measures from the LOMS. It resembles approaches used in dealing with stochastic SPCS (see, e.g., [9], [23], [34]) and makes the transition from the case $m = k$ to the case $m = k + 1$ ($\forall k = 1, 2, \dots$) very natural.

Different issues related to averaging over occupational measures in SPCS with one small parameter were discussed in [2], [3], [4], [5], [17], [32]. In [17], in particular, LOMS for control systems without small parameters were considered. In this paper, we introduce and study such sets for singularly perturbed control systems (as is the associated system if the original system is multiscale).

The paper is organized as follows. Section 1 is this introduction. In section 2 statements about approximation of the slow motions by the solutions of the averaged system are formulated under the assumption that the LOMS of the associated system exists. An application of these results to problems of optimal control is demonstrated and a special case concerning systems linear in fast variables and controls is considered. In section 3 issues of existence and structure of the LOMS are addressed and a multistage averaging procedure for the construction of the LOMS is presented. The procedure is then illustrated with a special case of control systems which have a triangular structure (similar to those studied in [20]). Proofs of most of the statements are provided in section 4.

2. Averaging of multiscale SPCS.

2.1. Preliminaries. Given a compact metric space W , $\mathcal{B}(W)$ will stand for the σ -algebra of its Borel subsets and $\mathcal{P}(W)$ will denote the set of probability measures defined on $\mathcal{B}(W)$. The set $\mathcal{P}(W)$ will always be treated as a compact metric space with a metric ρ , which is consistent with its weak convergence topology. That is, a sequence $\gamma^k \in \mathcal{P}(W)$, $k = 1, 2, \dots$, converges to $\gamma \in \mathcal{P}(W)$ in this metric if and only if

$$\lim_{k \rightarrow \infty} \int_W \phi(w) \gamma^k(dw) = \int_W \phi(w) \gamma(dw)$$

for any continuous $\phi(w) : W \rightarrow \mathbb{R}^1$.

Using the metric ρ , one can define the Hausdorff metric ρ_H on the set of subsets of $\mathcal{P}(W)$:

$$(2.1) \quad \rho_H(\Gamma_1, \Gamma_2) \stackrel{\text{def}}{=} \max \left\{ \sup_{\gamma \in \Gamma_1} \rho(\gamma, \Gamma_2), \sup_{\gamma \in \Gamma_2} \rho(\gamma, \Gamma_1) \right\} \quad \forall \Gamma_1, \Gamma_2 \in \mathcal{P}(W),$$

$$\text{where} \quad \rho(\gamma, \Gamma_i) \stackrel{\text{def}}{=} \inf_{\gamma' \in \Gamma_i} \rho(\gamma, \gamma'), \quad i = 1, 2.$$

We will deal with the convergence in the Hausdorff metric of sets in $\mathcal{P}(W)$ defined as unions of occupational measures. Given a measurable function $w(t) : [0, T] \rightarrow W$, the occupational measure $p^{w(\cdot)} \in \mathcal{P}(W)$ generated by this function is defined by taking

$$(2.2) \quad p^{w(\cdot)}(Q) \stackrel{\text{def}}{=} \frac{1}{T} \text{meas} \left\{ t \mid w(t) \in Q \right\} \quad \forall Q \in \mathcal{B}(W),$$

where $\text{meas} \{ \cdot \}$ stands for the Lebesgue measure on $[0, T]$.

2.2. Setting. Consider the SPCS

$$\begin{aligned}
 \epsilon_1 \epsilon_2 \dots \epsilon_{m-1} \epsilon_m \dot{y}_1(t) &= f_1(u(t), y_1(t), \dots, y_m(t), z(t)), \\
 &\vdots \quad \vdots \quad \vdots \\
 \epsilon_{m-1} \epsilon_m \dot{y}_{m-1}(t) &= f_{m-1}(u(t), y_1(t), \dots, y_m(t), z(t)), \\
 \epsilon_m \dot{y}_m(t) &= f_m(u(t), y_1(t), \dots, y_m(t), z(t)), \\
 \dot{z}(t) &= g(u(t), y_1(t), \dots, y_m(t), z(t)),
 \end{aligned}
 \tag{2.3}$$

where $\epsilon \stackrel{\text{def}}{=} (\epsilon_1, \epsilon_2, \dots, \epsilon_m)$ is a vector of small positive parameters, $t \in [0, T]$, and the functions $f_i : U \times \mathbb{R}^{M_1} \times \dots \times \mathbb{R}^{M_m} \times \mathbb{R}^N \rightarrow \mathbb{R}^{M_i}$, $i = 1, \dots, m$, and $g : U \times \mathbb{R}^{M_1} \times \dots \times \mathbb{R}^{M_m} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ are continuous and satisfy Lipschitz conditions in (y_1, \dots, y_m, z) . Admissible controls are Lebesgue measurable functions $u(t) : [0, T] \rightarrow U$, where U is a compact metric space.

Consider also the system

$$\begin{aligned}
 \epsilon_1 \epsilon_2 \dots \epsilon_{m-1} \dot{y}_1(\tau) &= f_1(u(\tau), y_1(\tau), \dots, y_m(\tau), z), \\
 &\vdots \quad \vdots \quad \vdots \\
 \epsilon_{m-1} \dot{y}_{m-1}(\tau) &= f_{m-1}(u(\tau), y_1(\tau), \dots, y_m(\tau), z), \\
 \dot{y}_m(\tau) &= f_m(u(\tau), y_1(\tau), \dots, y_m(\tau), z), \\
 z &= \text{constant},
 \end{aligned}
 \tag{2.4}$$

in which z is fixed and $\tau \in [0, S]$. This system will be referred to as an *associated system* with respect to SPCS (2.3). It is formally obtained from the “fast” subsystem of (2.3) via the replacement of the time scale $\tau = t\epsilon_m^{-1}$. Admissible controls for the associated system (2.4) are Lebesgue measurable functions $u(\tau) : [0, S] \rightarrow U$. The solutions of (2.3) and (2.4) which are obtained with admissible controls are called admissible trajectories.

ASSUMPTION 2.1. (i) *There exist compact sets $Y_i'' \subseteq Y_i' \subset \mathbb{R}^{M_i}$, $i = 1, \dots, m$, and $Z'' \subseteq Z' \subset \mathbb{R}^N$ such that the admissible trajectories of SPCS (2.3) which satisfy the initial conditions*

$$(y_1(0), \dots, y_m(0), z(0)) \in Y_1'' \times \dots \times Y_m'' \times Z''
 \tag{2.5}$$

do not leave the set $Y_1' \times \dots \times Y_m' \times Z'$ on the interval $[0, T]$.

(ii) *There exist compact sets Y_i ($Y_i' \subseteq Y_i$), $i = 1, \dots, m$, and Z ($Z' \in \text{int}Z$) such that for any z from Z , the admissible trajectories of system (2.4) which satisfy the initial conditions*

$$(y_1(0), \dots, y_m(0)) \in Y_1' \times \dots \times Y_m'
 \tag{2.6}$$

do not leave the set $Y_1 \times \dots \times Y_m$ on the interval $[0, \infty)$.

Note that to verify this assumption, one can use results from viability theory (see Chapter 5 in [6] and also [29] for further references).

Let us introduce the following notation:

$$y(\tau) \stackrel{\text{def}}{=} (y_1(\tau), \dots, y_m(\tau)), \quad Y \stackrel{\text{def}}{=} Y_1 \times \dots \times Y_m,$$

and also $Y' \stackrel{\text{def}}{=} Y_1' \times \dots \times Y_m'$, $Y'' \stackrel{\text{def}}{=} Y_1'' \times \dots \times Y_m''$.

Let $u(\tau)$ be an admissible control defined on the interval $[0, S]$ and let $y(\tau)$ be the solution of the associated system (2.4) obtained with this control and the initial conditions (2.6). Let $p^{(u(\cdot), y(\cdot))} \in \mathcal{P}(U \times Y)$ be the occupational measure generated by the pair $(u(\tau), y(\tau)) : [0, S] \rightarrow U \times Y$ and let

$$(2.7) \quad \Gamma(z, \epsilon_1, \dots, \epsilon_{m-1}, S, y(0)) \stackrel{\text{def}}{=} \bigcup_{(u(\cdot), y(\cdot))} \{p^{(u(\cdot), y(\cdot))}\},$$

where the union is taken over all admissible controls and the corresponding solutions of (2.4). Notice that the dependence on $(z, \epsilon_1, \dots, \epsilon_{m-1})$ in (2.7) is due to the dependence of the solutions of (2.4) on these parameters.

ASSUMPTION 2.2. For any $z \in Z$, there exists a convex and compact set $\Gamma(z) \subset \mathcal{P}(U \times Y)$ such that

$$(2.8) \quad \rho_H(\Gamma(z, \epsilon_1, \dots, \epsilon_{m-1}, S, y(0)), \Gamma(z)) \leq \nu(\epsilon_1, \dots, \epsilon_{m-1}, S) \quad \forall y(0) \in Y',$$

where $\lim_{(\epsilon_1, \dots, \epsilon_{m-1}, S) \rightarrow 0} \nu(\epsilon_1, \dots, \epsilon_{m-1}, S) = 0$.

The set $\Gamma(z)$ introduced in Assumption 2.2 will be referred to as the *limit occupational measures set* (LOMS). Some sufficient conditions for the existence of the LOMS are considered in section 3.

ASSUMPTION 2.3. For any $S > 0$, any absolutely continuous function $\tilde{z}(\tau) : [0, S] \rightarrow Z$, and any admissible control $u(\tau) : [0, S] \rightarrow U$,

$$(2.9) \quad \max_{\tau \in [0, S]} \|y^z(\tau) - \tilde{y}(\tau)\| \leq c \max_{\tau \in [0, S]} \|z - \tilde{z}(\tau)\| + \kappa(\epsilon_1, \dots, \epsilon_{m-1}), \quad c = \text{const},$$

where $y^z(\tau)$ is the solution of (2.4) obtained with a given $z \in Z$ and $\tilde{y}(\tau)$ is the solution of the same system obtained with the replacement of z by the function $\tilde{z}(\tau)$. Initial conditions for $y^z(\tau)$ and $\tilde{y}(\tau)$ are the same: $y^z(0) = \tilde{y}(0) \in Y'$ and the function $\kappa(\epsilon_1, \dots, \epsilon_{m-1})$ is either zero (for $m = 1$) or tends to zero as $(\epsilon_1, \dots, \epsilon_{m-1})$ tends to zero (for $m > 1$).

LEMMA 2.4. Let Assumptions 2.1–2.3 be satisfied. Then for any vector function $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^j$, $j = 1, 2, \dots$, which is continuous in (u, y, z) and satisfies Lipschitz conditions in (y, z) , there exists a constant c_h such that

$$(2.10) \quad d_H(V_h(z'), V_h(z'')) \leq c_h \|z' - z''\| \quad \forall z', z'' \in Z,$$

where

$$(2.11) \quad V_h(z) \stackrel{\text{def}}{=} \bigcup_{p \in \Gamma(z)} \int_{U \times Y} h(u, y, z) p(du, dy).$$

Note that $d_H(\cdot, \cdot)$ in (2.10) stands for the Hausdorff metric in a finite-dimensional space. That is, for arbitrary bounded subsets V_1, V_2 of \mathbb{R}^j ($j = 1, 2, \dots$),

$$(2.12) \quad d_H(V_1, V_2) \stackrel{\text{def}}{=} \max \left\{ \sup_{v \in V_1} d(v, V_2), \sup_{v \in V_2} d(v, V_1) \right\}, \quad d(v, V_i) \stackrel{\text{def}}{=} \inf_{v' \in V_i} \|v - v'\|,$$

where $\|\cdot\|$ is a norm in \mathbb{R}^j .

The proof of Lemma 2.4 is in section 4.1.

Note that Assumption 2.3 is satisfied automatically if the functions f_1, \dots, f_{m-1} defining the right-hand side of the associated systems (2.4) do not depend on z . In a

general case, Assumption 2.3 can be verified to be valid if the associated system (2.4) satisfies stability conditions similar to that introduced in [16] (see [16, Assumption 4.1, Lemma 4.1]), the latter being implied by the existence of a Lyapunov-like function (as in [17, p. 467]). For the case $m = 1$ (one singular perturbation parameter), Assumption 2.3 can be replaced by the assumption that the statement of Lemma 2.4 is valid (see [17]). A slightly different assumption which can replace Assumption 2.3 for $m > 1$ is discussed in Remark 4.1.

2.3. Approximation of the slow trajectories. Let the function $\tilde{g}(\gamma, z) : \mathcal{P}(U \times Y) \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ be defined as follows:

$$(2.13) \quad \tilde{g}(\gamma, z) \stackrel{\text{def}}{=} \int_{U \times Y} g(u, y, z) \gamma(du, dy).$$

We will assume that the metric ρ of $\mathcal{P}(U \times Y)$ is chosen in such a way that the function $\tilde{g}(\gamma, z)$ satisfies the Lipschitz conditions:

$$(2.14) \quad \|\tilde{g}(\gamma', z') - \tilde{g}(\gamma'', z'')\| \leq b(\rho(\gamma', \gamma'') + \|z' - z''\|) \quad \forall z', z'', \forall \gamma', \gamma'',$$

where b is a positive constant. Let us consider the system

$$(2.15) \quad \dot{z}(t) = \tilde{g}(\gamma(t), z(t)),$$

which will be referred to as the *averaged system*. The role of controls in the averaged system is played by functions $\gamma(t)$ satisfying the inclusion

$$(2.16) \quad \gamma(t) \in \Gamma(z(t)).$$

Note that the fact that the functions $\gamma(t)$ are measure valued underlines the similarity of our description with classical relaxed control setting (see [33]).

DEFINITION 2.5. A pair $(\gamma(t), z(t)) : [0, T] \rightarrow \mathcal{P}(U \times Y) \times \mathbb{R}^N$ is called *admissible* for the averaged system if $\gamma(t)$ is Lebesgue measurable, $z(t)$ is absolutely continuous, and (2.15)–(2.16) are satisfied for almost all $t \in [0, T]$.

THEOREM 2.6. Let Assumptions 2.1–2.3 be satisfied and let $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^j$, $j = 1, 2, \dots$, be an arbitrary Lipschitz continuous vector function. There exist $\mu(\epsilon, T)$ and $\mu_h(\epsilon, T)$,

$$(2.17) \quad \lim_{\epsilon \rightarrow 0} \mu(\epsilon, T) = 0, \quad \lim_{\epsilon \rightarrow 0} \mu_h(\epsilon, T) = 0,$$

such that the following two statements are valid:

(i) Let $u(t)$ be an admissible control and let $(y(t), z(t))$ be the corresponding trajectory of SPCS (2.3) which satisfies initial condition (2.5). There exists an admissible pair $(\gamma^a(t), z^a(t))$ of the averaged system (2.15) with the initial conditions $z^a(0) = z(0)$ such that

$$(2.18) \quad \max_{t \in [0, T]} \|z(t) - z^a(t)\| \leq \mu(\epsilon, T),$$

and also

$$(2.19) \quad \left\| \int_0^T h(u(t), y(t), z(t)) dt - \int_0^T \tilde{h}(\gamma^a(t), z^a(t)) dt \right\| \leq \mu_h(\epsilon, T),$$

where

$$(2.20) \quad \tilde{h}(\gamma, z) \stackrel{\text{def}}{=} \int_{U \times Y} h(u, y, z) \gamma(du, dy).$$

(ii) Conversely, let $(\gamma^a(t), z^a(t))$ be an admissible pair of the averaged system (2.15), which satisfies initial conditions $z^a(0) \in Z''$. One can construct an admissible control $u(t)$ such that the trajectory $(y(t), z(t))$ of SPCS (2.3) obtained with this control and initial conditions (2.5) ($z(0) = z^a(0)$) will satisfy (2.18)–(2.19).

The proof of the theorem is in section 4.1. Estimates (2.18)–(2.19) of Theorem 2.6 are not uniform with respect to the length T of the time interval. Additional assumptions are needed to make them uniform. The assumption we use in this paper is as follows.

ASSUMPTION 2.7. *There exist positive definite matrices C, D and a constant a such that corresponding to any z', z'' from Z and any $\gamma' \in \Gamma(z')$ there exists $\gamma'' \in \Gamma(z'')$ such that*

$$(2.21) \quad (\tilde{g}(\gamma', z') - \tilde{g}(\gamma'', z''))^T C (z' - z'') \leq -\|z' - z''\|_D^2$$

and

$$(2.22) \quad \rho(\gamma', \gamma'') \leq a\|z' - z''\|,$$

where $\|x\|_D^2$ in (2.21) (and in what follows) stands for $x^T D x$.

Note that Assumption 2.7 is satisfied if the inequality (2.21) is valid for any $\gamma' = \gamma''$ and the LOMS $\Gamma(z)$ is independent of z (that is, the associated system does not depend on z).

THEOREM 2.8. *Let Assumptions 2.1–2.3 and 2.7 be satisfied. Assume also that all the admissible trajectories of averaged system (2.15) which start in Z'' do not leave Z' and those which start in Z' do not leave $\text{int}Z$ on the infinite time horizon. Then there exist $\mu(\epsilon)$ and $\mu_h(\epsilon)$,*

$$\lim_{\epsilon \rightarrow 0} \mu(\epsilon) = 0, \quad \lim_{\epsilon \rightarrow 0} \mu_h(\epsilon) = 0,$$

such that statements (i) and (ii) of Theorem 2.6 remain valid with

$$(2.23) \quad \sup_{t > 0} \|z(t) - z^a(t)\| \leq \mu(\epsilon)$$

replacing (2.18) and

$$(2.24) \quad \sup_{T > T_0} \left\| T^{-1} \int_0^T h(u(t), y(t), z(t)) dt - T^{-1} \int_0^T \tilde{h}(\gamma^a(t), z^a(t)) dt \right\| \leq \mu_h(\epsilon), \quad T_0 = \text{const}$$

replacing (2.19) for any Lipschitz continuous vector function $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^j$, $j = 1, 2, \dots$, such that the corresponding $\tilde{h}(\gamma, z)$ defined by (2.20) satisfies the Lipschitz condition

$$(2.25) \quad \|\tilde{h}(\gamma', z') - \tilde{h}(\gamma'', z'')\| \leq a_h(\rho(\gamma', \gamma'') + \|z' - z''\|) \quad \forall z', z'', \forall \gamma', \gamma'',$$

where a_h is some positive constant.

The proof of the theorem is in section 4.1.

2.4. Application to optimal control. Let $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^1$ be continuous and satisfy the Lipschitz conditions in (y, z) . Consider the optimal control problem

$$(2.26) \quad \inf_{(u(\cdot), y(\cdot), z(\cdot))} \left\{ \int_0^T h(u(t), y(t), z(t)) dt \right\},$$

where inf is sought over all admissible controls and trajectories of (2.3). Under the assumptions of Theorem 2.6, the optimal value of this problem converges to the optimal value of the problem

$$(2.27) \quad \inf_{(\gamma(\cdot), z(\cdot))} \left\{ \int_0^T \tilde{h}(\gamma(t), z(t)) dt \right\},$$

where $\tilde{h}(\gamma, z)$ is defined according to (2.20) and inf is over the admissible pairs of the averaged system (2.15). Near optimal controls of (2.26) can also be constructed on the basis of the solution of (2.27). These will be the controls which provide the validity of (2.18)–(2.19) for the admissible pair $(\gamma^a(t), z^a(t))$ which delivers the optimal (or near optimal) value to (2.27) (see statement (ii) of Theorem 2.6). If the assumptions of Theorem 2.8 are satisfied, then a similar approximation of a problem on the infinite time horizon with a time average criterion is possible.

In some cases the “limit” problem (2.27) can be significantly simplified with the help of the following proposition.

PROPOSITION 2.9. *Let $\phi(y_i) : Y_i \rightarrow \mathbb{R}^1$ be continuously differentiable. Then*

$$(2.28) \quad \int_{U \times Y} (\phi'(y_i))^T f_i(u, y, z) \gamma(du, dy) = 0 \quad \forall \gamma \in \Gamma(z),$$

and, in particular,

$$(2.29) \quad \int_{U \times Y} f_i(u, y, z) \gamma(du, dy) = 0 \quad \forall \gamma \in \Gamma(z),$$

where $f_i(u, y, z)$, $i = 1, \dots, m$, are the functions defining the right-hand side of (2.4).

The proof of the proposition is in section 4.1. To illustrate how this proposition can be applied let us consider the following special case. Assume that the set U is convex and the functions $f_i(u, y, z)$, $g(z, y, u)$ are linear in fast variables and controls. That is,

$$(2.30) \quad f_i(u, y, z) = \sum_{j=1}^m A_{i,j}(z) y_j + A_{i,m+1}(z) u + A_{i,m+2}(z), \quad i = 1, \dots, m,$$

$$(2.31) \quad g(u, y, z) = \sum_{j=1}^m A_{0,j}(z) y_j + A_{0,m+1}(z) u + A_{0,m+2}(z),$$

where $A_{i,j}$ are matrix functions of the corresponding dimensions. By (2.31), the averaged system is equivalent to

$$(2.32) \quad \dot{z}(t) = g(\bar{u}(t), \bar{y}(t), z(t)), \quad (\bar{u}(t), \bar{y}(t)) \in \Omega(z(t)),$$

where $\Omega(z)$ is the set of the first moments corresponding to the probability measures from the LOMS $\Gamma(z)$:

$$\Omega(z) \stackrel{\text{def}}{=} \left\{ (\bar{u}, \bar{y}) \mid (\bar{u}, \bar{y}) = \int_{Y \times U} (u, y) \gamma(du, dy), \quad \gamma \in \Gamma(z) \right\}.$$

By (2.29) and (2.30), this set allows the representation

$$(2.33) \quad \Omega(z) = \{(\bar{u}, \bar{y}) \mid f_i(\bar{u}, \bar{y}, z) = 0, \quad i = 1, \dots, m, \quad \bar{u} \in U\},$$

and thus (2.32) is equivalent to the control system

$$(2.34) \quad \dot{z}(t) = g(\bar{u}(t), \psi(\bar{u}(t), z(t)), z(t)), \quad \bar{u}(t) \in U,$$

where $\bar{y} = \psi(\bar{u}, z)$ is the root of the system of equations $f_i(\bar{u}, \bar{y}, z) = 0, i = 1, \dots, m$. This is a so-called reduced system and can be obtained from (2.3) via formally equating ϵ to zero. If, in addition, the function $h(u, y, z)$ used in (2.26) is convex in (u, y) , then limit problem (2.27) becomes equivalent to

$$(2.35) \quad \inf_{(\bar{u}(\cdot), z(\cdot))} \left\{ \int_0^T h(\bar{u}(t), \psi(\bar{u}(t), z(t)), z(t)) dt \right\},$$

where \inf is over the admissible controls and corresponding trajectories of (2.34). Notice that the reasoning above is valid if Assumptions 2.1–2.3 are satisfied. It can be shown (although it is quite technical and we do not prove it in this paper) that these assumptions are satisfied if the eigenvalues of the matrices $A_{i,l}^{(l-1)}(z), l = 1, \dots, m$, defined below have negative real parts for all z from a sufficiently large domain. The matrices are defined recursively for $l = 1, \dots, m$ by the equations

$$(2.36) \quad A_{i,j}^{(l)}(z) = A_{i,j}^{(l-1)}(z) - A_{i,l}^{(l-1)}(z)(A_{l,l}^{(l-1)}(z))^{-1}A_{l,j}^{(l-1)}(z)$$

($i = l + 1, \dots, m, j = l + 1, \dots, m + 2$), with $A_{i,j}^{(0)}(z) \stackrel{\text{def}}{=} A_{i,j}(z)$ ($i = 1, \dots, m, j = 1, \dots, m + 2$). Note that the condition that the matrices (2.36) have negative real parts is similar to that used in [12] to asymptotically describe the reachability set of a multiscale linear SPCS.

3. Existence of LOMS.

3.1. Approximation of the occupational measures set. Let $u(t)$ be an admissible control and let $(y(t), z(t))$ be the corresponding admissible trajectory of SPCS (2.3) which satisfies initial conditions (2.5). Let $p^{(u(\cdot), y(\cdot), z(\cdot))} \in \mathcal{P}(U \times Y \times Z)$ be the occupational measure generated by the vector function $(u(\cdot), y(\cdot), z(\cdot)): [0, T] \rightarrow U \times Y' \times Z' \subset U \times Y \times Z$ and let

$$(3.1) \quad \Gamma(\epsilon, T, y(0), z(0)) \stackrel{\text{def}}{=} \bigcup_{(u(\cdot), y(\cdot), z(\cdot))} \left\{ p^{(u(\cdot), y(\cdot), z(\cdot))} \right\},$$

where the union is taken over all admissible controls and the corresponding trajectories of SPCS (2.3). In this section, we will describe the asymptotics of this set as the vector of small parameters $\epsilon = (\epsilon_1, \dots, \epsilon_{m-1}, \epsilon_m)$ tends to zero.

Let $(\gamma(t), z(t)) : [0, T] \rightarrow \mathcal{P}(U \times Y) \times Z$ be an admissible pair of the averaged system (2.15) with the initial condition

$$(3.2) \quad z(0) \in Z''.$$

Let $p^{(\gamma(t), z(t))} \in \mathcal{P}(\mathcal{P}(U \times Y) \times Z)$ be the occupational measure generated by this pair and let $\tilde{\Gamma}(T, z(0))$ be the union of the occupational measures generated by all such pairs

$$(3.3) \quad \tilde{\Gamma}(T, z(0)) \stackrel{\text{def}}{=} \bigcup_{(\gamma(\cdot), z(\cdot))} \left\{ p^{(\gamma(\cdot), z(\cdot))} \right\}.$$

We will use $\tilde{\Gamma}(T, z(0))$ to specify the limit of (3.1) as ϵ tends to zero. To do that let us define a map $\psi(p) : p \in \mathcal{P}(\mathcal{P}(U \times Y) \times Z) \rightarrow \mathcal{P}(U \times Y \times Z)$ in such a way that for any $Q \in \mathcal{B}(U \times Y)$ and any $F \in \mathcal{B}(Z)$,

$$(3.4) \quad \psi(p)(Q \times F) = \int_{\mathcal{P}(U \times Y) \times Z} \gamma(Q) \chi_F(z) p(d\gamma, dz),$$

where $\chi_F(\cdot)$ is the indicator function of F . The integration in (3.4) is legitimate since the function

$$(3.5) \quad \gamma(Q) \chi_F(z) : (\gamma, z) \in \mathcal{P}(U \times Y) \times Z \rightarrow [0, 1]$$

is measurable with respect to $\mathcal{B}(\mathcal{P}(U \times Y) \times Z)$ (see [10, Proposition 7.25, p. 133]). Notice that for any $p \in \mathcal{P}(\mathcal{P}(U \times Y) \times Z)$ and any continuous function $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^j, j = 1, 2, \dots,$

$$(3.6) \quad \int_{U \times Y \times Z} h(u, y, z) \psi(p)(du, dy, dz) = \int_{\mathcal{P}(U \times Y) \times Z} \tilde{h}(\gamma, z) p(d\gamma, dz),$$

where $\tilde{h}(\gamma, z)$ is defined by (2.20). For $p = p^{(\gamma(\cdot), z(\cdot))}$ (that is, for p being the occupational measure generated by an admissible pair $(\gamma(\cdot), z(\cdot))$ of (2.15))

$$(3.7) \quad \int_{U \times Y \times Z} h(u, y, z) \psi\left(p^{(\gamma(\cdot), z(\cdot))}\right)(du, dy, dz) = \frac{1}{T} \int_0^T \tilde{h}(\gamma(t), z(t)) dt.$$

Let us now define the set $\Gamma(T, z(0)) \subset \mathcal{P}(U \times Y \times Z)$ as follows:

$$(3.8) \quad \Gamma(T, z(0)) \stackrel{\text{def}}{=} \bigcup_{p \in \tilde{\Gamma}(T, z(0))} \left\{ \psi(p) \right\} = \bigcup_{(\gamma(\cdot), z(\cdot))} \left\{ \psi\left(p^{(\gamma(\cdot), z(\cdot))}\right) \right\},$$

where the second union is taken over all admissible pairs of (2.15) satisfying initial conditions (3.2). (The second equality follows from the definition (3.3) of the set $\tilde{\Gamma}(T, z(0))$.)

THEOREM 3.1. (i) *Let the assumptions of Theorem 2.6 be satisfied. Then there exists $\nu(\epsilon, T), \lim_{\epsilon \rightarrow 0} \nu(\epsilon, T) = 0,$ such that*

$$(3.9) \quad \rho_{\text{H}}\left(\Gamma(\epsilon, T, y(0), z(0)), \Gamma(T, z(0))\right) \leq \nu(\epsilon, T) \quad \forall (y(0), z(0)) \in Y'' \times Z''.$$

(ii) *Let the assumptions of Theorem 2.8 be satisfied and let there be a sequence $q_k(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^1, k = 1, 2, \dots,$ of Lipschitz continuous functions such that it is dense in $C(U \times Y \times Z)$ and for any*

$$(3.10) \quad h(z, y, u) \stackrel{\text{def}}{=} (q_1(u, y, z), \dots, q_j(u, y, z)), \quad j = 1, 2, \dots,$$

in which both y_m and z are fixed. For the sake of convenience, in this section we will refer to (2.4) and (3.16) as to y_m - and y_{m-1} -associated systems, respectively (by the name of the group of variables changing their values with rates of the order $O(1)$). One can also consider y_{m-2}, \dots, y_2 - and y_1 -associated systems, the latter two being of the form

$$(3.17) \quad \begin{aligned} \epsilon_1 \dot{y}_1(\tau) &= f_1(u(\tau), y_1(\tau), y_2(\tau), y_3, \dots, y_m, z), \\ \dot{y}_2(\tau) &= f_2(u(\tau), y_1(\tau), y_2(\tau), y_3, \dots, y_m, z), \\ (y_3, \dots, y_m, z) &= \text{constant} \end{aligned}$$

and

$$(3.18) \quad \begin{aligned} \dot{y}_1(\tau) &= f_1(u(\tau), y_1(\tau), y_2, y_3, \dots, y_m, z), \\ (y_2, y_3, \dots, y_m, z) &= \text{constant}. \end{aligned}$$

Assume that the LOMS $\Gamma_1(y_2, y_3, \dots, y_m, z) \subset \mathcal{P}(U \times Y_1)$ of system (3.18) exists (sufficient conditions for the existence of LOMS of systems which, like (3.18), do not involve small parameters were discussed in [17]) and that Theorem 2.6 is applicable to system (3.17). Then y_2 -components of the trajectories of this system are approximated by the trajectories of the averaged system

$$(3.19) \quad \dot{y}_2(\tau) = \tilde{f}_2(\gamma_1(\tau), y_2(\tau), y_3, \dots, y_m, z), \quad \gamma_1(\tau) \in \Gamma_1(y_2(\tau), y_3, \dots, y_m, z),$$

where (y_3, \dots, y_m, z) are fixed and

$$(3.20) \quad \tilde{f}_2(\gamma_1, y_2, y_3, \dots, y_m, z) \stackrel{\text{def}}{=} \int_{U \times Y_1} f_2(u, y_1, y_2, y_3, \dots, y_m, z) \gamma_1(du, dy_1).$$

Suppose that the LOMS $\tilde{\Gamma}_2(y_3, \dots, y_m, z) \subset \mathcal{P}(\mathcal{P}(U \times Y_1) \times Y_2)$ of system (3.19) exists and that the other assumptions of Proposition 3.2 or Theorem 3.3 are satisfied. One then can come to the conclusion that the LOMS $\Gamma_2(y_3, \dots, y_m, z) \subset \mathcal{P}(U \times Y_1 \times Y_2)$ of system (3.17) exists and is presented in the form

$$(3.21) \quad \Gamma_2(y_3, \dots, y_m, z) = \bigcup_{p \in \tilde{\Gamma}_2(y_3, \dots, y_m, z)} \{ \psi_1(p) \},$$

where the map $\psi_1(p) : p \in \mathcal{P}(\mathcal{P}(U \times Y_1) \times Y_2) \rightarrow \mathcal{P}(U \times Y_1 \times Y_2)$ is such (compare with (3.4) above) that for any $Q \in \mathcal{B}(U \times Y_1)$ and any $F \in \mathcal{B}(Y_2)$,

$$(3.22) \quad \psi_1(p)(Q \times F) = \int_{\mathcal{P}(U \times Y_1) \times Y_2} \gamma_1(Q) \chi_F(y_2) p(d\gamma_1, dy_2),$$

$\chi_F(\cdot)$ being the indicator function of F . Assuming further that Proposition 3.2 or Theorem 3.3 can be applied step by step to y_3, \dots, y_m -associated systems, one can establish the existence of the LOMS $\Gamma(z) \stackrel{\text{def}}{=} \Gamma_m(z)$ of system (2.4), which is presented in the form

$$(3.23) \quad \Gamma_m(z) = \bigcup_{p \in \tilde{\Gamma}_m(z)} \{ \psi_{m-1}(p) \},$$

with the corresponding definition of $\psi_{m-1}(p)$ and $\tilde{\Gamma}_m(z)$ being the LOMS of the averaged system

$$(3.24) \quad \dot{y}_m(\tau) = \tilde{f}_m(\gamma_{m-1}(\tau), y_m(\tau), z), \quad \gamma_{m-1}(\tau) \in \Gamma_{m-1}(y_m(\tau), z),$$

where $z = \text{const}$, $\Gamma_{m-1}(y_m, z)$ is the LOMS of the y_{m-1} -associated system, and

$$\tilde{f}_m(\gamma_{m-1}, y_m, z) \stackrel{\text{def}}{=} \int_{U \times Y_1 \times \dots \times Y_{m-1}} f_m(u, y_1, \dots, y_{m-1}, y_m, z) \gamma_{m-1}(du, dy_1, \dots, dy_{m-1}).$$

The applicability of Theorem 3.3 to each of the above systems is easy to verify, for example, if

$$(3.25) \quad f_i(u, y, z) \stackrel{\text{def}}{=} f_i(u, y_1, \dots, y_i), \quad i = 1, \dots, m.$$

That is, the dynamics of y_i -components in (2.3) is not influenced by the dynamics of y_{i+1}, \dots, y_m and z -components. Assuming that this is the case, let us also introduce the following assumption about the functions $f_i(\cdot)$.

ASSUMPTION 3.4. *There exist positive definite matrices C_i, D_i ($i = 1, \dots, m$) such that for any $u \in U$ and any $y_1, \dots, y_{i-1}, y'_i, y''_i$,*

$$(3.26) \quad (f_i(u, y_1, \dots, y_{i-1}, y'_i) - f_i(u, y_1, \dots, y_{i-1}, y''_i))^T C_i (y'_i - y''_i) \leq -\|y'_i - y''_i\|_{D_i}^2.$$

By (3.25), the y_1 -associated system (3.18) does not depend on (y_2, \dots, y_m, z) and, by (3.26) with $i = 1$, the LOMS Γ_1 of this system exists (see Proposition 3.3 in [17]). Again, by (3.25), the dependence on (y_3, \dots, y_m, z) in the function (3.20) defining the right-hand side of (3.19) disappears and, by (3.26) with $i = 2$, this function satisfies the inequality

$$(\tilde{f}_2(\gamma_1, y'_2) - \tilde{f}_2(\gamma_1, y''_2))^T C_2 (y'_2 - y''_2) \leq -\|y'_2 - y''_2\|_{D_2}^2 \quad \forall \gamma_1 \in \mathcal{P}(U \times Y_1),$$

$\forall y'_2, y''_2 \in \mathbb{R}^{M_2}$ and $\forall \gamma_1 \in \mathcal{P}(U \times Y_1)$. This implies the applicability of Theorem 3.3 according to which the LOMS $\tilde{\Gamma}_2$ of averaged system (3.19) and the LOMS Γ_2 of the y_2 -associated system both exist and the representation (3.21) is valid. Continuing in a similar way, one can finally verify that the LOMS $\tilde{\Gamma}_m$ of averaged system (3.24) and the LOMS Γ_m of y_m -associated system (2.4) exist and that the representation (3.23) is valid. The applicability of Theorem 3.3 at this final stage can be verified by using the fact that the function $\tilde{f}_m(\gamma_{m-1}, y_m)$ defining the right-hand side of the averaged system (3.24) (which, by (3.25), does not involve the dependence on z) satisfies the inequality

$$(\tilde{f}_m(\gamma_{m-1}, y'_m) - \tilde{f}_m(\gamma_{m-1}, y''_m))^T C_m (y'_m - y''_m) \leq -\|y'_m - y''_m\|_{D_m}^2$$

$\forall y'_m, y''_m \in \mathbb{R}^{M_m}$ and $\forall \gamma_{m-1} \in \mathcal{P}(U \times Y_1 \times \dots \times Y_{m-1})$.

Note that a different multistage averaging procedure for SPCS with $f_i(\cdot)$ having the form (3.25) and satisfying an assumption similar to Assumption 3.4 (with C_i, D_i being identity matrices) was suggested in [20].

3.4. Basic lemma and the proof of Theorem 3.1. The proofs of Theorems 3.1 and 3.3 are based on the lemma and its corollaries presented below. Let W be a compact metric space and $q_k(w) : W \rightarrow \mathbb{R}^1, k = 1, 2, \dots$, be a sequence of Lipschitz continuous functions which is dense in $C(W)$.

LEMMA 3.5. *Let $\Gamma^i(\alpha, \beta) \subset \mathcal{P}(W), i = 1, 2$, where α and β take values in some metric spaces \mathcal{A} and \mathcal{B} . Assume that corresponding to any vector function*

$$(3.27) \quad h(w) = (q_1(w), \dots, q_j(w)), \quad j = 1, 2, \dots,$$

there exists a function

$$(3.28) \quad \nu_h(\alpha) : \mathcal{A} \rightarrow \mathbb{R}^1, \quad \lim_{\alpha \rightarrow \alpha_0} \nu_h(\alpha) = 0,$$

such that

$$(3.29) \quad \sup_{v \in V_h^1(\alpha, \beta)} d(v, V_h^2(\alpha, \beta)) \leq \nu_h(\alpha),$$

where

$$(3.30) \quad V_h^i(\alpha, \beta) \stackrel{\text{def}}{=} \bigcup_{\gamma \in \Gamma^i(\alpha, \beta)} \left\{ \int_W h(w) \gamma(dw) \right\}, \quad i = 1, 2, \dots$$

Then there also exists another function

$$(3.31) \quad \nu(\alpha) : \mathcal{A} \rightarrow \mathbb{R}^1, \quad \lim_{\alpha \rightarrow \alpha_0} \nu(\alpha) = 0,$$

such that

$$(3.32) \quad \sup_{\gamma \in \Gamma^1(\alpha, \beta)} \rho(\gamma, \Gamma^2(\alpha, \beta)) \leq \nu(\alpha).$$

COROLLARY 3.6. *If for any $h(w) : W \rightarrow \mathbb{R}^j$ as in (3.27) there exists a function (3.28) such that*

$$(3.33) \quad d_H(V_h^1(\alpha, \beta), V_h^2(\alpha, \beta)) \leq \nu_h(\alpha),$$

then there also exists a function (3.31) such that

$$(3.34) \quad \rho_H(\Gamma^1(\alpha, \beta), \Gamma^2(\alpha, \beta)) \leq \nu(\alpha).$$

COROLLARY 3.7. *Let $\Gamma(\alpha, \beta) \subset \mathcal{P}(W)$ for $(\alpha, \beta) \in \mathcal{A} \times \mathcal{B}$, and for any $h(w) : W \rightarrow \mathbb{R}^j$ as in (3.27) there exists a convex and compact set $V_h \subset \mathbb{R}^j$ and a function (3.28) such that*

$$(3.35) \quad d_H(V_h(\alpha, \beta), V_h) \leq \nu_h(\alpha),$$

where

$$(3.36) \quad V_h(\alpha, \beta) = \bigcup_{\gamma \in \Gamma(\alpha, \beta)} \left\{ \int_W h(w) \gamma(dw) \right\}.$$

Then there exists a function (3.31) such that

$$(3.37) \quad \rho_H(\Gamma(\alpha, \beta), \Gamma) \leq \nu(\alpha),$$

where Γ is a convex and compact subset of $\mathcal{P}(W)$ defined by

$$(3.38) \quad \Gamma \stackrel{\text{def}}{=} \left\{ \gamma \mid \gamma \in \mathcal{P}(W), \int_W h(w) \gamma(dw) \in V_h \ \forall h(w) : W \rightarrow \mathbb{R}^j \text{ as in (3.27)} \right\}.$$

The proof of Lemma 3.5 is in section 4.2. Corollary 3.6 is implied by Lemma 3.5 in an obvious way. The proof of Corollary 3.7 is similar to the proof of Theorem 3.1(i) in [17].

Proof of Theorem 3.1. Let $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^j, j = 1, 2, \dots$, be an arbitrary Lipschitz continuous vector function. Let $u(t)$ be an admissible control and let $(y(t), z(t))$ be the corresponding admissible trajectory of SPCS (2.3) which satisfies initial conditions (2.5). Let $V_h(\epsilon, T, y(0), z(0))$ be the set of time averages

$$(3.39) \quad V_h(\epsilon, T, y(0), z(0)) \stackrel{\text{def}}{=} \bigcup_{(u(\cdot), y(\cdot), z(\cdot))} \left\{ \frac{1}{T} \int_0^T h(u(t), y(t), z(t)) dt \right\},$$

where the union is taken over all admissible controls and the corresponding trajectories of (2.3). Notice that by definition (3.1) of $\Gamma(\epsilon, T, y(0), z(0))$, the set (3.39) also allows the representation

$$(3.40) \quad V_h(\epsilon, T, y(0), z(0)) = \bigcup_{\gamma \in \Gamma(\epsilon, T, y(0), z(0))} \left\{ \int h(u, y, z) \gamma(du, dy, dz) \right\}.$$

Let the set $\tilde{V}_h(T, z(0))$ be defined as follows:

$$(3.41) \quad \begin{aligned} \tilde{V}_h(T, z(0)) &\stackrel{\text{def}}{=} \bigcup_{\gamma \in \Gamma(T, z(0))} \left\{ \int h(u, y, z) \gamma(du, dy, dz) \right\} \\ &= \bigcup_{(\gamma(\cdot), z(\cdot))} \left\{ \int h(u, y, z) \psi(p^{(\gamma(\cdot), z(\cdot))})(du, dy, dz) \right\}, \end{aligned}$$

where, as in (3.8), the second union is taken over all admissible pairs of (2.15) which satisfy the initial conditions (3.2).

By (3.7), the set $\tilde{V}_h(T, z(0))$ can also be represented in the form

$$(3.42) \quad \tilde{V}_h(T, z(0)) = \bigcup_{(\gamma(\cdot), z(\cdot))} \left\{ \frac{1}{T} \int_0^T \tilde{h}(\gamma(\cdot), z(\cdot)) \right\}.$$

Using estimate (2.19) from Theorem 2.6 and comparing (3.39) and (3.42), one obtains

$$(3.43) \quad d_H(V_h(\epsilon, T, y(0), z(0)), \tilde{V}_h(T, z(0))) \leq \frac{1}{T} \mu_h(\epsilon, T)$$

$\forall (y(0), z(0)) \in Y'' \times Z''$. Having in mind representations (3.40), (3.41) and applying Corollary 3.6, one proves (3.9). Under the conditions of Theorem 2.8, estimate (3.43) can be rewritten in the uniform with respect to the $T \geq T_0$ form

$$d_H(V_h(\epsilon, T, y(0), z(0)), \tilde{V}_h(T, z(0))) \leq \mu_h(\epsilon) \quad \forall T \geq T_0,$$

where $h(\cdot)$ is as in (3.10). This, by Corollary 3.6, proves (3.11). \square

4. Proofs and auxiliary results.

4.1. Proofs for section 2.

Proof of Lemma 2.4. Consider the set of the time averages

$$(4.1) \quad V_h(z, \bar{\epsilon}, S, y(0)) \stackrel{\text{def}}{=} \bigcup_{(u(\cdot), y(\cdot))} \left\{ \frac{1}{S} \int_0^S h(u(\tau), y(\tau), z) \right\},$$

where $\bar{\epsilon} \stackrel{\text{def}}{=} (\epsilon_1, \dots, \epsilon_{m-1})$ and the union is taken over all admissible controls and the corresponding trajectories of (2.4). By Assumption 2.3,

$$(4.2) \quad \max_{\tau \in [0, S]} \|y^{z'}(\tau) - y^{z''}(\tau)\| \leq c\|z' - z''\| + \kappa(\bar{\epsilon}) \quad \forall z', z'' \in Z,$$

where $y^{z'}(\tau)$ and $y^{z''}(\tau)$ are solutions of (2.4) obtained with the same control and initial conditions and with $z = z'$ and $z = z''$, respectively. Hence,

$$(4.3) \quad d_H\left(V_h(z', \bar{\epsilon}, S, y(0)), V_h(z'', \bar{\epsilon}, S, y(0))\right) \leq c_h\|z' - z''\| + c_h\kappa(\bar{\epsilon}) \quad \forall z', z'' \in Z,$$

where c_h is a constant which is expressed via the Lipschitz constant of $h(\cdot)$ and c from (4.2) in an obvious way.

By definition (2.7) of $\Gamma(z, \bar{\epsilon}, S, y(0))$, the set $V_h(z, \bar{\epsilon}, S, y(0))$ defined in (4.1) allows also the representation

$$(4.4) \quad V_h(z, \bar{\epsilon}, S, y(0)) = \bigcup_{p \in \Gamma(z, \bar{\epsilon}, S, y(0))} \left\{ \int_{U \times Y} h(u, y)p(du, dy) \right\}.$$

It follows from Assumption 2.2 that there exists a function $\nu_h(\bar{\epsilon}, S)$ such that

$$\lim_{(\bar{\epsilon}, S^{-1}) \rightarrow 0} \nu_h(\bar{\epsilon}, S) = 0$$

and

$$(4.5) \quad d_H\left(V_h(z, \bar{\epsilon}, S, y(0)), V_h(z)\right) \leq \nu_h(\bar{\epsilon}, S) \quad \forall z \in Z, \forall y(0) \in Y'.$$

Passing to the limit as $(\bar{\epsilon}, S^{-1})$ tends to zero in (4.3), one obtains (2.10). \square

Proof of Theorem 2.6. Let $\bar{g}(u, y, z) \stackrel{\text{def}}{=} (g(u, y, z), h(u, y, z))$. Consider the set of time averages

$$V(z, \bar{\epsilon}, S, y(0)) = \bigcup_{(u(\cdot), y(\cdot))} \left\{ \frac{1}{S} \int_0^T \bar{g}(u(\tau), y(\tau), z) d\tau \right\} \subset \mathbb{R}^{N+j},$$

where, as in (4.1), the union is taken over all admissible controls and corresponding trajectories of (2.4). From Assumption 2.2 it follows (similarly to (4.5)) that there exists $\bar{\nu}(\bar{\epsilon}, S)$, $\lim_{(\bar{\epsilon}, S^{-1}) \rightarrow 0} \bar{\nu}(\bar{\epsilon}, S) = 0$ such that

$$(4.6) \quad d_H\left(V(z, \bar{\epsilon}, S, y(0)), V(z)\right) \leq \bar{\nu}(\bar{\epsilon}, S) \quad \forall z \in Z, \forall y(0) \in Y',$$

where

$$(4.7) \quad V(z) \stackrel{\text{def}}{=} \left\{ (v, w) \mid (v, w) = (\tilde{g}(\gamma, z), \tilde{h}(\gamma, z)), \gamma \in \Gamma(z) \right\} \subset \mathbb{R}^{N+j},$$

with \tilde{g} and \tilde{h} being defined by (2.13) and (2.20), respectively.

Let us augment the averaged system (2.15) with the equation

$$(4.8) \quad \dot{\theta}(t) = \tilde{h}(\gamma(t), z(t)), \quad \theta(0) = 0.$$

The map $V(z): Z \rightarrow 2^{\mathbb{R}^{N+j}}$ defined by (4.7) is convex and compact valued. It also satisfies the Lipschitz conditions (Lemma 2.4)

$$(4.9) \quad d_H(V(z'), V(z'')) \leq \bar{c}\|z' - z''\| \quad \forall z', z'' \in Z, \quad \bar{c} = \text{const.}$$

By the Filippov theorem (see, e.g., [7, Theorem 8.2.10, p. 316]), the set of admissible trajectories $(z(t), \theta(t)) \stackrel{\text{def}}{=} \bar{z}(t)$ of systems (2.15) and (4.8) coincides with the set of solutions of the differential inclusion

$$(4.10) \quad \dot{\bar{z}}(t) \in V(z(t)).$$

Let us augment system (2.3) with the equation

$$(4.11) \quad \dot{\theta}(t) = h(u(t), y_1(t), \dots, y_m(t), z(t)), \quad \theta(0) = 0,$$

and again denote $\bar{z}(\tau) \stackrel{\text{def}}{=} (z(\tau), \theta(\tau))$. To prove the theorem it is enough to show that, corresponding to any admissible trajectory $(y(t), \bar{z}(t))$ of (2.3) and (4.11), there exists a solution $\bar{z}^a(t)$ of (4.10) satisfying the inequality

$$(4.12) \quad \max_{t \in [0, T]} \|\bar{z}(t) - \bar{z}^a(t)\| \leq \bar{\mu}(\epsilon, T), \quad \lim_{\epsilon \rightarrow 0} \bar{\mu}(\epsilon, T) = 0,$$

and, conversely, for any solution $\bar{z}^a(t)$ of (4.10), there exists an admissible trajectory $(y(\tau), \bar{z}(\tau))$ of (2.3) and (4.11) which satisfies (4.12).

The proof of these statements is similar to Lemma 2.1 in [16] or Theorem 3.1 in [19]. \square

REMARK 4.1. *Note that an important step of the proof is an introduction of the new time scale $\tau \stackrel{\text{def}}{=} t\epsilon_m^{-1}$ and a partition of the interval $[0, T\epsilon_m^{-1}]$ by the points $\tau_l = lS(\epsilon_m)$, $l = 0, 1, \dots$, where $S(\epsilon_m) > 0$ is a function of ϵ_m such that $\lim_{\epsilon_m \rightarrow 0} S(\epsilon_m) = \infty$ and $\lim_{\epsilon_m \rightarrow 0} \epsilon_m S(\epsilon_m) = 0$. At the cost of making the proof slightly more involved, one can replace Assumption 2.3 by the assumption that the statement of Lemma 2.4 is valid and that the y_{m-1} -associated system (3.16) has a property similar to (2.9), with (y_m, z) playing the role of z .*

Proof of Theorem 2.8. The proof is based on the following result.

PROPOSITION 4.2. *Given a solution $(z^1(t), \theta^1(t))$ of the differential inclusion (4.10) satisfying the initial condition $(z^1(0), \theta^1(0)) = (z^1, \theta^1) \in Z' \times \mathbb{R}^j$ and a vector $(z^2, \theta^2) \in Z' \times \mathbb{R}^j$, there exists a solution $(z^2(t), \theta^2(t))$ of (4.10) which satisfies the initial condition $(z^2(0), \theta^2(0)) = (z^2, \theta^2)$, and the following inequalities hold:*

$$(4.13) \quad \|z^1(t) - z^2(t)\| \leq b_1 e^{-\beta t} \|z^1 - z^2\|,$$

$$(4.14) \quad \|\theta^1(t) - \theta^2(t)\| \leq \|\theta^1 - \theta^2\| + b_2 \|z^1 - z^2\|,$$

where b_1, b_2, β are some positive constants.

Proof of Proposition 4.2. As mentioned above, the map $V(z)$ defined in (4.7) is convex and compact valued and satisfies Lipschitz conditions. Also, from Assumption 2.7 (see (2.21)–(2.22)) it follows that it has the following property: for any $z' \in Z$, $(v', w') \in V(z')$ and any $z'' \in Z$, there exists $(v'', w'') \in V(z'')$ such that

$$(4.15) \quad (v' - v'')^T C(z' - z'') \leq -\|z' - z''\|_D^2,$$

$$(4.16) \quad \|w' - w''\| \leq b_h \|z' - z''\|.$$

The claim of the proposition follows now from Lemma A.2 in [18]. \square

To prove Theorem 2.8 let us choose T_0 in such a way that

$$(4.17) \quad b_1 e^{-\beta T_0} \stackrel{\text{def}}{=} \delta < 1,$$

and let $(y(t), z(t), \theta(t))$ be an admissible trajectory of the systems (2.3) and (4.11) which satisfies the initial conditions $(y(0), z(0)) \in Y'' \times Z''$, $\theta(0) = 0$. By Theorem 2.6, there exists a solution $(z^a(t), \theta^a(t))$ of the differential inclusion (4.10) satisfying the initial condition $(z^a(0), \theta^a(0)) = (z(0), 0)$ such that

$$(4.18) \quad \|z(t) - z^a(t)\| \leq \mu(\epsilon, T_0), \quad \|\theta(t) - \theta^a(t)\| \leq \mu_h(\epsilon, T_0) \quad \forall t \in [0, T_0].$$

When Theorem 2.6 is applied again, one can establish that there exists a solution $(\tilde{z}^a(t), \tilde{\theta}^a(t))$ of (4.10) on the interval $[T_0, 2T_0]$ such that it satisfies the initial conditions $(\tilde{z}^a(T_0), \tilde{\theta}^a(T_0)) = (z(T_0), \theta(T_0))$ and, also, such that the following estimates are valid:

$$(4.19) \quad \|z(t) - \tilde{z}^a(t)\| \leq \mu(\epsilon, T_0), \quad \|\theta(t) - \tilde{\theta}^a(t)\| \leq \mu_h(\epsilon, T_0) \quad \forall t \in [T_0, 2T_0].$$

It follows from Proposition 4.2 that the solution $(z^a(t), \theta^a(t))$ used in (4.18) can be extended to the interval $[T_0, 2T_0]$ in such a way that for any $t \in [T_0, 2T_0]$,

$$\begin{aligned} \|\tilde{z}^a(t) - z^a(t)\| &\leq b_1 e^{-\beta(t-T_0)} \|z(T_0) - z^a(T_0)\|, \\ \|\tilde{\theta}^a(t) - \theta^a(t)\| &\leq \|\theta(T_0) - \theta^a(T_0)\| + b_2 \|z(T_0) - z^a(T_0)\|. \end{aligned}$$

These along with (4.19) allow us to establish that for any $t \in [T_0, 2T_0]$,

$$\begin{aligned} \|z(t) - z^a(t)\| &\leq \mu(\epsilon, T_0) + b_1 e^{-\beta(t-T_0)} \|z(T_0) - z^a(T_0)\|, \\ \|\theta(t) - \theta^a(t)\| &\leq \mu_h(\epsilon, T_0) + \|\theta(T_0) - \theta^a(T_0)\| + b_2 \|z(T_0) - z^a(T_0)\|. \end{aligned}$$

Continuing in a similar fashion, one can construct a solution of (4.10) such that the following inequalities are satisfied $\forall t \in [lT_0, (l+1)T_0]$, $l = 1, 2, \dots$:

$$(4.20) \quad \|z(t) - z^a(t)\| \leq \mu(\epsilon, T_0) + b_1 e^{-\beta(t-lT_0)} \|z(lT_0) - z^a(lT_0)\|,$$

$$(4.21) \quad \|\theta(t) - \theta^a(t)\| \leq \mu_h(\epsilon, T_0) + \|\theta(lT_0) - \theta^a(lT_0)\| + b_2 \|z(lT_0) - z^a(lT_0)\|.$$

It follows now from (4.17) and (4.20)–(4.21) that

$$\begin{aligned} \|z((l+1)T_0) - z^a((l+1)T_0)\| &\leq \mu(\epsilon, T_0) + \delta \|z(lT_0) - z^a(lT_0)\|, \\ \|\theta((l+1)T_0) - \theta^a((l+1)T_0)\| &\leq \mu_h(\epsilon, T_0) + \|\theta(lT_0) - \theta^a(lT_0)\| + b_2 \|z(lT_0) - z^a(lT_0)\|, \end{aligned}$$

which imply that

$$\begin{aligned} \|z(lT_0) - z^a(lT_0)\| &\leq \frac{\mu(\epsilon, T_0)}{1 - \delta}, \quad l = 1, 2, \dots, \\ \|\theta(lT_0) - \theta^a(lT_0)\| &\leq l \left(\mu_h(\epsilon, T_0) + \frac{b_2}{1 - \delta} \mu(\epsilon, T_0) \right), \quad l = 1, 2, \dots \end{aligned}$$

These and (4.20)–(4.21) lead to statement (i) of the theorem (see also the proof of Lemma 3.2 in [18]). The proof of (ii) is similar. \square

Proof of Proposition 2.9. Let $\gamma \in \Gamma(z)$. By (2.8), there exist sequences $\bar{\epsilon}^k$, S^k , and $\gamma^k \in \Gamma(z, \bar{\epsilon}^k, S^k, y(0))$ such that $(\bar{\epsilon}^k, (S^k)^{-1}) \rightarrow 0$ and $\gamma^k \rightarrow \gamma$ as k tends to infinity. The latter convergence is in the metric consistent with the weak convergence topology of $\mathcal{P}(U \times Y)$ and, hence, it implies in particular that

$$(4.22) \quad \lim_{k \rightarrow \infty} \int_{U \times Y} (\phi^j(y_i))^T f_i(u, y, z) \gamma^k(du, dy) = \int_{U \times Y} (\phi^j(y_i))^T f_i(u, y, z) \gamma(du, dy).$$

According to the definition of the set $\Gamma(z, \bar{\epsilon}^k, S^k, y(0))$ (see (2.7)) and the fact that $\gamma^k \in \Gamma(z, \bar{\epsilon}^k, S^k, y(0))$, there exists an admissible control $u^k(\tau)$ and the corresponding trajectory $y^k(\tau)$ of system (2.4) such that

$$\int_{U \times Y} (\phi'(y_i))^T f_i(u, y, z) \gamma^k(du, dy) = \frac{1}{S^k} \int_0^{S^k} (\phi'(y_i^k(\tau)))^T f_i(u^k(\tau), y^k(\tau), z) d\tau.$$

The second integral is apparently equal to $\frac{\phi(y_i^k(S^k)) - \phi(y_i^k(0))}{S^k}$, which tends to zero as S^k tends to infinity (since, by Assumption 2.1, the solutions of (2.4) stay in the bounded area). This and (4.22) imply the validity of the proposition. \square

4.2. Proofs for section 3.

LEMMA 4.3. *The map $\psi(p)$ defined by (3.4) is continuous. That is, $\psi(p_l)$ converges to $\psi(p)$ in the weak convergence topology of $\mathcal{P}(U \times Y \times Z)$ if p_l converges to p in the weak convergence topology of $\mathcal{P}(\mathcal{P}(U \times Y) \times Z)$.*

Proof of Lemma 4.3. Let $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^1$ be a continuous function. Then

$$\begin{aligned} \lim_{p_l \rightarrow p} \int h(u, y, z) \psi(p_l)(du, dy, dz) &= \lim_{p_l \rightarrow p} \int \left(\int h(u, y, z) \gamma(du, dy) \right) p_l(d\gamma, dz) \\ &= \int \left(\int h(u, y, z) \gamma(du, dy) \right) p(d\gamma, dz) = \int h(u, y, z) \psi(p)(du, dy, dz), \end{aligned}$$

where it is taken into account that, because $h(u, y, z)$ is continuous, it follows that the function $\hat{h}(\gamma, z)$ defined by (2.20) is continuous as well. Since the last equalities are valid for any continuous $h(\cdot)$, it follows that $\lim_{p_l \rightarrow p} \psi(p_l) = \psi(p)$. \square

Proof of Lemma 3.5. Let the metric ρ of $\mathcal{P}(W)$ be defined as follows:

$$(4.23) \quad \rho(\gamma', \gamma'') = \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{|\langle \gamma', q_k \rangle - \langle \gamma'', q_k \rangle|}{1 + |\langle \gamma', q_k \rangle - \langle \gamma'', q_k \rangle|} \quad \forall \gamma', \gamma'' \in \mathcal{P}(W)$$

where $q_k : W \rightarrow \mathbb{R}^1, k = 0, 1, \dots$, is a sequence of Lipschitz continuous functions which is dense in the space of continuous functions $C(W)$ and $\langle \gamma, q_k \rangle = \int_W q_k(w) \gamma(dw)$. Note that this metric is consistent with the weak convergence topology of $\mathcal{P}(W)$. Define

$$(4.24) \quad \nu(\alpha) \stackrel{\text{def}}{=} \sup_{\beta \in \mathcal{B}} \sup_{\gamma \in \Gamma^1(\alpha, \beta)} \rho(\gamma, \Gamma^2(\alpha, \beta))$$

and show that $\nu(\alpha)$ tends to zero as α tends to α_0 . Assume that it does not. Then there exists a number $\delta > 0$ and sequences $(\alpha_l, \beta_l) \in \mathcal{A} \times \mathcal{B}, \gamma^l \in \Gamma^1(\alpha_l, \beta_l), l = 1, 2, \dots$, such that $\lim_{l \rightarrow \infty} \alpha_l = \alpha_0$ and $\rho(\gamma^l, \gamma) \geq \delta \forall \gamma \in \Gamma^2(\alpha, \beta)$. That is,

$$(4.25) \quad \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{|\langle \gamma^l, q_k \rangle - \langle \gamma, q_k \rangle|}{1 + |\langle \gamma^l, q_k \rangle - \langle \gamma, q_k \rangle|} \geq \delta \quad \forall \gamma \in \Gamma^2(\alpha, \beta).$$

Hence, for some integer K ,

$$(4.26) \quad \sum_{k=1}^K |\langle \gamma^l, q_k \rangle - \langle \gamma, q_k \rangle| \geq \frac{\delta}{2} \quad \forall \gamma \in \Gamma^2(\alpha, \beta).$$

Let $h(w) \stackrel{\text{def}}{=} (q_1(w), \dots, q_k(w)) : W \rightarrow \mathbb{R}^K$. Assume that the norm of a vector in (2.12) is defined as the sum of the absolute values of its components. Then, by (3.30), one can rewrite (4.26) in the form

$$d(v^l, v) \geq \frac{\delta}{2} \quad \forall v \in V_h^2(\alpha_l, \beta_l), \quad \text{where } v^l \stackrel{\text{def}}{=} \int_W h(w)\gamma^l(dw) \in V_h^1(\alpha_l, \beta_l).$$

Hence, $d(v^l, V_h^2(\alpha_l, \beta_l)) \geq \frac{\delta}{2}$, $l = 1, 2, \dots$, which contradicts (3.29) and thus proves the lemma. \square

Proof of Theorem 3.3(i). Let $\tilde{h}(\gamma, z) : \mathcal{P}(U \times Y) \times Z \rightarrow \mathbb{R}^j$, $j = 1, 2, \dots$, be an arbitrary Lipschitz continuous vector function. That is,

$$(4.27) \quad \|\tilde{h}(\gamma', z') - \tilde{h}(\gamma'', z'')\| \leq c_{\tilde{h}}(\|z' - z''\| + \rho(\gamma', \gamma'')), \quad c_{\tilde{h}} = \text{const}.$$

Consider a set-valued map $V(z)$ defined by (4.7) with $\tilde{h}(\gamma, z)$ as above. Note that this map is not necessarily convex valued since $\tilde{h}(\gamma, z)$ may not be represented as the integral (2.20). By (2.14), (4.27), and (2.22) (see Assumption 2.7), it satisfies Lipschitz conditions (4.9). Hence, by the relaxation theorem (see, e.g., [7, Theorem 10.4.4, p. 402]), the set of solutions of the differential inclusion (4.10) is dense in the set of solutions of the differential inclusion

$$(4.28) \quad \dot{\bar{z}}(t) \in \text{co}V(z(t)),$$

where $\text{co}V(z)$ is the convex hull of $V(z)$.

By Corollary 3.7, to establish the existence of a convex and compact set $\tilde{\Gamma} \subset \mathcal{P}(\mathcal{P}(U \times Y) \times Z)$ satisfying (3.12) it is enough to show that for any Lipschitz continuous $\tilde{h}(\gamma, z) : \mathcal{P}(U \times Y) \times Z \rightarrow \mathbb{R}^j$, $j = 1, 2, \dots$, there exist a convex and compact set $V_{\tilde{h}} \subset \mathbb{R}^j$ and a function $\mu_{\tilde{h}}(T)$ such that

$$(4.29) \quad d_H(V_{\tilde{h}}(T, z(0)), V_{\tilde{h}}) \leq \mu_{\tilde{h}}(T) \quad \forall z(0) \in Z'', \quad \lim_{T \rightarrow \infty} \mu_{\tilde{h}}(T) = 0,$$

where

$$(4.30) \quad \begin{aligned} V_{\tilde{h}}(T, z(0)) &= \bigcup_{p \in \tilde{\Gamma}(T, z(0))} \left\{ \int_{\mathcal{P}(U \times Y) \times Z} \tilde{h}(\gamma, z) p(d\gamma, dz) \right\} \\ &= \bigcup_{(\gamma(\cdot), z(\cdot))} \left\{ \frac{1}{T} \int_0^T \tilde{h}(\gamma(t), z(t)) dt \right\}, \end{aligned}$$

with the second union being taken over all admissible pairs of averaged system (2.15). The closure of the set (4.30), $\text{cl}V_{\tilde{h}}(T, z(0))$, allows also the representations

$$(4.31) \quad \text{cl}V_{\tilde{h}}(T, z(0)) = \text{cl} \bigcup_{\bar{z}(\cdot)} \left\{ \frac{\theta(T)}{T} \right\} = \bigcup_{\bar{z}(\cdot)} \left\{ \frac{\theta(T)}{T} \right\},$$

where the first union is taken over the solutions of (4.10) and the second over the solutions of (4.28), which satisfy the initial conditions $\bar{z}(0) = (z(0), 0)$.

As in the proof of Proposition 4.2, from Assumption 2.7 it follows that for any $z' \in Z$, $(v', w') \in V(z')$, and $z'' \in Z$, there exists $(v'', w'') \in V(z'')$ such that (4.15)–(4.16) are satisfied. It can be verified that the map $\text{co}V(z)$ has a similar property. That is, for any $z' \in Z$, $(v', w') \in \text{co}V(z')$, and $z'' \in Z$, there exists

$(v'', w'') \in \text{co}V(z'')$ such that (4.15)–(4.16) are satisfied. As with Proposition 4.2, this allows us to establish that, given a solution $(z^1(t), \theta^1(t))$ of the differential inclusion (4.28) satisfying the initial condition $(z^1(0), \theta^1(0)) = (z^1, \theta^1) \in Z' \times \mathbb{R}^j$ and a vector $(z^2, \theta^2) \in Z' \times \mathbb{R}^j$, there exists a solution $(z^2(t), \theta^2(t))$ of (4.28) which satisfies the initial condition $(z^2(0), \theta^2(0)) = (z^2, \theta^2)$ such that estimates (4.13)–(4.14) will be valid.

It follows from (4.14) that

$$(4.32) \quad d_H(\text{cl}V_{\bar{h}}(T, z^1), \text{cl}V_{\bar{h}}(T, z^2)) \leq b_2 T^{-1} \quad \forall z^i \in Z', \quad i = 1, 2, \quad \forall T \geq 0.$$

Now applying results from [15] or [19, Proposition 3.2], one can establish the existence of a convex and compact set $V_{\bar{h}}$ and a function $\mu_{\bar{h}}(T) = O(T^{-1/2})$ which satisfy (4.29). This completes the proof of the theorem. \square

REFERENCES

- [1] O. ALVAREZ AND M. BARDI, *Viscosity solutions methods for singular perturbations in deterministic and stochastic controls*, SIAM J. Control Optim., 40 (2001), pp. 1159–1188.
- [2] Z. ARTSTEIN, *Invariant measures of differential inclusions*, J. Differential Equations, 152 (1999), pp. 289–307.
- [3] Z. ARTSTEIN, *The chattering limit of singularly perturbed optimal control problems*, in Proceedings of CDC-2000, Control and Decision Conference, Sydney, 2000, pp. 564–569.
- [4] Z. ARTSTEIN AND V. GAITSGORY, *Tracking fast trajectories along a slow dynamics: A singular perturbations approach*, SIAM J. Control Optim., 35 (1997), pp. 1487–1507.
- [5] Z. ARTSTEIN AND A. VIGODNER, *Singularly perturbed ordinary differential equations with dynamical limits*, Proc. Roy. Soc. Edinburgh Sect. A, 126 (1996), pp. 541–569.
- [6] J.-P. AUBIN, *Viability Theory*, Birkhäuser, Boston, 1991.
- [7] J.-P. AUBIN AND H. FRANKOWSKA, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
- [8] F. BAGAGIOLIO AND M. BARDI, *Singular perturbation of a finite horizon problem with state-space constraints*, SIAM J. Control Optim., 36 (1998), pp. 2040–2060.
- [9] A. BENSOUSSAN, *Perturbations Methods in Optimal Control Problems*, Wiley, New York, 1984.
- [10] D.P. BERTSEKAS AND S.E. SHREVE, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1978.
- [11] T. DONCHEV AND I. SLAVOV, *Averaging method for one-sided Lipschitz differential inclusions with generalized solutions*, SIAM J. Control Optim., 37 (1999), pp. 1600–1613.
- [12] A. DONTCHEV, *Time-scale decomposition of the reachable set of constrained linear systems*, Math. Control Signals Systems, 5 (1992), pp. 327–340.
- [13] A. DONTCHEV, T. DONCHEV, AND I. SLAVOV, *A Tichonov-type theorem for singularly perturbed differential inclusions*, Nonlinear Analysis, 26 (1996), pp. 1547–1554.
- [14] O.P. FILATOV AND M.M. HAPAEV, *Averaging of Systems of Differential Inclusions*, Moscow University Publishing House, Moscow, 1998 (in Russian).
- [15] V. GAITSGORY, *Use of the averaging method in control problems*, Differential Equations, 22 (1986), pp. 1290–1299 (translated from Russian).
- [16] V. GAITSGORY, *Suboptimization of singularly perturbed control systems*, SIAM J. Control Optim., 30 (1992), pp. 1228–1249.
- [17] V. GAITSGORY AND A. LEIZAROWITZ, *Limit occupational measures set for a control system and averaging of singularly perturbed control systems*, J. Math. Anal. Appl., 233 (1999), pp. 461–475.
- [18] V. GAITSGORY AND M.-T. NGUYEN, *Averaging of three time scale singularly perturbed control systems*, Systems Control Lett., 42 (2001), pp. 395–403.
- [19] G. GRAMMEL, *Averaging of singularly perturbed systems*, Nonlinear Analysis, 28 (1997), pp. 1855–1865.
- [20] G. GRAMMEL, *Order Reduction for Nonlinear Systems by Re-iterated Averaging*, in Proceedings of the IEEE Singapore International Symposium on Control Theory and Applications, Singapore, 1997, pp. 1855–1865.
- [21] P.V. KOKOTOVIĆ, *Applications of singular perturbations techniques to control problems*, SIAM Rev., 26 (1984), pp. 501–550.
- [22] P.V. KOKOTOVIĆ, H.K. KHALIL, AND J. O'REILLY, *Singular Perturbation Methods in Control: Analysis and Design*, Academic Press, London, 1986.

- [23] H. KUSHNER, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, 1990.
- [24] R.E. O'MALLEY, JR., *Introduction to Singular Perturbations*, Academic Press, New York, 1974.
- [25] R.E. O'MALLEY, JR., *Singular perturbations and optimal control*, in *Mathematical Control Theory*, Lecture Notes in Math. 680, W.A. Copel, ed., Springer-Verlag, Berlin, 1978, pp. 170–218.
- [26] Z. PAN AND T. BASAR, *Multi-time scale zero-sum differential games with perfect state measurements*, *Dynam. Control*, 5 (1995), pp. 7–29.
- [27] V.A. PLOTNIKOV, *Differential Equations with Multivalued Right-Hand Sides: Asymptotic Methods*, AstroPrint, Odessa, 1999 (in Russian).
- [28] M. QUINCAMPOIX AND H. ZHANG, *Singular perturbations in non-linear optimal control systems*, *Differential Integral Equations*, 8 (1995), pp. 931–944.
- [29] M. QUINCAMPOIX AND M. VELIOV, *Open-loop viable control under uncertain initial state information*, *Set-Valued Anal.*, 7 (1999), pp. 55–87.
- [30] A.B. VASIL'ÉVA AND A.F. BUTUZOV, *Asymptotic Expansions of Solutions of Singularly Perturbed Equations*, Nauka, Moscow, 1973 (in Russian).
- [31] V. VELIOV, *A generalization of Tikhonov theorem for singularly perturbed differential inclusions*, *J. Dynam. Control Systems*, 3 (1997), pp. 291–319.
- [32] A. VIGODNER, *Limits of singularly perturbed control problems with statistical dynamics of fast motions*, *SIAM J. Control Optim.*, 35 (1997), pp. 1–28.
- [33] J. WARGA, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.
- [34] G.G. YIN AND Q. ZHANG, *Continuous-Time Markov Chains and Applications. A Singular Perturbations Approach*, Springer-Verlag, New York, 1998.