Practical Exponential Stability and Closeness of Solutions for Singularly Perturbed Systems via Averaging*

Mohammad Deghat[a], Saeed Ahmadizadeh[b], Dragan Nešić[b], Chris Manzie[b]

[a]*School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney, NSW 2052, Australia*
[b]*Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia*

Abstract

This paper studies the behavior of singularly perturbed nonlinear differential equations with boundary-layer solutions that do not necessarily converge to an equilibrium. Using the average for the derivative of the slow state variables and assuming the boundary-layer solutions converge exponentially fast to a bounded set, which is possibly parameterized by the slow variable, results on the closeness of solutions of the singularly perturbed system to the solutions of the reduced average and boundary-layer systems over a finite time interval are presented. The closeness of solution error is shown to be of order $O(\sqrt{\varepsilon})$ where ε is the perturbation parameter. Moreover, under the additional assumption of exponential stability of the reduced average system, practical exponential stability of the solutions of the singularly perturbed system is provided.

Keywords: Singular perturbation, Averaging, Closeness of solutions.

1. Introduction

The singular perturbation method is a common technique to analyze a two-time-scale system via the behavior of two auxiliary systems, namely the reduced (slow) system and the boundary-layer (fast) system. Two-time-scale systems arise in a variety of engineering and applied science applications such as electro-mechanical systems [1], power electronic systems with DC-DC or PFC converters [2], combustion engines [3], classical mechanical and quantum mechanical systems [4], fast sensors and/or actuators [1] and complex systems, such as networks consisting of many agents [5].

The study of singular perturbation systems was started in the mathematical literature by Tikhonov [6, 7] and was followed by [8–11]. In general, the results using the singular perturbation method either relate the stability properties of the original system with the above-mentioned auxiliary systems or show the closeness of solutions of the original system to the solutions of the auxiliary systems; see e.g. [1], [12, Sec. 11] for results on stability and closeness of solutions of the classical singular perturbation with applications to systems and control problems. It is usually assumed in the classical singular perturbation results that the solutions of the boundary-layer system converge to a unique equilibrium manifold. The case where the solutions converge to a bounded set, e.g. a set of limit cycles, has been studied using the averaging method [13, 14]. In these results, the derivative of the slow state is averaged over a finite or infinite time interval and the behavior of the reduced averaged slow system, together with the behavior of the boundary-layer system, is used to describe the behavior of the full-order system. This idea can be found in the work of Gaitsgory *et al.* [15–17], Grammel [18–20], Artstein *et al.* [21–24], Teel *et al.* [25], and others [26, 27].

The problem of exponential stability of this general class of singular perturbation is not well studied in the literature. Among the above-mentioned results, Grammel showed in [20], using a trajectory-based proof, that under the exponential stability of the origin of the reduced average system and under some other conditions on the system model, the slow state of a delayed singularly perturbed system is exponentially stable. However, the behavior of the fast state and also the closeness of solutions of the singularly perturbed system to the solutions of the reduced average and boundary-layer systems when the reduced average system is not exponentially stable are not studied in [20]. Furthermore, some of the assumptions in [20] are strong and not satisfied even by linear systems.

This paper studies non-delayed singularly perturbed systems and presents results on both the stability and closeness of solutions. In particular, it is shown that under the exponential stability of the boundary-layer system and some other conditions on the system model and over a finite time interval, the solutions to the singularly perturbed system are approximated by the solutions of the reduced average and boundary-layer systems when the perturbation parameter, ε, is small. Although Grammel did not study closeness of solutions in [20], Teel *et. al* presented a closeness of solution result in [25] which can be applied to a general class of singular perturbation systems. However, the order of magnitude of error is not studied in [25]. Compared to [25], we propose stronger conditions on the system model and obtain stronger closeness of solution results; we show under our conditions that the approximation error is of order $O(\sqrt{\varepsilon})$. If the reduced average system is also exponentially stable, then it

*This work was supported by the ARC Discovery Scheme, grant number DP170104102.

is shown that over the infinite time interval, the fast and slow states of the singularly perturbed system are both practically exponentially stable.

An abbreviated conference version of this paper has been presented in [28]. This paper has the following key additions over the conference version. It includes complete proofs which were not included in [28]. It also studies a more general case where the reduced average system is defined as a *differential inclusion*, while [28] assumes the reduced average system is defined as a *differential equation*. A new section has also been added to the paper that studies the practical stability of the singularly perturbed system over an infinite time interval. Moreover, some of the assumptions in the conference version such as the global Lipschitz assumption of the average system and the forward invariance assumption of a closed set of initial conditions with respect to an auxiliary slow system are relaxed in this paper.

The structure of the paper is as follows. In Section 2, the general singular perturbation problem is explained and some of the main assumptions on the system model are presented. Section 3 contains the main results of the paper. A numerical example is presented in Section 4 with concluding remarks provided in Section 5.

Notation:

- $B_R(0)$ denotes a ball of radius $R > 0$ centered at the origin.
- [12, Definition 10.1]: A function $\delta_1(\varepsilon)$ is of order $O(\delta_2(\varepsilon))$, denoted by $\delta_1(\varepsilon) = O(\delta_2(\varepsilon))$, if there exist positive constants k and c such that

$$|\delta_1(\varepsilon)| \leqslant k|\delta_2(\varepsilon)|, \quad \forall|\varepsilon| < c. \tag{1}$$

  If $\delta_1(\varepsilon)$ and $\delta_2(\varepsilon)$ are continuous at $\varepsilon = 0$, then (1) implies that

$$\limsup_{\varepsilon\to 0} \frac{|\delta_1(\varepsilon)|}{|\delta_2(\varepsilon)|} \leqslant k < \infty. \tag{2}$$

- The standard Euclidean norm is denoted by $\|\cdot\|$. The distance between a point x and a non-empty set A is denoted by $d(x, A)$, i.e.

$$d(x, A) = \inf_{y\in A} \|x - y\|. \tag{3}$$

- The Hausdorff distance of compact subsets A and B of a compact metric space X is denoted by $d_{\mathrm{H}}(A, B)$ and is defined as

$$d_{\mathrm{H}}(A, B) := \max\left\{\sup_{x\in A}\inf_{y\in B}\|x - y\|, \sup_{y\in B}\inf_{x\in A}\|x - y\|\right\}.$$

2. Preliminaries

Consider a singularly perturbed system

$$\dot{x} = f(x, z, \varepsilon), \quad x(0) = x_0, \tag{4a}$$
$$\varepsilon\dot{z} = g(x, z, \varepsilon), \quad z(0) = z_0, \tag{4b}$$

where $\varepsilon > 0$ is a small perturbation parameter, and $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^m$ are respectively the slow and fast variables. Define the *fast-time* variable $\tau = t/\varepsilon$. Then in the τ-domain, (4) can be written as

$$\frac{dx}{d\tau} = \varepsilon f(x, z, \varepsilon), \tag{5a}$$
$$\frac{dz}{d\tau} = g(x, z, \varepsilon). \tag{5b}$$

Following the standard procedure for the analysis of the singularly perturbed systems, we decompose the system into two auxiliary reduced-order systems (namely, the boundary-layer system and the reduced system), each one associated with a different time scale, and then prove a certain stability property for the singularly perturbed system, or a closeness of solution result, using the stability properties of the reduced-order systems. These types of results are usually valid when the perturbation parameter ε is sufficiently small.

Letting $\varepsilon = 0$, (5a) becomes $dx/d\tau = 0$ which implies that the slow variable x is fixed, i.e. $x(\tau) = x_0$, $\forall\tau \geqslant 0$. Then the *boundary-layer system* is obtained by setting $\varepsilon = 0$ in (5b) as

$$\frac{dz_b}{d\tau} = g(x, z_b, 0), \quad z_b(0) = z_0, \tag{6}$$

where $z_b(\tau)$ denotes the state of the boundary-layer system and $x(\tau) = x_0$ is treated as a fixed parameter.

Unlike the classical singular perturbation problem, we assume the solutions to the boundary-layer system, denoted by $\phi_b(\tau, x, z_0)$, or by $\phi_b(\tau)$ for the ease of notation, do not necessarily converge to a unique equilibrium, but may converge to a bounded set. For example, the solutions to the boundary-layer system may converge to a set of limit cycles parameterized by the slow variable x. We denote this set throughout the paper by H_x where the sub-script x is used to highlight the fact that H_x is, in general, parameterized by x. We formally state this assumption and some other assumptions on the system model below. But we first define an auxiliary signal $\xi_\varepsilon(t)$ which will be used later in the proof of Theorem 1 and the definition of the signals $\Delta_l(t)$ and $\delta_l(t)$ in (24) and (25), and has properties which are given below in Assumption 1, item 2.

Denote the solution of (4) by $\big(x_\varepsilon(t), z_\varepsilon(t)\big)$ where the subscript ε shows the dependence of the solution on ε. The signal $\xi_\varepsilon(t)$ is defined for $t \in [t_l, t_{l+1}]$ as

$$\xi_\varepsilon(t) := \xi_\varepsilon(t_l) + \int_{t_l}^{t} f(\xi_\varepsilon(t_l), y_\varepsilon(s), 0)ds, \tag{7}$$

where $\xi_\varepsilon(0) = x_0$ and $y_\varepsilon(t) : [t_l, t_{l+1}] \to \mathbb{R}^m$ is the unique solution to

$$\varepsilon\dot{y}(t) = g(\xi_\varepsilon(t_l), y(t), 0), \quad y(t_l) = z_\varepsilon(t_l). \tag{8}$$

Note that $\xi_\varepsilon(t)$ is a continuous-time signal, however $y_\varepsilon(t)$ is piecewise continuous as its value at time instants t_l changes to $z_\varepsilon(t_l)$. More information on how the time instants t_l are chosen will be given in Section 3.1.

Assumption 1. *There exist $\varepsilon_1 > 0$ and compact sets $B_{\bar{R}}(0) \subset \mathbb{R}^n$ and $\bar{M} \subset \mathbb{R}^m$ such that*

1. $f(x,z,\varepsilon)$ and $g(x,z,\varepsilon)$ are Lipschitz on $B_{\bar{R}}(0)\times\bar{M}\times[0,\varepsilon_1]$.

2. For any given $T>0$ and for all $t\in[0,T]$, there exist initial condition sets $B_R(0)\subset B_{\bar{R}}(0)\subset\mathbb{R}^n$ and $M\subset\bar{M}\subset\mathbb{R}^m$ such that the solutions to (4) and (7) satisfy $x_\varepsilon(t),\xi_\varepsilon(t)\in B_{\bar{R}}(0)$ and $z_\varepsilon(t)\in\bar{M}$.

3. For any given $x\in B_R(0)$ and any initial condition $z_0\in M$, a bounded set H_x to which the solutions of the boundary-layer system (6) converge exists and satisfies $H_x\subset\bar{M}$.

4. For all x_1, $x_2\in B_{\bar{R}}(0)$, there exists a constant $L_H>0$ such that

$$d_{\mathrm{H}}(H_{x_1},H_{x_2})\leqslant L_H\|x_1-x_2\|. \tag{9}$$

Assumption 2 (Boundary-layer solutions). *The solutions of (6), denoted by $\phi_b(\tau,x,z_0)$, converge locally exponentially fast to H_x stated in Assumption 1. More precisely, there exist constants $r_y>0$ and $\beta_y>0$ and a forward invariant set M, stated in Assumption 1, with respect to (6) such that for all $x\in\mathbb{R}^n$ and all initial values $z_0\in M$,*

$$d\big(\phi_b(\tau,x,z_0),H_x\big)\leqslant r_y e^{-\beta_y\tau}d(z_0,H_x). \tag{10}$$

The above assumption will be used to guarantee the existence of set-valued averages and also in the closeness of solutions and stability analysis. We define the following set-valued averages which depend on the forward invariant set $M\subset\mathbb{R}^m$.

Definition 1 (Set-valued average [20]). *For $\mathcal{T}>0$ and $x\in\mathbb{R}^n$, the finite-time average $F_{\mathcal{T}}(x)$ is defined as*

$$F_{\mathcal{T}}(x)=\operatorname{conv}\left(\bigcup_{z_0\in M}\left\{\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} f\big(x,\phi_b(s,x,z_0),0\big)ds\right\}\right), \tag{11}$$

where x is treated as a parameter and $\operatorname{conv}(S)$ denotes the closed convex hull of a set S.

Proposition 1 ([20, Proposition 2.4.]). *Under the Lipschitz condition of f and g in Assumption 1 and the forward-invariance property of M in Assumption 2, for any $x\in\mathbb{R}^n$, there is a convex compact set $F_{av}(x)\subset\mathbb{R}^n$ with*

$$\mathrm{d_H}\big(F_{\mathcal{T}}(x),F_{av}(x)\big)\to 0,\quad \text{as } \mathcal{T}\to\infty. \tag{12}$$

Furthermore, the set-valued mapping $F_{av}(\cdot)$ takes nonvoid, compact, convex values and is upper semi-continuous.

Remark 1. *To define the reduced average system, it is not required to know the set H_x. The structure of the fast dynamics is also not directly employed in the definition of $F_{\mathcal{T}}$ and F_{av}, namely, the set-valued maps $F_{\mathcal{T}}$ and F_{av} are stated in terms of the integrals of the fast dynamics mapped by the function f.*

Remark 2. *A condition for the existence of the set-valued map $F_{av}(x)$ is given in Proposition 1. The set-valued map F_{av} is defined via the convergence of the integrals in Definition 1; however, it can also be obtained via averaging the slow subsystem over invariant measures from the limit occupational measures set constructed for the boundary-layer system (6). The reader is referred to [29–31] for more information.*

The *reduced average system* (or what is called the *reduced system* in the rest of the paper) is defined by the following differential inclusion

$$\dot{x}_{av}\in F_{av}(x_{av}),\quad x_{av}(0)=x_0. \tag{13}$$

We make the following assumptions on the set valued mapping F_{av}.

Assumption 3. *There is a function $\gamma(\mathcal{T},x,z)$ with $\lim_{\mathcal{T}\to\infty}\gamma(\mathcal{T},x,z)=0$ such that*

$$\mathrm{d_H}\big(F_{\mathcal{T}}(x),F_{av}(x)\big)\leqslant\gamma(\mathcal{T},x,z) \tag{14}$$

for all $x\in B_{\bar{R}}(0)$, $z\in\bar{M}$ and $\mathcal{T}>0$.

Remark 3. *Since $B_{\bar{R}}(0)$ and $\bar{M}$ are bounded sets, there exists a function $\bar{\gamma}(\mathcal{T})$, with $\lim_{\mathcal{T}\to\infty}\bar{\gamma}(\mathcal{T})=0$, such that $\gamma(\mathcal{T},x,z)\leqslant\bar{\gamma}(\mathcal{T})$ $\forall x\in B_{\bar{R}}(0)$, $z\in\bar{M}$. Therefor, (14) can be written as*

$$\mathrm{d_H}\big(F_{\mathcal{T}}(x),F_{av}(x)\big)\leqslant\bar{\gamma}(\mathcal{T}). \tag{15}$$

Assumption 4 (Lipschitz continuity of F_{av}). *The set-valued mapping F_{av} is Lipschitz on $B_{\hat{R}}(0)$ for some $\hat{R}>\bar{R}$.* [1]

The assumptions in this paper are generally standard in singular perturbation literature. As stated earlier, the exponential stability assumptions are used to establish the main result of the paper. While Assumption 2 is stronger than asymptotic stability assumptions considered in [25], we obtain stronger results than asymptotic stability ones. Unlike the classical singular perturbation results, the boundary-layer solutions are assumed here to converge to a family of bounded sets. Hence, we considered Assumption 3 and 4 which are similar to the assumptions in [20, 25].

Throughout the paper, we denote $L>0$ as the Lipschitz constant of $f(x,z,\varepsilon)$ and $g(x,z,\varepsilon)$ on $B_{\bar{R}}(0)\times\bar{M}\times[0,\varepsilon_1]$ and denote $L_{av}>0$ as the Lipschitz constant of $F_{av}(x)$ on $B_{\hat{R}}(0)$. We also define the bound P such that

$$\|f(x,z,\varepsilon)\|\leqslant P,\quad \|g(x,z,\varepsilon)\|\leqslant P, \tag{16}$$

for all $x\in B_{\bar{R}}(0)$, $z\in\bar{M}$, $\varepsilon\in[0,\varepsilon_1]$ and all $t\in[0,T]$. The existence of P can be concluded from Assumption 1.

3. Main result

3.1. Closeness of solutions over a finite time interval

In this subsection, we analyze the closeness of solutions of the singularly perturbed system and the reduced and boundary-layer systems over a finite time interval. This result is independent of any stability properties of the reduced system (13).

[1] [32, Definition 4, Chapter 1]: A set-valued map F_{av} from a metric space X to a metric space Y is locally Lipschitz if for any $x_0\in X$, there exist a neighborhood $N(x_0)\subset X$ and a Lipschitz constant $L_{av}\geqslant 0$ such that

$$\forall x,x'\in N(x_0),\qquad F_{av}(x)\subset B(F_{av}(x'),L_{av}\|x-x'\|)$$

where $B(K,\alpha):=\{x\in X\mid d(x,K)\leqslant\alpha\}$. F_{av} is Lipshitz with Lipschiz constant L_{av} if

$$\forall x,x'\in X,\qquad F_{av}(x)\subset B(F_{av}(x'),L_{av}\|x-x'\|).$$

We aim to investigate the system on a finite time horizon $t \in [kT,(k+1)T]$ where $k \in \mathbb{Z}$, $T > 0$ and $t_0 := kT$. We divide this time interval into sub intervals of the form $[t_l, t_{l+1}]$ which all have the same length $\varepsilon S_\varepsilon$, except possibly the last interval with length smaller than or equal to the length $\varepsilon S_\varepsilon$, and the index l is an element of the index set $I_\varepsilon = \{0, 1, \cdots, \lfloor T/\varepsilon S_\varepsilon \rfloor\}$, where $\lfloor \cdot \rfloor$ denotes the floor function. The last time in the sequence is equal to $(k+1)T$.

In order to state the main result of this subsection, we first highlight that the set H_x, which is parameterized by x, is the set to which the solutions of the boundary-layer system converge (note that x is fixed for the boundary-layer system). Given a solution of the singularly perturbed system $(x_\varepsilon(t), z_\varepsilon(t))$ for $t \geqslant 0$, we use $H_{x_\varepsilon(t)}$ to refer to time-varying sets where, at each time instant $t = t^*$, the set $H_{x_\varepsilon(t^*)}$ is the set to which the solution of the boundary-layer system with the fixed value of $x = x_\varepsilon(t^*)$ converges.

Theorem 1 (closeness of solutions over a finite time)**.** *For any given finite $T > 0$, consider a $\mathcal{K}_\infty$ function $q(S)$ given by*

$$q(S) := S e^{TL(1+SLe^{SL})} \tag{17}$$

and define the map $\varepsilon \to S_\varepsilon$ as [2]

$$S_\varepsilon = q^{-1}(\varepsilon^{-1/4}). \tag{18}$$

Let Assumptions 1-4 hold. Then there exists a sufficiently small $\hat{\varepsilon} > 0$ such that for for every $(x_0, z_0, \varepsilon) \in B_R(0) \times M \times (0, \hat{\varepsilon}]$,

(i) given the solution $(x_\varepsilon(t), z_\varepsilon(t))$ of (4), there exists a solution $x_{av}(t)$ of the differential inclusion (13)*, such that for $t \in [0, T]$*

$$\|x_\varepsilon(t) - x_{av}(t)\| \leqslant K(\varepsilon), \tag{19}$$

where $K(\varepsilon) : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ satisfies $\lim_{\varepsilon \to 0} K(\varepsilon) = 0$. Furthermore,

$$d(z_\varepsilon(t), H_{x_\varepsilon(t)}) \leqslant r_y e^{-(\beta_y - \delta_y)t/\varepsilon} d(z_0, H_{x_0}) + F(\varepsilon) \tag{20}$$

where $F(\varepsilon) : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ satisfies $\lim_{\varepsilon \to 0} F(\varepsilon) = 0$.

(ii) if there exist $\varepsilon^* \in (0, \hat{\varepsilon}]$, $r' > 0$ and $\alpha_1 : \alpha_1 > 2$ such that for $S_\varepsilon \geqslant S_{\varepsilon^*}$, the function $\bar{\gamma}(S_\varepsilon)$ satisfies

$$\bar{\gamma}(S_\varepsilon) \leqslant r' e^{-\alpha_1 TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)}, \tag{21}$$

then

$$\|x_\varepsilon(t) - x_{av}(t)\| = O(\sqrt{\varepsilon}) \tag{22}$$

holds for $\varepsilon \in (0, \varepsilon^*]$, uniformly on $t \in [0, T]$. Moreover, given any $t_a : t_a \in (0, T)$, there exists $\varepsilon^{**} \leqslant \varepsilon^*$ such that

$$\left| d(z_\varepsilon(t), H_{x_\varepsilon(t)}) - d(\phi_b(t/\varepsilon), H_{x_0}) \right| = O(\sqrt{\varepsilon}) \tag{23}$$

holds uniformly on $t \in [t_a, T]$ when $\varepsilon \in (0, \varepsilon^{**}]$.

To prove Theorem 1, we need to define some auxiliary signals $\Delta_l(t)$, $d_l(t)$ and $D_l(t)$. We define these signals for $t \in [t_l, t_{l+1}]$ as

$$\Delta_l(t) := \max_{t_l \leqslant s \leqslant t} \|x_\varepsilon(s) - \xi_\varepsilon(s)\|, \tag{24}$$

$$d_l(t) := \max_{t_l \leqslant s \leqslant t} \|x_\varepsilon(s) - \xi_\varepsilon(t_l)\|, \tag{25}$$

$$D_l(t) := \max_{t_l \leqslant s \leqslant t} \|z_\varepsilon(s) - y_\varepsilon(s)\|. \tag{26}$$

We state the following lemmas which gives the limit value of S_ε and the upper bounds for $\Delta_l(t)$ and $D_l(t)$ over a finite time interval. The idea for the lemma is taken from [19].

Lemma 1. *For any given $L > 0$ and $T > 0$, the map $\varepsilon \to S_\varepsilon$ defined in* (18) *has the following properties*

$$\lim_{\varepsilon \to 0} S_\varepsilon = \infty, \tag{27a}$$

$$\lim_{\varepsilon \to 0} \varepsilon^{1/4} S_\varepsilon = 0. \tag{27b}$$

Proof. See the Appendix. □

Lemma 2. *Consider S_ε defined in* (18) *and suppose Assumption 1 holds for $(x_0, z_0, \varepsilon) \in B_R(0) \times M \times (0, \varepsilon_1]$. Then for $t \in [kT, (k+1)T]$ where $k \in \mathbb{Z}$, the signals $\Delta_l(t)$ and $D_l(t)$, $l \in I_\varepsilon$, defined respectively in* (24) *and* (26) *are upper bounded by $\bar{\Delta}(\varepsilon)$ and $\bar{D}(\varepsilon)$ defined as*

$$\bar{\Delta}(\varepsilon) := \Big(2\varepsilon S_\varepsilon P + TL(\varepsilon S_\varepsilon P + \varepsilon)\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)\Big) \times e^{TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)}, \tag{28}$$

$$\bar{D}(\varepsilon) := S_\varepsilon L\big(\bar{\Delta}(\varepsilon) + \varepsilon S_\varepsilon P + \varepsilon\big) e^{S_\varepsilon L}. \tag{29}$$

Furthermore, $\bar{\Delta}(\varepsilon)$ and $\bar{D}(\varepsilon)$ are of order $O(\sqrt{\varepsilon})$ and therefore converge to zero as $\varepsilon \to 0$.

Proof. See the Appendix. □

We are now ready to state the proof of Theorem 1.

Proof of Theorem 1. (i) By Assumption 3, we obtain for all $l \in I_\varepsilon$ and for any given initial conditions $x_0 \in B_R(0)$ and $z_0 \in M$ that

$$d\left(\frac{1}{S_\varepsilon}\int_{\tau=t_l/\varepsilon}^{\tau=t_{l+1}/\varepsilon} f\big(\xi_\varepsilon(t_l), y_\varepsilon(\varepsilon\tau), 0\big)d\tau\ ,\ F_{av}(\xi_\varepsilon(t_l))\right) \leqslant \bar{\gamma}(S_\varepsilon). \tag{30}$$

Note that $\phi_b(\tau)$ is a solution in the fast-time domain while $y_\varepsilon(t)$ is the solution in the slow-time domain. So by a change of variable $s = \varepsilon\tau$, the inequality (30) can be written as

$$d\left(\frac{1}{\varepsilon S_\varepsilon}\int_{s=t_l}^{s=t_{l+1}} f\big(\xi_\varepsilon(t_l), y_\varepsilon(s), 0\big)ds\ ,\ F_{av}(\xi_\varepsilon(t_l))\right) \leqslant \bar{\gamma}(S_\varepsilon). \tag{31}$$

Then we choose a $v_l \in F_{av}(\xi_\varepsilon(t_l))$ for $l \in I_\varepsilon$ and obtain that

$$\left\| \frac{1}{\varepsilon S_\varepsilon}\int_{t_l}^{t_{l+1}} f\big(\xi_\varepsilon(t_l), y_\varepsilon(s), 0\big)ds - v_l \right\| \leqslant \bar{\gamma}(S_\varepsilon). \tag{32}$$

[2] Note that $q(S)$ is invertible as it is a $\mathcal{K}_\infty$ function. Combining (17) and (18), we obtain that $\varepsilon^{-1/4} = S_\varepsilon e^{TL(1+S_\varepsilon L e^{S_\varepsilon L})}$. This definition is inspired from [19] and is slightly different. The map $\varepsilon \to S_\varepsilon$ cannot be defined similarly to [19] as the definition in [19] does not guarantee the order of magnitude of error is $O(\sqrt{\varepsilon})$.

Define a new family $(\eta_l)_{l\in I_\varepsilon} \subset \mathbb{R}^n$ by

$$\eta_{l+1} := \eta_l + \varepsilon S_\varepsilon v_l, \qquad \eta_0 = x_0, \tag{33}$$

and define a piecewise interpolating curve by

$$\eta_\varepsilon(t) := \eta_l + v_l(t - t_l) \tag{34}$$

for $t \in [t_l, t_{l+1}]$ and $l \in I_\varepsilon$. Using (7) and (33) we have

$$\begin{aligned}
\|\xi_\varepsilon(t_{l+1}) - \eta_{l+1}\| &\leqslant \|\xi_\varepsilon(t_l) - \eta_l\| \\
&\quad + \left\| \int_{t_l}^{t_{l+1}} f\big(\xi_\varepsilon(t_l), y_\varepsilon(s), 0\big) ds - \varepsilon S_\varepsilon v_l \right\| \\
\overset{(32)}{\Longrightarrow} \quad &\leqslant \|\xi_\varepsilon(t_l) - \eta_l\| + \varepsilon S_\varepsilon \bar{\gamma}(S_\varepsilon),
\end{aligned} \tag{35}$$

and we obtain by induction that

$$\|\xi_\varepsilon(t_l) - \eta_l\| \leqslant T\bar{\gamma}(S_\varepsilon), \tag{36}$$

for all $l \in I_\varepsilon$.

Inequality (36) implies that, for a sufficiently small $\hat{\varepsilon}$, the signal $\eta_\varepsilon(t)$ remains in $B_{\hat{R}}(0)$ and therefore we can use the Lipschitz property of F_{av} in Assumption 1. Given $\eta_\varepsilon(t)$ defined in (34), we can calculate an upper bound for $d\big(\dot{\eta}_\varepsilon(t), F_{av}(\eta_\varepsilon(t))\big)$ as follows

$$\begin{aligned}
&d\big(\dot{\eta}_\varepsilon(t), F_{av}(\eta_\varepsilon(t))\big) \\
&\quad \leqslant d\big(v_l, F_{av}(\xi_\varepsilon(t_l))\big) + d_{\mathrm{H}}\big(F_{av}(\xi_\varepsilon(t_l)), F_{av}(\eta_l)\big) \\
&\qquad + d_{\mathrm{H}}\big(F_{av}(\eta_l), F_{av}(\eta_\varepsilon(t))\big) \\
&\quad \leqslant L_{av}T\bar{\gamma}(S_\varepsilon) + L_{av}\varepsilon S_\varepsilon P,
\end{aligned} \tag{37}$$

where we use the fact that $d\big(v_l, F_{av}(\xi_\varepsilon(t_l))\big) = 0$ for $v_l \in F_{av}(\xi_\varepsilon(t_l))$. Then according to Filippov theorem [33, Theorem 11.3.9], there exists a solution $x_{av}(t) : [0, T] \to \mathbb{R}^n$ to the differential inclusion (13) such that

$$\begin{aligned}
\|x_{av}(t) - \eta_\varepsilon(t)\| &\leqslant e^{L_{av}T}\Big(\|x_0 - \eta_0\| \\
&\qquad + \int_0^T d\big(\dot{\eta}_\varepsilon(s), F_{av}(\eta_\varepsilon(s))\big) e^{-L_{av}s} ds\Big) \\
\overset{(33),(37)}{\Longrightarrow} \quad &\leqslant e^{L_{av}T} \frac{(1 - e^{-L_{av}T})}{L_{av}} \big(L_{av}T\bar{\gamma}(S_\varepsilon) + L_{av}\varepsilon S_\varepsilon P\big) \\
&\leqslant Te^{L_{av}T}\Big(L_{av}T\bar{\gamma}(S_\varepsilon) + L_{av}\varepsilon S_\varepsilon P\Big),
\end{aligned} \tag{38}$$

where we used the fact that $1 - e^{-aT} \leqslant aT$ for all $a \geqslant 0$ and $T \geqslant 0$. We now estimate an upper bound for $\|x_\varepsilon(t) - x_{av}(t)\|$.

$$\begin{aligned}
\|x_\varepsilon(t) - x_{av}(t)\| &\leqslant \|x_\varepsilon(t) - \xi_\varepsilon(t)\| + \|\xi_\varepsilon(t) - \xi_\varepsilon(t_l)\| \\
&\quad + \|\xi_\varepsilon(t_l) - \eta_l\| + \|\eta_l - \eta_\varepsilon(t)\| \\
&\quad + \|\eta_\varepsilon(t) - x_{av}(t)\|.
\end{aligned} \tag{39}$$

From (24) and Lemma 2, for any l in the index set I_ε, the first term on the right hand side of (39) is less than or equal to $\bar{\Delta}(\varepsilon)$. The rest of the terms can also be upper bounded, respectively using (7), (36), (34) and (38) as

$$\|x_\varepsilon(t) - x_{av}(t)\| \leqslant K(\varepsilon), \tag{40}$$

where $K(\varepsilon)$ is

$$\begin{aligned}
K(\varepsilon) := &\bar{\Delta}(\varepsilon) + \varepsilon S_\varepsilon P + T\bar{\gamma}(S_\varepsilon) + \varepsilon S_\varepsilon v_l \\
&+ Te^{L_{av}T}\Big(L_{av}T\bar{\gamma}(S_\varepsilon) + L_{av}\varepsilon S_\varepsilon P\Big).
\end{aligned} \tag{41}$$

Finally, we conclude from Lemma 1 and Lemma 2 that $\lim_{\varepsilon\to 0} K(\varepsilon) = 0$.

We now study the behavior of the solution of the fast state, $z_\varepsilon(t)$, and calculate its distance to the set $H_{x_\varepsilon(t)}$.

Define $\bar{y}_\varepsilon(t)$ as the solution to the following system for $t \in [t_l, t_{l+1}]$,

$$\varepsilon\dot{\bar{y}}(t) = g\big(x_\varepsilon(t_l), \bar{y}(t), 0\big); \qquad \bar{y}(t_l) = z_\varepsilon(t_l). \tag{42}$$

Using the triangle inequality, we obtain for $t \in [t_l, t_{l+1}]$ that

$$\begin{aligned}
d(z_\varepsilon(t), H_{x_\varepsilon(t)}) &\leqslant d(y_\varepsilon(t), H_{x_\varepsilon(t)}) + \|z_\varepsilon(t) - y_\varepsilon(t)\| \\
&\leqslant \|y_\varepsilon(t) - \bar{y}_\varepsilon(t)\| + d(\bar{y}_\varepsilon(t), H_{x_\varepsilon(t)}) \\
&\quad + \|z_\varepsilon(t) - y_\varepsilon(t)\| \\
&\leqslant \|y_\varepsilon(t) - \bar{y}_\varepsilon(t)\| + d(\bar{y}_\varepsilon(t), H_{x_\varepsilon(t_l)}) \\
&\quad + d_{\mathrm{H}}(H_{x_\varepsilon(t_l)}, H_{x_\varepsilon(t)}) + \|z_\varepsilon(t) - y_\varepsilon(t)\|.
\end{aligned} \tag{43}$$

We now find an upper-bound for the terms on the right-hand-side of the last inequality in (43).

$$\begin{aligned}
&\|y_\varepsilon(t) - \bar{y}_\varepsilon(t)\| \\
\overset{(8),(42)}{\Longrightarrow} \quad &\leqslant \frac{1}{\varepsilon}\left\| \int_{t_l}^t \Big(g(\xi_\varepsilon(t_l), y_\varepsilon(s), 0)\Big) - g\big(x_\varepsilon(t_l), \bar{y}_\varepsilon(s), 0\big)\Big) ds \right\| \\
&\leqslant \frac{L}{\varepsilon}\int_{t_l}^t \Big(\|\xi_\varepsilon(t_l) - x_\varepsilon(t_l)\| + \|y_\varepsilon(s) - \bar{y}_\varepsilon(s)\|\Big) ds \\
\overset{(25)}{\Longrightarrow} \quad &\leqslant S_\varepsilon L\, d_l(t_l) + \frac{L}{\varepsilon}\int_{t_l}^t \|y_\varepsilon(s) - \bar{y}_\varepsilon(s)\| ds \\
&\leqslant S_\varepsilon L\, d_l(t_l)\, e^{S_\varepsilon L},
\end{aligned} \tag{44}$$

where the last inequality is obtained by applying the Gronwall-Bellman inequality. To upper-bound the second term on the right-hand-side of (43), we use Assumption 2. Note that $\bar{y}_\varepsilon(t)$ is the solution to (42) and is different from $\phi_b(t/\varepsilon)$ which is the solution to the boundary-layer system (6). Indeed, the signal $\bar{y}_\varepsilon(t)$ is defined such that its value at the time instant t_l, $l \in I_\varepsilon$ is equal to $z_\varepsilon(t_l)$ and changes according to (42) over the interval $[t_l, t_{l+1}]$. However, over this time interval, (42) can be represented as a boundary-layer model of the form (6) with $x_\varepsilon(t_l)$ as the frozen parameter and possibly with a different initial condition. Hence, from Assumption 2 [3], the second term on the right-hand-side of (43) can be bounded as

$$\begin{aligned}
d(\bar{y}_\varepsilon(t), H_{x_\varepsilon(t_l)}) &\leqslant r_y e^{-\beta_y t/\varepsilon} d(\bar{y}_\varepsilon(t_l), H_{x_\varepsilon(t_l)}) \\
&\overset{(42)}{=} r_y e^{-\beta_y t/\varepsilon} d(z_\varepsilon(t_l), H_{x_\varepsilon(t_l)}),
\end{aligned} \tag{45}$$

for $t \in [t_l, t_{l+1}]$. The third term can also be bounded using (9) in Assumption 1 as

$$d_{\mathrm{H}}(H_{x_\varepsilon(t_l)}, H_{x_\varepsilon(t)}) \leqslant L_H\|x_\varepsilon(t_l) - x_\varepsilon(t)\|$$

[3]Note that $\tau = t/\varepsilon$ as we are writing the equations in slow-time domain.

$$\leqslant L_H \int_{t_l}^{t} \|f(x_\varepsilon(s), z_\varepsilon(s), \varepsilon)\| ds$$
$$\leqslant \varepsilon S_\varepsilon L_H P. \tag{46}$$

where we assumed a bound P for the norm of f as explained in (16). Finally, the last term is bounded using (26) and Lemma 2,

$$\|z_\varepsilon(t) - y_\varepsilon(t)\| \leqslant \bar{D}(\varepsilon). \tag{47}$$

So (43) can be written as

$$d(z_\varepsilon(t), H_{x_\varepsilon(t)}) \leqslant S_\varepsilon L\, d_l(t_l)\, e^{S_\varepsilon L} + r_y e^{-\beta_y t/\varepsilon} d(z_\varepsilon(t_l), H_{x_\varepsilon(t_l)}) + \varepsilon S_\varepsilon L_H P + \bar{D}(\varepsilon). \tag{48}$$

Define $\hat{D}(\varepsilon)$ as

$$\hat{D}(\varepsilon) := S_\varepsilon L\, d_l(t_l)\, e^{S_\varepsilon L} + \varepsilon S_\varepsilon L_H P + \bar{D}(\varepsilon). \tag{49}$$

Then for $t = t_{l+1}$, (48) can be written as that

$$d(z_\varepsilon(t_{l+1}), H_{x_\varepsilon(t_{l+1})}) \leqslant r_y e^{-\beta_y S_\varepsilon} d(z_\varepsilon(t_l), H_{x_\varepsilon(t_l)}) + \hat{D}(\varepsilon). \tag{50}$$

Choose $\delta_y \in (0, \beta_y)$ and $\bar{\varepsilon} > 0$ such that[4]

$$e^{-\delta_y S_{\bar{\varepsilon}}} \leqslant \frac{1}{r_y}. \tag{51}$$

Then

$$d(z_\varepsilon(t_{l+1}), H_{x_\varepsilon(t_{l+1})}) \leqslant e^{-(\beta_y-\delta_y)S_{\bar{\varepsilon}}} d(z_\varepsilon(t_l), H_{x_\varepsilon(t_l)}) + \hat{D}(\varepsilon), \tag{52}$$

and we obtain by induction for all $l \in I_\varepsilon$ and all $\varepsilon \in (0, \min\{\hat{\varepsilon}, \bar{\varepsilon}\}]$ that

$$\begin{aligned} d(z_\varepsilon(t_{l+1}), H_{x_\varepsilon(t_{l+1})}) &\leqslant e^{-(l+1)(\beta_y-\delta_y)S_\varepsilon} d(z_0, H_{x_0}) \\ &\quad + \hat{D}(\varepsilon) \sum_{k=0}^{l} e^{-k(\beta_y-\delta_y)S_\varepsilon} \\ &= e^{-(l+1)(\beta_y-\delta_y)S_\varepsilon} d(z_0, H_{x_0}) \\ &\quad + \hat{D}(\varepsilon) \frac{1 - e^{-(\beta_y-\delta_y)(l+1)S_\varepsilon}}{1 - e^{-(\beta_y-\delta_y)S_\varepsilon}}, \end{aligned} \tag{53}$$

and also obtain for $t \in [t_l, t_{l+1}]$ that

$$\begin{aligned} d(z_\varepsilon(t), H_{x_\varepsilon(t)}) &\leqslant r_y e^{-\beta_y S_\varepsilon} d(z_\varepsilon(t_l), H_{x_\varepsilon(t_l)}) + \hat{D}(\varepsilon) \\ \overset{(53)}{\Longrightarrow} &\leqslant r_y e^{-\beta_y S_\varepsilon} e^{-l(\beta_y-\delta_y)S_\varepsilon} d(z_0, H_{x_0}) \\ &\quad + \hat{D}(\varepsilon) r_y e^{-\beta_y S_\varepsilon} \frac{1 - e^{-(\beta_y-\delta_y)lS_\varepsilon}}{1 - e^{-(\beta_y-\delta_y)S_\varepsilon}} + \hat{D}(\varepsilon) \\ &\leqslant r_y e^{-(\beta_y-\delta_y)t/\varepsilon} d(z_0, H_{x_0}) \\ &\quad + \hat{D}(\varepsilon) r_y e^{-\beta_y S_\varepsilon} \frac{1 - e^{-(\beta_y-\delta_y)lS_\varepsilon}}{1 - e^{-(\beta_y-\delta_y)S_\varepsilon}} + \hat{D}(\varepsilon), \end{aligned} \tag{54}$$

where we used $l = t_l/(\varepsilon S_\varepsilon)$ and $t_l \leqslant t \leqslant t_{l+1}$. Define $F(\varepsilon)$ as

$$F(\varepsilon) := \hat{D}(\varepsilon) \left(1 + \frac{r_y e^{-\beta_y S_\varepsilon}}{1 - e^{-(\beta_y-\delta_y)S_\varepsilon}}\right). \tag{55}$$

Then we obtain that

$$d(z_\varepsilon(t), H_{x_\varepsilon(t)}) \leqslant r_y e^{-(\beta_y-\delta_y)t/\varepsilon} d(z_0, H_{x_0}) + F(\varepsilon), \tag{56}$$

where $\lim_{\varepsilon\to 0} F(\varepsilon) = 0$. The proof of the first part of the theorem is complete.

(ii) In order to prove the equality (22), we show that each term of $K(\varepsilon)$, defined by (41), is of order $O(\sqrt{\varepsilon})$. Since P, v_l, L_{av} and T are bounded, it suffices to show $\bar{\Delta}(\varepsilon)$ and $\bar{\gamma}(S_\varepsilon)$ are of order $O(\sqrt{\varepsilon})$. From Lemma 2, $\bar{\Delta}(\varepsilon)$ is of order $O(\sqrt{\varepsilon})$. Furthermore, we obtain using Lemma 1 that

$$\begin{aligned} \lim_{\varepsilon\to 0} \frac{\bar{\gamma}(S_\varepsilon)}{\sqrt{\varepsilon}} &\overset{(18)}{=} \lim_{\varepsilon\to 0} \bar{\gamma}(S_\varepsilon) S_\varepsilon^2 e^{2TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)} \\ \overset{(21)}{\Longrightarrow} &\leqslant \lim_{\varepsilon\to 0} S_\varepsilon^2 e^{-(\alpha_1-2)TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)} \\ &= 0, \end{aligned} \tag{57}$$

which means $\bar{\gamma}(S_\varepsilon)$ is of order $O(\sqrt{\varepsilon})$. Hence, equality (22) holds uniformly for $t \in [0, T]$ and for all $\varepsilon \in (0, \varepsilon^*]$.

In order to show the equality (23), we have from (10) and (56) that

$$\begin{aligned} &\left| d(z_\varepsilon(t), H_{x_\varepsilon(t)}) - d(\phi_b(t/\varepsilon), H_{x_0}) \right| \\ &\quad \leqslant r_y e^{-(\beta_y-\delta_y)t/\varepsilon} d(z_0, H_{x_0}) + F(\varepsilon) + r_y e^{-\beta_y t/\varepsilon} d(z_0, H_{x_0}) \\ &\quad \leqslant 2 r_y e^{-(\beta_y-\delta_y)t/\varepsilon} d(z_0, H_{x_0}) + F(\varepsilon). \end{aligned} \tag{58}$$

For any given $t_a \in (0, T)$, choose ε^{**} such that[5]

$$(\beta_y - \delta_y) t_a \geqslant \varepsilon^{**} \ln\left(\frac{1}{\sqrt{\varepsilon^{**}}}\right). \tag{59}$$

Then for all $\varepsilon \in [0, \varepsilon^{**}]$, we have

$$e^{-(\beta_y-\delta_y)t/\varepsilon} \leqslant e^{-(\beta_y-\delta_y)t_a/\varepsilon} \leqslant \sqrt{\varepsilon}, \tag{60}$$

which means the first term of (58) is of order $O(\sqrt{\varepsilon})$. Furthermore, the term $F(\varepsilon)$, defined in (55), is of order $O(\sqrt{\varepsilon})$ as $\hat{D}(\varepsilon) = O(\sqrt{\varepsilon})$ (the proof is similar to the proof of Lemma 2) and

$$\lim_{\varepsilon\to 0} \left(1 + \frac{r_y e^{-\beta_y S_\varepsilon}}{1 - e^{-(\beta_y-\delta_y)S_\varepsilon}}\right) = 1 < \infty. \tag{61}$$

This completes the proof. □

3.2. Practical exponential stability

In this subsection, we study the stability of (4) over the infinite time interval. To this end, we require an additional stability condition on the reduced system which is stated below.

[4] Note that there always exist some δ_y and $\bar{\varepsilon}$ that satisfy (51) as according to (27a), $lim_{\bar{\varepsilon}\to 0} S_{\bar{\varepsilon}} = 0$ and thus for any r_y, one could choose $\bar{\varepsilon}$ sufficiently small such that $e^{-\delta_y S_{\bar{\varepsilon}}}$ is less than or equal to $1/r_y$.

[5] It is always possible to choose an ε^{**} for any given $t_a > 0$ as $\lim_{\varepsilon^{**}\to 0} \varepsilon^{**} \ln(\frac{1}{\sqrt{\varepsilon^{**}}}) = 0$.

Assumption 5. *The reduced system* (13) *has a unique equilibrium point at the origin which is uniformly exponentially stable with a domain of attraction containing $B_R(0)$. In other words, there are constants $r_x > 0$ and $\beta_x > 0$ such that for all initial conditions $x_0 \in B_R(0)$, the trajectories of* (13) *satisfy*

$$\|x_{av}(t)\| \leqslant r_x e^{-\beta_x t}\|x_0\|. \tag{62}$$

We are interested in conditions under which the exponential stability properties of the boundary-layer system (6) and the reduced system (13), stated respectively in Assumption 2 and Assumption 5, guarantee practical exponential stability of (4). We propose the following theorem.

Theorem 2 (Practical exponential stability)**.** *Adopt the hypothesis of Theorem 1 along with Assumption 5. Then, for each $\delta > 0$, there exists an $\varepsilon^* : 0 < \varepsilon^* \leqslant \hat{\varepsilon}$ such that for $\varepsilon \in (0, \varepsilon^*]$ and for all initial conditions (x_0, z_0) in $B_R(0) \times M$,*

$$\begin{aligned} \|x_\varepsilon(t)\| &\leqslant r_1 e^{-\beta_1 t}\|x_0\| + \delta, \\ d(z_\varepsilon(t), H_{x_\varepsilon(t)}) &\leqslant r_2 e^{-\beta_2 t/\varepsilon} d(z_0, H_{x_0}) + \delta, \end{aligned} \tag{63}$$

for all $t \geqslant 0$, where $\big(x_\varepsilon(t), z_\varepsilon(t)\big)$ denotes the solution of (4).

Proof. By virtue of Assumption 5, there exist positive constants r_x and β_x such that $\|x_{av}(t)\| \leqslant r_x e^{-\beta_x t}\|x_0\|$. Let δ_x be $\delta_x \in (0, \beta_x)$ and choose T such that [6]

$$e^{-\delta_x T} \leqslant \frac{1}{r_x}. \tag{64}$$

By applying Theorem 1, we have $\|x_\varepsilon(t) - x_{av}(t)\| \leqslant K(\varepsilon)$ for all $t \in [0, T]$. Hence, we can upper bound $\|x_\varepsilon(t)\|$ as

$$\begin{aligned} \|x_\varepsilon(t)\| &\leqslant \|x_\varepsilon(t) - x_{av}(t)\| + \|x_{av}(t)\| \\ &\leqslant r_x e^{-\beta_x t}\|x_0\| + K(\varepsilon), \quad \forall t \in [0, T]. \end{aligned} \tag{65}$$

Using (64) and (65), we have

$$\|x(T)\| \leqslant e^{-(\beta_x - \delta_x)T}\|x_0\| + K(\varepsilon). \tag{66}$$

Since x_0 is assumed to be in $B_R(0)$ (i.e. $\|x_0\| \leqslant R$), we obtain from (66) that $x(T)$ is also in $B_R(0)$ if

$$e^{-(\beta_x - \delta_x)T} R + K(\varepsilon) \leqslant R. \tag{67}$$

We can choose ε_x as the maximum value that satisfies

$$K(\varepsilon_x) \leqslant R\left(1 - e^{-(\beta_x - \delta_x)T}\right). \tag{68}$$

Then for all $\varepsilon \in [0, \varepsilon_x]$, (67) holds and therefore $x(T)$ is in $B_R(0)$. We can now consider $x_\varepsilon(kT)$, $k \in \mathbb{N}$, as a new initial condition for the system for $t \in [kT, (k+1)T]$ and obtain from (66) that

$$\|x_\varepsilon\big((k+1)T\big)\| \leqslant e^{-(\beta_x - \delta_x)T}\|x_\varepsilon(kT)\| + K(\varepsilon). \tag{69}$$

Then we obtain by induction that

$$\begin{aligned} \|x_\varepsilon\big((k+1)T\big)\| &\leqslant e^{-(\beta_x - \delta_x)(k+1)T}\|x_0\| \\ &\quad + K(\varepsilon)\sum_{\ell=0}^{k} e^{-(\beta_x - \delta_x)\ell T} \\ &= e^{-(\beta_x - \delta_x)(k+1)T}\|x_0\| \\ &\quad + K(\varepsilon)\frac{1 - e^{-(\beta_x - \delta_x)(k+1)T}}{1 - e^{-(\beta_x - \delta_x)T}}, \end{aligned} \tag{70}$$

and for any $t \in [kT, (k+1)T]$, $k \in \mathbb{Z}$, we have

$$\begin{aligned} \|x_\varepsilon(t)\| &\leqslant r_x e^{-\beta_x(t - kT)}\|x_\varepsilon(kT)\| + K(\varepsilon) \\ &\leqslant r_x e^{-\beta_x(t - kT)} e^{-(\beta_x - \delta_x)kT}\|x_0\| \\ &\quad + r_x e^{-\beta_x(t - kT)} K(\varepsilon)\frac{1 - e^{-(\beta_x - \delta_x)kT}}{1 - e^{-(\beta_x - \delta_x)T}} + K(\varepsilon) \\ &\leqslant r_x e^{-(\beta_x t - \delta_x kT)}\|x_0\| \\ &\quad + K(\varepsilon)\left(1 + r_x e^{-\beta_x(t - kT)}\frac{1 - e^{-(\beta_x - \delta_x)kT}}{1 - e^{-(\beta_x - \delta_x)T}}\right). \end{aligned} \tag{71}$$

Since $kT \leqslant t \leqslant (k+1)T$ we have $e^{-(\beta_x t - \delta_x kT)} \leqslant e^{-(\beta_x - \delta_x)t}$ and therefore (71) can be written as

$$\begin{aligned} \|x_\varepsilon(t)\| &\leqslant r_x e^{-(\beta_x - \delta_x)t}\|x_0\| \\ &\quad + K(\varepsilon)\left(1 + r_x e^{-\beta_x(t - kT)}\frac{1 - e^{-(\beta_x - \delta_x)kT}}{1 - e^{-(\beta_x - \delta_x)T}}\right). \end{aligned} \tag{72}$$

Define $\bar{K}(\varepsilon)$ as

$$\bar{K}(\varepsilon) := K(\varepsilon)\left(1 + \frac{r_x}{1 - e^{-(\beta_x - \delta_x)T}}\right). \tag{73}$$

Then we obtain for all $t \geqslant 0$ that

$$\|x_\varepsilon(t)\| \leqslant r_x e^{-(\beta_x - \delta_x)t}\|x_0\| + \bar{K}(\varepsilon), \tag{74}$$

where $\lim_{\varepsilon \to 0} \bar{K}(\varepsilon) = 0$.

We now study the behavior of the fast state, $z_\varepsilon(t)$. Using (53) in the proof of Theorem 1, we obtain for $t = T$ (or $l + 1 = T/\varepsilon S_\varepsilon$) that

$$\begin{aligned} d(z_\varepsilon(T), H_{x_\varepsilon(T)}) &\leqslant e^{-(\beta_y - \delta_y)T/\varepsilon} d(z_0, H_{x_0}) \\ &\quad + \hat{D}(\varepsilon)\frac{1 - e^{-(\beta_y - \delta_y)T/\varepsilon}}{1 - e^{-(\beta_y - \delta_y)S_\varepsilon}}. \end{aligned} \tag{75}$$

Similarly to the calculations for the slow state, we first choose ε_z such that for all $\varepsilon \in [0, \varepsilon_z]$, the signal $z_\varepsilon(T)$ does not leave the set M. To this end, we use (75) and the fact [7] that there exists $R_z > 0$ such that $d(z_0, H_{x_0}) \leqslant R_z$, and choose ε_z as the largest value that satisfies

$$\hat{D}(\varepsilon_z)\frac{1 - e^{-(\beta_y - \delta_y)T/\varepsilon_z}}{1 - e^{-(\beta_y - \delta_y)S_{\varepsilon_z}}} \leqslant R_z\left(1 - e^{-(\beta_y - \delta_y)T/\varepsilon_z}\right). \tag{76}$$

[6] Note there always exists T that satisfies (64) which might be large if r_x is large or if δ_x is too small.

[7] It is assumed that the initial condition z_0 and the set H_{x_0} are in the compact sets M and $\bar{M}$, respectively. Since $M \subset \bar{M}$, the distance $d(z_0, H_{x_0})$ is bounded.

Then $d(z_\varepsilon(T), H_{x_\varepsilon(T)}) \leqslant R_z$ and we can consider $z\big((k-1)T\big)$ as an initial condition for the system on the time interval $t \in \big[(k-1)T, kT\big]$, $k \in \mathbb{N}$, and obtain by induction that

$$d(z_\varepsilon(kT), H_{x_\varepsilon(kT)}) \leqslant e^{-(\beta_y-\delta_y)kT/\varepsilon} d(z_0, H_{x_0}) + \hat{D}(\varepsilon)\frac{1-e^{-(\beta_y-\delta_y)kT/\varepsilon}}{1-e^{-(\beta_y-\delta_y)S_\varepsilon}}. \tag{77}$$

Repeating the derivation that led to (54), it can be shown for $t \in [kT + t_l, kT + t_{l+1}]$, $k \in \mathbb{Z}$, that

$$\begin{aligned} d(z_\varepsilon(t), H_{x_\varepsilon(t)}) &\leqslant r_y e^{-\beta_y S_\varepsilon} d(z_\varepsilon(kT+t_l), H_{x_\varepsilon(kT+t_l)}) + \hat{D}(\varepsilon) \\ &\leqslant r_y e^{-\beta_y S_\varepsilon} e^{-l(\beta_y-\delta_y)S_\varepsilon} d(z_\varepsilon(kT), H_{x_\varepsilon(kT)}) \\ &\quad + \hat{D}(\varepsilon) r_y e^{-\beta_y S_\varepsilon}\frac{1-e^{-(\beta_y-\delta_y)lS_\varepsilon}}{1-e^{-(\beta_y-\delta_y)S_\varepsilon}} + \hat{D}(\varepsilon) \\ \overset{(77)}{\Longrightarrow} &\leqslant r_y e^{-\beta_y S_\varepsilon} e^{-(\beta_y-\delta_y)(l\varepsilon S_\varepsilon + kT)/\varepsilon} d(z_0, H_{x_0}) \\ &\quad + \hat{D}(\varepsilon) r_y e^{-\beta_y S_\varepsilon} e^{-l(\beta_y-\delta_y)S_\varepsilon}\frac{1-e^{-(\beta_y-\delta_y)kT/\varepsilon}}{1-e^{-(\beta_y-\delta_y)S_\varepsilon}} \\ &\quad + \hat{D}(\varepsilon) r_y e^{-\beta_y S_\varepsilon}\frac{1-e^{-(\beta_y-\delta_y)lS_\varepsilon}}{1-e^{-(\beta_y-\delta_y)S_\varepsilon}} + \hat{D}(\varepsilon). \end{aligned} \tag{78}$$

Define $\bar{F}(\varepsilon)$ as

$$\bar{F}(\varepsilon) := \hat{D}(\varepsilon)\left\{1 + \frac{2r_y e^{-\beta_y S_\varepsilon}}{1-e^{-(\beta_y-\delta_y)S_\varepsilon}}\right\}, \tag{79}$$

and note that for $kT + t_l \leqslant t \leqslant kT + t_{l+1}$ (i.e. $kT + l\varepsilon S_\varepsilon \leqslant t \leqslant kT + (l+1)\varepsilon S_\varepsilon$), we have $e^{\delta_y(kT+l\varepsilon S_\varepsilon)/\varepsilon} \leqslant e^{\delta_y t/\varepsilon}$ and $e^{-\beta_y(kT+(l+1)\varepsilon S_\varepsilon)/\varepsilon} \leqslant e^{-\beta_y t/\varepsilon}$. So we obtain for $t \geqslant 0$ that

$$d(z_\varepsilon(t), H_{x_\varepsilon(t)}) \leqslant r_y e^{-(\beta_y-\delta_y)t/\varepsilon} d(z_0, H_{x_0}) + \bar{F}(\varepsilon), \tag{80}$$

where $\lim_{\varepsilon\to 0}\bar{F}(\varepsilon) = 0$. Given δ, choose $\varepsilon^* \in (0, \min\{\varepsilon_1, \hat{\varepsilon}, \bar{\varepsilon}, \varepsilon_x, \varepsilon_z\}]$, where $\bar{\varepsilon}$ is defined in (51), such that $\bar{K}(\varepsilon^*) \leqslant \delta$ and $\bar{F}(\varepsilon^*) \leqslant \delta$. Note that according to Theorem 1 and the definition of $\hat{D}(\varepsilon)$ given in (49), it is possible to find such an ε^* as $\hat{D}(\varepsilon)$ and $\bar{\Delta}(\varepsilon)$ converge to zero as $\varepsilon \to 0$. Hence, the proof is complete and the singularly perturbed system (4) is practically exponentially stable with $r_1 = r_x$, $r_2 = r_y$, $\beta_1 = \beta_x - \delta_x$ and $\beta_2 = \beta_y - \delta_y$. □

Remark 4. *Grammel has proposed an exponential stability result for delayed singularly perturbed systems in [20]. Compared to [20], we have relaxed the assumption which required that the origin is a uniform equilibrium point of the slow system (i.e. $f(0, z, \varepsilon) = 0$ for all $z \in M$) [20, Assumption 2.5]. We further studied the behavior of the fast variable z and showed that the closeness of solution error is of order $O(\sqrt{\varepsilon})$.*

4. Simulations

In this section, we present a numerical example in which the solutions of the boundary-layer system converge to a limit cycle. Consider the following system

$$\dot{x} = -x - \left(\big(z_1 + \sin(x)\big)^2 + z_2^2\right)(x + \varepsilon)$$

$$\begin{aligned} \varepsilon\dot{z}_1 &= z_2 + \varepsilon x \\ \varepsilon\dot{z}_2 &= -z_1 - \sin(x). \end{aligned} \tag{81}$$

By writing (81) in τ-domain and letting $\varepsilon = 0$, the boundary-layer system can be written as

$$\frac{dx}{d\tau} = 0 \tag{82a}$$

$$\frac{dz_1}{d\tau} = z_2 \tag{82b}$$

$$\frac{dz_2}{d\tau} = -z_1 - \sin(x). \tag{82c}$$

The solution to (82) from initial conditions $z_1(0)$ and $z_2(0)$ is $\varphi_b(\tau) = [z_1(\tau),\ z_2(\tau)]^T$ where

$$\begin{aligned} z_1(\tau) &= c_1\cos(\tau) + c_2\sin(\tau) - \sin(x) \\ z_2(\tau) &= -c_1\sin(\tau) + c_2\cos(\tau), \end{aligned} \tag{83}$$

in which x is constant according to (82a) and c_1 and c_2 are defined as

$$\begin{aligned} c_1 &= z_1(0) + \sin(x) \\ c_2 &= z_2(0). \end{aligned} \tag{84}$$

Using (83) and (84), the set H_x to which the solutions of the boundary-layer system converge is

$$H_x = \Big\{z_1, z_2 :\ \big(z_1 + \sin(x)\big)^2 + z_2^2 = c_1^2 + c_2^2 = \big(z_1(0) + \sin(x)\big)^2 + z_2^2(0)\Big\}. \tag{85}$$

Note that H_x is parameterized by both x and the initial conditions of the system. Indeed, H_x is a circle in the z-plane whose radius and center depend on x, $z_1(0)$ and $z_2(0)$.

The distance between the boundary-layer solutions and the circles comprising the set H_x, centered at $(-\sin(x), 0)$ with radius $c_1^2 + c_2^2 = \big(z_1(0) + \sin(x)\big)^2 + z_2^2(0)$, can be obtained as

$$d(\varphi_b(\tau), H_x) = \sqrt{\big(z_1(\tau) + \sin(x)\big)^2 + z_2^2(\tau)} - \sqrt{c_1^2 + c_2^2} \overset{(83),(84)}{=} 0. \tag{86}$$

So Assumption 2 holds for any compact set M.

We now define the reduced system and check the validity of the assumptions on this system. From Definition 1, $F_{\mathcal{T}}(x)$ can be written as

$$\begin{aligned} F_{\mathcal{T}}(x) &= \operatorname{conv}\left(\bigcup_{z_0\in M}\left\{\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} f\big(x, \phi_b(s, x, z_0), 0\big)ds\right\}\right) \\ &= -x\left(1 + \operatorname{conv}\left(\bigcup_{z_0\in M}\left\{\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}}(c_1^2 + c_2^2)ds\right\}\right)\right) \\ &= -x\left(1 + \operatorname{conv}\left(\bigcup_{z_0\in M}\{c_1^2 + c_2^2\}\right)\right). \end{aligned} \tag{87}$$

Choose $F_{av}(x)$ as

$$F_{av}(x) = F_{\mathcal{T}}(x). \tag{88}$$

Then Assumption 3 holds with $\gamma = 0$. Also Assumptions 4 and 5 hold as $F_{av}(x)$ is Lipschitz and the reduced average system (13) with $F_{av}(x)$ defined in (88) is exponentially stable.

It is straightforward to choose compact sets $B_R(0)$, $B_{\bar{R}}(0)$, M, $\bar{M}$ and a positive constant ε_1 such that Assumption 1 holds. For example, let $\varepsilon_1 = 0.1$, $R = 1$ and $M = \{z \in \mathbb{R}^2 : 0.5 \leqslant \|z\| \leqslant 1\}$. Then for any given $T > 0$ and for all $t \in [0, T]$, $|x_\varepsilon(t)|$ and $|\xi_\varepsilon(t)|$ are in $B_{\bar{R}}(0)$ with $\bar{R} = 1$, and $z_\varepsilon(t)$ is in $\bar{M} := \{z \in \mathbb{R}^2 : 0.5 \leqslant \|z\| \leqslant 3\}$. Therefore the second condition of Assumption 1 is satisfied. The rest of conditions in Assumption 1 also hold for these sets.

So all conditions of Theorem 2 hold and we conclude that $\|x_\varepsilon(t)\|$ and $d\big(z_\varepsilon(t), H_{x_\varepsilon(t)}\big)$ converge exponentially fast to neighborhoods of zero, and the size of these neighborhoods shrinks to zero as $\varepsilon \to 0$. This is shown in Figure 1 and Figure 2 where the trajectories of (81) are depicted for $\varepsilon = 0.1$ and $\varepsilon = 0.01$. The set $H_{x_\varepsilon(t)}$ in Figure 2 is

$$H_{x_\varepsilon(t)} = \Big\{z_1, z_2 : \big(z_1 + \sin(x_\varepsilon(t))\big)^2 + z_2^2 = \big(z_1(0) + \sin(x_0)\big)^2 + z_2^2(0)\Big\}. \quad (89)$$

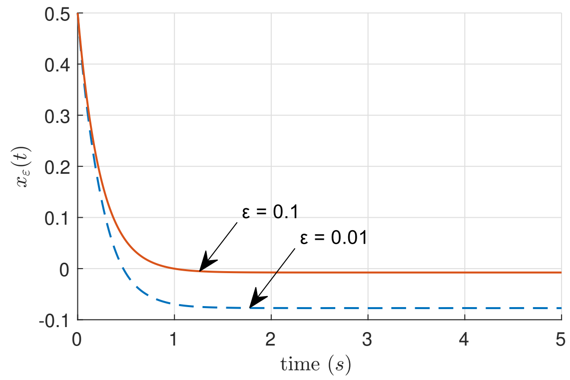

Figure 1: The slow part of the solution of the full-order system (4) for different values of ε.

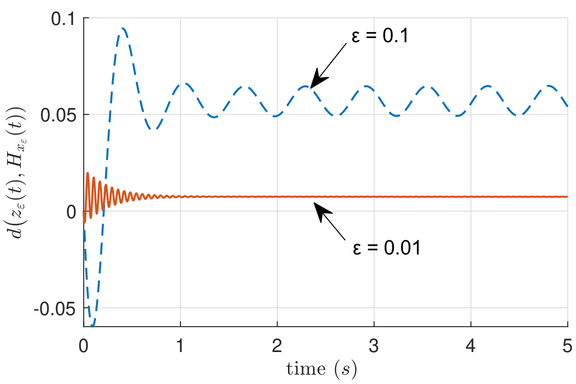

Figure 2: The distance between the the fast part of the solution of the full-order system and the set $H_{x_\varepsilon}(t)$ for different values of ε.

5. Conclusion

In this paper, we have studied the behavior of a general singularly perturbed system with solutions of the boundary-layer system converging exponentially fast to a bounded set. We used averaging to eliminate the fast oscillations of the fast state, and presented results on the behavior of the singularly perturbed system based on the behavior of the average and the boundary-layer system.

6. Appendix

Proof of Lemma 1

Consider the definition of S_ε in (18) and note that as ε goes to zero, $S_\varepsilon e^{TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)}$ goes to infinity which implies that S_ε goes to infinity. Therefore $\lim_{\varepsilon\to 0} S_\varepsilon = \infty$.

To show that $\lim_{\varepsilon\to 0} \varepsilon^{1/4} S_\varepsilon = 0$, observe that

$$\lim_{\varepsilon\to 0} \varepsilon^{1/4} S_\varepsilon \overset{(18)}{=} \lim_{\varepsilon\to 0} e^{-TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)}. \quad (90)$$

Since $\lim_{\varepsilon\to 0} S_\varepsilon = \infty$, we obtain that $\lim_{\varepsilon\to 0} e^{-TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)} = 0$.

Proof of Lemma 2.

Consider $\Delta_l(t)$ and $d_l(t)$ defined in (24) and (25) and note there is a bound P on the norm of f (see (16)). Then for $t \in [t_l, t_{l+1}]$, we have

$$\begin{aligned} d_l(t) &= \max_{t_l\leqslant s\leqslant t} \|x_\varepsilon(s) - \xi_\varepsilon(t_l)\| \\ &= \max_{t_l\leqslant s\leqslant t} \|x_\varepsilon(s) - \xi_\varepsilon(s) + \xi_\varepsilon(s) - \xi_\varepsilon(t_l)\| \\ &\leqslant \max_{t_l\leqslant s\leqslant t} \|x_\varepsilon(s) - \xi_\varepsilon(s)\| + \max_{t_l\leqslant s\leqslant t} \|\xi_\varepsilon(s) - \xi_\varepsilon(t_l)\| \\ &\leqslant \Delta_l(t) + \max_{t_l\leqslant s\leqslant t} \int_{t_l}^{t} \|f(\xi_\varepsilon(t_l), y_\varepsilon(s), 0)\| ds \\ &\leqslant \Delta_l(t) + \varepsilon S_\varepsilon P. \end{aligned} \quad (91)$$

From (4b) and (8) we have

$$\|z_\varepsilon(t) - y_\varepsilon(t)\| = \frac{1}{\varepsilon} \left\| \int_{t_l}^{t} \Big(g(x_\varepsilon(s), z_\varepsilon(s), \varepsilon) - g(\xi_\varepsilon(t_l), y_\varepsilon(s), 0) \Big) ds \right\|.$$

Then using the Lipschitz property of g in Assumption 1, we obtain

$$\begin{aligned} D_l(t) &= \max_{t_l\leqslant s\leqslant t} \frac{1}{\varepsilon} \left\| \int_{t_l}^{s} \Big(g(x_\varepsilon(s), z_\varepsilon(s), \varepsilon) - g(\xi_\varepsilon(t_l), y_\varepsilon(s), 0) \Big) ds \right\| \\ &\leqslant \max_{t_l\leqslant s\leqslant t} \frac{L}{\varepsilon} \int_{t_l}^{s} \Big(\|x_\varepsilon(s) - \xi_\varepsilon(t_l)\| + \|z_\varepsilon(s) - y_\varepsilon(s)\| + \varepsilon \Big) ds \\ &\xRightarrow{(25),(26)} \leqslant S_\varepsilon L\big(d_l(t) + \varepsilon\big) + \frac{L}{\varepsilon} \int_{t_l}^{t} D_l(s) ds, \end{aligned} \quad (92)$$

and by applying the Gronwall-Bellman inequality [12, Lemma A.1] we have

$$D_l(t) \leqslant S_\varepsilon L\big(d_l(t) + \varepsilon\big) e^{S_\varepsilon L}. \quad (93)$$

Also from (4a) and (7) we have

$$\max_{t_l \leqslant s \leqslant t} \|x_\varepsilon(s) - \xi_\varepsilon(s)\| \leqslant \|x_\varepsilon(t_l) - \xi_\varepsilon(t_l)\| + \max_{t_l \leqslant s \leqslant t} \left\| \int_{t_l}^{s} \Big(f\big(x_\varepsilon(s), z_\varepsilon(s), \varepsilon\big) - f\big(\xi_\varepsilon(t_l), y_\varepsilon(s), 0\big) \Big) ds \right\| \tag{94}$$

and thus we obtain using the Lipschitz property of f in Assumption 1 and the Gronwall-Bellman inequality that

$$\begin{aligned}
\Delta_l(t) &\leqslant \Delta_l(t_l) + L \int_{t_l}^{t} \big(d_l(s) + D_l(s) + \varepsilon\big) ds \\
\overset{(93)}{\Longrightarrow} \quad &\leqslant \Delta_l(t_l) + L \int_{t_l}^{t} \big(d_l(s) + \varepsilon\big) \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right) ds \\
\overset{(91)}{\Longrightarrow} \quad &\leqslant \Delta_l(t_l) + \varepsilon S_\varepsilon L\big(\varepsilon S_\varepsilon P + \varepsilon\big)\big(1 + S_\varepsilon L e^{S_\varepsilon L}\big) \\
&\quad + L\big(1 + S_\varepsilon L e^{S_\varepsilon L}\big) \int_{t_l}^{t} \Delta_l(s) ds \\
\overset{\text{Gronwall-Bellman}}{\Longrightarrow} \quad &\leqslant \Big(\Delta_l(t_l) + \varepsilon S_\varepsilon L\big(\varepsilon S_\varepsilon P + \varepsilon\big)\big(1 + S_\varepsilon L e^{S_\varepsilon L}\big)\Big) \\
&\quad e^{\varepsilon S_\varepsilon L\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)}.
\end{aligned} \tag{95}$$

Specifically, for $t = t_{l+1}$ we have

$$\Delta_l(t_{l+1}) \leqslant \Big(\Delta_l(t_l) + \varepsilon S_\varepsilon L(\varepsilon S_\varepsilon P + \varepsilon) \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)\Big) e^{\varepsilon S_\varepsilon L\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)}. \tag{96}$$

From the definition of $\Delta_l(t_l)$ in (24), we have $\Delta_l(t_l) \leqslant \Delta_{l-1}(t_l)$ and therefore (96) can be written as

$$\Delta_l(t_{l+1}) \leqslant \Big(\Delta_{l-1}(t_l) + \varepsilon S_\varepsilon L(\varepsilon S_\varepsilon P + \varepsilon) \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)\Big) e^{\varepsilon S_\varepsilon L\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)}. \tag{97}$$

We can now find an expression for $\Delta_l(t_{l+1})$ using the initial value $\Delta_0(t_1)$ which can be upper bounded as follows:

$$\begin{aligned}
\Delta_0(t_1) &= \max_{0 \leqslant s \leqslant t_1} \|x_\varepsilon(s) - \xi_\varepsilon(s)\| \\
&= \max_{0 \leqslant s \leqslant t_1} \left\| \int_0^s \Big(f\big(x_\varepsilon(s), z_\varepsilon(s), \varepsilon\big) - f(\xi_\varepsilon(t_l), y_\varepsilon(s), 0) \Big) ds \right\| \\
&\leqslant 2\varepsilon S_\varepsilon P,
\end{aligned} \tag{98}$$

where we assumed a bound P for the norm of f as explained in (16). So using (96) and (98), we obtain by induction that

$$\Delta_l(t) \leqslant \Delta_l(t_{l+1}) \leqslant \bar{\Delta}(\varepsilon), \quad \forall l \in I_\varepsilon, \tag{99}$$

where $\bar{\Delta}(\varepsilon)$ is defined as (28). Also from (91), (93) and (99), we obtain that

$$D_l(t) \leqslant D_l(t_{l+1}) \leqslant \bar{D}(\varepsilon), \quad \forall l \in I_\varepsilon \tag{100}$$

with $\bar{D}(\varepsilon)$ defined in (29).

To show that $\bar{\Delta}(\varepsilon) = O(\sqrt{\varepsilon})$, we split the right hand side of (28) into the following three terms and show that they are all $O(\sqrt{\varepsilon})$. We use (2) to check the order of magnitude of each of these terms.

$$\begin{aligned}
(i): \quad & \lim_{\varepsilon \to 0} \frac{2\varepsilon S_\varepsilon P e^{TL\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)}}{\sqrt{\varepsilon}} \\
& \overset{(18)}{=} \lim_{\varepsilon \to 0} 2\varepsilon^{1/4} P = 0,
\end{aligned} \tag{101}$$

$$\begin{aligned}
(ii): \quad & \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} TL\varepsilon S_\varepsilon P \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right) e^{TL\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)} \\
& \overset{(18)}{=} P \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} \varepsilon S_\varepsilon \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \frac{1}{\varepsilon^{1/4} S_\varepsilon} \\
& = P \lim_{\varepsilon \to 0} \frac{1}{S_\varepsilon} \varepsilon^{1/4} S_\varepsilon \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \overset{(27)}{=} 0,
\end{aligned} \tag{102}$$

where we used the fact that $\lim_{x \to 0} x \ln \frac{1}{x} = 0$.

$$\begin{aligned}
(iii): \quad & \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} TL\varepsilon \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right) e^{TL\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)} \\
& \overset{(18)}{=} \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} \varepsilon \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \frac{1}{\varepsilon^{1/4} S_\varepsilon} \\
& = \lim_{\varepsilon \to 0} \frac{1}{(S_\varepsilon)^2} \varepsilon^{1/4} S_\varepsilon \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \overset{(27)}{=} 0.
\end{aligned} \tag{103}$$

We now show that $\bar{D}(\varepsilon) = O(\sqrt{\varepsilon})$. We obtain from (18) that

$$S_\varepsilon L e^{S_\varepsilon L} = \frac{1}{TL} \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) - 1. \tag{104}$$

Similarly to the above calculations for $\bar{\Delta}(\varepsilon)$, it can be shown using (104) that the term $(\varepsilon S_\varepsilon P + \varepsilon) S_\varepsilon L e^{S_\varepsilon L}$ in (29) is $O(\sqrt{\varepsilon})$. We show below that $S_\varepsilon L e^{S_\varepsilon L} \bar{\Delta}(\varepsilon) = O(\sqrt{\varepsilon})$. Equation (104) implies that $S_\varepsilon L e^{S_\varepsilon L} \bar{\Delta}(\varepsilon) = \frac{1}{TL} \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \bar{\Delta}(\varepsilon) - \bar{\Delta}(\varepsilon)$. Given (28), we split $\ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \bar{\Delta}(\varepsilon)$ into the following three terms and show they are $O(\sqrt{\varepsilon})$

$$\begin{aligned}
(i): \quad & \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} 2 \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \varepsilon S_\varepsilon P e^{TL\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)} \\
& \overset{(18)}{=} 2P \lim_{\varepsilon \to 0} \varepsilon^{1/4} \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \\
& = 2P \lim_{\varepsilon \to 0} \frac{1}{S_\varepsilon} \varepsilon^{1/4} S_\varepsilon \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) \overset{(27)}{=} 0.
\end{aligned} \tag{105}$$

$$\begin{aligned}
(ii): \quad & \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) TL\varepsilon S_\varepsilon P \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right) \\
& \quad e^{TL\left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)} \\
& \overset{(18)}{=} P \lim_{\varepsilon \to 0} \sqrt{\varepsilon} S_\varepsilon \left(\ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right)\right)^2 \frac{1}{\varepsilon^{1/4} S_\varepsilon} \\
& = P \lim_{\varepsilon \to 0} \frac{1}{S_\varepsilon} \varepsilon^{1/4} S_\varepsilon \left(\ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right)\right)^2 \overset{(27)}{=} 0,
\end{aligned} \tag{106}$$

where we used $\lim_{x \to 0} x (\ln \frac{1}{x})^2 = 0$.

$$(iii): \quad \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} \ln\left(\frac{1}{\varepsilon^{1/4} S_\varepsilon}\right) TL\varepsilon \left(1 + S_\varepsilon L e^{S_\varepsilon L}\right)$$

$$
\begin{aligned}
& e^{TL\left(1+S_\varepsilon L e^{S_\varepsilon L}\right)} \\
& \stackrel{(18)}{=} \lim_{\varepsilon \to 0} \sqrt{\varepsilon} \left(\ln \left(\frac{1}{\varepsilon^{1/4} S_\varepsilon} \right) \right)^2 \frac{1}{\varepsilon^{1/4} S_\varepsilon} \\
& = \lim_{\varepsilon \to 0} \frac{1}{(S_\varepsilon)^2} \varepsilon^{1/4} S_\varepsilon \left(\ln \left(\frac{1}{\varepsilon^{1/4} S_\varepsilon} \right) \right)^2 \stackrel{(27)}{=} 0. \quad (107)
\end{aligned}
$$

The proof of Lemma 2 is now complete. □

References

[1] P. Kokotovic, H. K. Khalil, and J. O'reilly, *Singular perturbation methods in control: analysis and design*. Siam, 1999.

[2] J. W. Kimball and P. T. Krein, "Singular perturbation theory for dc–dc converters and application to pfc converters," *IEEE Transactions on Power Electronics*, vol. 23, no. 6, pp. 2970–2981, 2008.

[3] R. Sharma, D. Nešić, and C. Manzie, "Model reduction of turbocharged (tc) spark ignition (si) engines," *IEEE transactions on control systems technology*, vol. 19, no. 2, pp. 297–310, 2010.

[4] F. Bornemann, *Homogenization in time of singularly perturbed mechanical systems*. Springer, 2006.

[5] E. Bıyık and M. Arcak, "Area aggregation and time-scale modeling for sparse nonlinear networks," *Systems & Control Letters*, vol. 57, no. 2, pp. 142–149, 2008.

[6] A. Tikhonov, "On the dependence of the solutions of differential equations on a small parameter," *Matematicheskii sbornik*, vol. 64, no. 2, pp. 193–204, 1948.

[7] A. N. Tikhonov, "Systems of differential equations containing small parameters in the derivatives," *Matematicheskii sbornik*, vol. 73, no. 3, pp. 575–586, 1952.

[8] N. Levinson, "Perturbations of discontinuous solutions of non-linear systems of differential equations," *Acta Mathematica*, vol. 82, no. 1, pp. 71–106, 1950.

[9] A. B. Vasil'eva, "Asymptotic behaviour of solutions to certain problems involving non-linear differential equations containing a small parameter multiplying the highest derivatives," *Russian Mathematical Surveys*, vol. 18, no. 3, p. 13, 1963.

[10] F. C. Hoppensteadt, "Singular perturbations on the infinite interval," *Transactions of the American Mathematical Society*, vol. 123, no. 2, pp. 521–535, 1966.

[11] R. O'Malley, "Boundary layer methods for nonlinear initial value problems," *SIAM Review*, vol. 13, no. 4, pp. 425–434, 1971.

[12] H. K. Khalil, *Nonlinear systems*, 3rd ed. Prentice Hall, 2002.

[13] N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic methods in the theory of non-linear oscillations*. CRC Press, 1961.

[14] J. K. Hale, "Ordinary differential equations, robert e," *Krieer, New York*, 1980.

[15] V. Gaitsgory, "Suboptimal control of singularly perturbed systems and periodic optimization," *IEEE Transactions on Automatic Control*, vol. 38, no. 6, pp. 888–903, 1993.

[16] V. Gaitsgory *et al.*, "Averaging and near viability of singularly perturbed control systems," *Journal of Convex Analysis*, vol. 13, no. 2, p. 329, 2006.

[17] V. Gaitsgory and S. Rossomakhine, "Averaging and linear programming in some singularly perturbed problems of optimal control," *Applied Mathematics & Optimization*, vol. 71, no. 2, pp. 195–276, 2015.

[18] G. Grammel, "Singularly perturbed differential inclusions: an averaging approach," *Set-Valued Analysis*, vol. 4, no. 4, pp. 361–374, 1996.

[19] ——, "Averaging of singularly perturbed systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 28, no. 11, pp. 1851–1865, 1997.

[20] ——, "Robustness of exponential stability to singular perturbations and delays," *Systems & Control Letters*, vol. 57, no. 6, pp. 505–510, 2008.

[21] Z. Artstein and A. Vigodner, "Singularly perturbed ordinary differential equations with dynamic limits," *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, vol. 126, no. 3, pp. 541–569, 1996.

[22] Z. Artstein and V. Gaitsgory, "Tracking fast trajectories along a slow dynamics: A singular perturbations approach," *SIAM Journal on Control and Optimization*, vol. 35, no. 5, pp. 1487–1507, 1997.

[23] Z. Artstein, "Stability in the presence of singular perturbations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 34, no. 6, pp. 817–827, 1998.

[24] ——, "Asymptotic stability of singularly perturbed differential equations," *Journal of Differential Equations*, vol. 262, no. 3, pp. 1603–1616, 2017.

[25] A. R. Teel, L. Moreau, and D. Nesic, "A unified framework for input-to-state stability in systems with two time scales," *IEEE Transactions on Automatic Control*, vol. 48, no. 9, pp. 1526–1544, 2003.

[26] W. Wang, A. R. Teel, and D. Nešić, "Analysis for a class of singularly perturbed hybrid systems via averaging," *Automatica*, vol. 48, no. 6, pp. 1057–1068, 2012.

[27] Y. Yang, Y. Lin, and Y. Wang, "Stability analysis via averaging for singularly perturbed nonlinear systems with delays," in *12th IEEE International Conference on Control and Automation (ICCA)*, 2016, pp. 92–97.

[28] M. Deghat, S. Ahmadizadeh, D. Nesic, and C. Manzie, "Closeness of solutions for singularly perturbed systems via averaging," in *2018 57th IEEE Conference on Decision and Control (CDC)*. IEEE, 2018.

[29] V. Gaitsgory and A. Leizarowitz, "Limit occupational measures set for a control system and averaging of singularly perturbed control systems," *Journal of mathematical analysis and applications*, vol. 233, no. 2, pp. 461–475, 1999.

[30] V. Gaitsgory, "On a representation of the limit occupational measures set of a control system with applications to singularly perturbed control systems," *SIAM journal on control and optimization*, vol. 43, no. 1, pp. 325–340, 2004.

[31] V. Gaitsgory and M.-T. Nguyen, "Multiscale singularly perturbed control systems: Limit occupational measures sets and averaging," *SIAM Journal on Control and Optimization*, vol. 41, no. 3, pp. 954–974, 2002.

[32] J.-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory*. Springer Science & Business Media, 1984, vol. 264.

[33] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability theory: new directions*. Springer Science & Business Media, 2011.