Systems & Control Letters 191 (2024) 105881

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Initialization-free Lie-bracket Extremum Seeking ☆

Mahmoud Abdelgalil *, Jorge I. Poveda

Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, 92093, CA, USA

ARTICLE INFO

Keywords:
Extremum seeking
Second-order averaging
Adaptive systems
Zeroth-order optimization

ABSTRACT

Stability results for extremum seeking control in $\mathbb{R}^n$ have predominantly been restricted to local or, at best, semi-global practical stability. Extending semi-global stability results of extremum-seeking systems to unbounded sets of initial conditions often demands a stringent global Lipschitz condition on the cost function, which is rarely satisfied by practical applications. In this paper, we address this challenge by leveraging tools from higher-order averaging theory. In particular, we establish a novel second-order averaging result with *global* (practical) stability implications. By leveraging this result, we characterize sufficient conditions on cost functions under which uniform global (i.e., under any initialization) practical asymptotic stability can be established for a class of extremum-seeking systems acting on static maps. Our sufficient conditions include the case when the gradient of the cost function, rather than the cost function itself, satisfies a global Lipschitz condition, which covers quadratic cost functions. Our results are also applicable to vector fields that are not necessarily Lipschitz continuous at the origin, opening the door to non-smooth Lie-bracket ES dynamics. We illustrate all the results via different analytical and/or numerical examples.

1. Introduction

Extremum Seeking (ES) systems are some of the most popular real-time model-free optimization and stabilization algorithms developed during the last century [1]. The stability and robustness guarantees, simplicity of implementation, and model-agnostic nature of ES make it an attractive option for numerous practical control problems, especially when the plant model is unknown and real-time adaptation and optimization are necessary, see [2–4].

The classical tool for analyzing the stability properties of ES systems is (first-order) averaging theory [5, Ch. 10], [6], which enables local or semi-global practical stability results under mild assumptions [7–10]. These ideas have been extended to study ES systems that emulate Newton-like flows [11,12], as well as ES schemes for control and optimization problems involving delays [13], partial differential equations [14], and hybrid dynamical systems [15–17], among other examples. Recently, ES systems have also been studied via higher-order averaging theory [18–20], which can offer some flexibility in the design and analysis of the exploration–exploitation mechanism in problems that involve geometric constraints [17,21–23], or when additional structure is imposed on the exploration dynamics, e.g. when the exploration is done through a Levi-Civita connection associated with a mechanical system [24–27]. Such tools have led to the discovery of new ES algorithms with desirable properties such as bounded update rates [28], vanishing amplitudes [29,30], and even local exponential/asymptotic stability properties [31].

On the other hand, irrespective of the nature of the averaging tool used for the analysis and design of continuous-time ES algorithms, when the extremum seeking problem is defined on smooth compact boundaryless manifolds, achieving uniform *global* stability results (either practical or asymptotic) is, in general, not possible due to the topological obstructions that apply to continuous-time systems (time-invariant or periodic) evolving on such sets [32], [33, Sec. 4.1]. However, when the ES problem is defined in $\mathbb{R}^n$, such obstructions do not emerge, and, in principle, it might be possible to achieve global extremum seeking. Nevertheless, the majority of results on ES in $\mathbb{R}^n$ have achieved, at best, semi-global practical asymptotic stability [8, 11,18,20,28]. Such results enable convergence from arbitrarily large pre-defined compact sets of initial conditions by appropriately tuning the parameters of the controller. However, without further re-tuning of these parameters, solutions initialized (or pushed via perturbations) outside of these pre-defined compact sets might exhibit finite escape times. Recently, global practical convergence properties were studied in [34] using a normalized scheme, and also in [35] using tools from quasi-stochastic approximation theory. However, results that assert *uniform global practical asymptotic stability* (characterized by, e.g., $\mathcal{KL}$ bounds) in ES controllers remain absent in the literature. One of the

☆ This work was supported in part by the grants National Science Foundation (NSF), United States ECCS CAREER 2305756 and AFOSR, United States YIP: FA9550-22-1-0211.

* Corresponding author.
E-mail addresses: mabdelgalil@ucsd.edu (M. Abdelgalil), poveda@ucsd.edu (J.I. Poveda).

https://doi.org/10.1016/j.sysconle.2024.105881
Received 20 January 2024; Received in revised form 16 May 2024; Accepted 6 July 2024
Available online 30 July 2024

main challenges in achieving such a result using standard averaging theory is the requirement for global Lipschitz conditions in the vector fields of the dynamics (see [5, Ch.10], [17, Sec. 6.1]). This condition is often violated even in the simplest ES problems, which involve cost functions characterized by quadratic maps. This limitation raises the question of whether ES systems can achieve global convergence results in a representative class of problems—a property that could be highly valuable in practical applications by removing any restrictions on the initialization of the algorithms, thus rendering them not only model-free but also "initialization-free".

In this paper, we address the above question and provide a positive answer by showing that certain ES systems can achieve uniform global (practical) stability properties. Such properties are achieved by shifting from *first-order* averaging-based feedback designs, such as those considered in [7,8,15,34], to *second-order* averaging-based feedback designs, akin to those explored in [17,21–25], but utilizing a different averaging tool for the purpose of analysis. In particular, the main contribution of this paper is twofold:

(a) First, we introduce a novel *second-order* averaging theorem with *global* practical stability implications for a class of highly-oscillatory continuous-time systems under appropriate assumptions on the maps involved. For standard (i.e., first-order) averaging, global stability results with applications to control have been studied in [36] for ODEs, and in [17] for hybrid systems. However, to the best of our knowledge, a result of this nature was absent in the literature of second-order averaging. Furthermore, unlike existing results on second-order averaging [18], our results allow for the relaxation of the local Lipschitz condition on the vector field at the origin, requiring only continuity instead. This relaxation opens the door to new non-smooth dynamics that could potentially lead to improved transient performance away from the origin.

(b) Second, we use the aforementioned second-order averaging results to establish uniform global practical asymptotic stability properties for a class of ES systems for which a variety of "typical" cost functions apply, including quadratic maps, and, more generally, strongly convex functions with smooth gradients. However, we also show that convexity of the cost function is, in general, not a necessary condition to achieve global ES under the algorithms studied in this paper. Different analytical and numerical examples are presented to illustrate our results, as well as the limitations and generality of our assumptions.

The rest of the manuscript is organized as follows. We begin by introducing our notation in Section 2. Global averaging results are presented in Section 3. In Section 4, we apply the results of Section 3 to study a class of extremum seeking systems that attain global (practical) stability properties. All the proofs of the results are presented in Section 5. Finally, the conclusions and future work are discussed in Section 6.

2. Preliminaries

2.1. Notation

We use $\mathbb{R}_{\geq 0}$ to denote the set of non-negative real numbers and $\mathbb{R}_{>0}$ to denote the set of positive real numbers. Similarly, we use $\mathbb{Q}_{>0}$ to denote the set of positive rational numbers and $\mathbb{N}_{\geq 1}$ to denote the set of positive integers. The 2-norm of a vector $x \in \mathbb{R}^n$ is denoted by $|x| := \sqrt{x^\top x}$, and the operator 2-norm of a matrix $A \in \mathbb{R}^{m\times n}$ is also denoted as $|A| := \sup\{|Ax| : x \in \mathbb{R}^n, |x| = 1\}$. Given functions $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^l$, we use $g\circ f : \mathbb{R}^n \to \mathbb{R}^l$ to denote their composition, i.e. $g\circ f(x) = g(f(x))$. We use $\mathcal{C}^0$ to denote the class of continuous functions, and $\mathcal{C}^k$ to denote the class of functions that are k-times continuously differentiable, for $k \geq 1$. Given a closed set $K \subset \mathbb{R}^n$, the function f is said to be $\mathcal{C}^k$ on K if there exists an open neighborhood $\mathcal{U} \subset \mathbb{R}^n$ such that $K \subset \mathcal{U}$ and f is $\mathcal{C}^k$ on $\mathcal{U}$. For each $\delta \in \mathbb{R}_{>0}$, we denote the closed ball of radius δ, centered at the origin, by $\delta\mathbb{B}$, i.e. $\delta\mathbb{B} := \{x \in \mathbb{R}^n : |x| \leq \delta\}$. Given a set $\mathcal{A} \subset \mathbb{R}^n$, we use $\mathrm{cl}(\mathcal{A})$ to denote the closure of $\mathcal{A}$ with respect to the natural topology in $\mathbb{R}^n$. When $f \in \mathcal{C}^1$ is a vector-valued map, $\mathrm{D}f$ denotes the Jacobian of f. If $f \in \mathcal{C}^2$ is a real-valued function, then ∇f denotes the gradient of f, and $\nabla^2 f$ is the Hessian of f, i.e. $\nabla^2 f = \mathrm{D}(\nabla f)$. If $f \in \mathcal{C}^1$ and $f = f(x_1,\ldots,x_n)$ is vector-valued, then $\mathrm{D}_{x_i} f$ denotes the Jacobian of f with respect to the ith argument. The map $\pi_i : \mathbb{R}^{n_1}\times\cdots\times\mathbb{R}^{n_k} \to \mathbb{R}^{n_i}$ is the canonical projection onto the x_i-factor, which is defined by $\pi_i(x_1,\ldots,x_k) = x_i$. A class $\mathcal{K}$-function is a strictly increasing continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\alpha(0) = 0$. A class $\mathcal{K}_\infty$-function is a class $\mathcal{K}$-function with the additional requirement that $\lim_{\rho\to+\infty}\alpha(\rho) = +\infty$. A class $\mathcal{KL}$-function $\beta : \mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a continuous function such that, for every $s \in \mathbb{R}_{\geq 0}$, the function $\beta(\cdot, s)$ is a class $\mathcal{K}_\infty$-function, and, for every $r \in \mathbb{R}_{\geq 0}$, the function $\beta(r,\cdot)$ is a strictly decreasing function and $\lim_{s\to+\infty}\beta(r,s) = 0$. To simplify notation, given two (or more) vectors $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}$, we use $(x_1, x_2) \in \mathbb{R}^{n_1+n_2}$ to denote the concatenation of x_1 and x_2. Finally, a map $\Psi : \mathcal{D} \to \mathcal{R}$, where $\mathcal{D},\mathcal{R} \subset \mathbb{R}^n$ are closed, is called a *diffeomorphism* if: (i) it is $\mathcal{C}^1$, (ii) there exists a $\mathcal{C}^1$ map $\Psi^{-1} : \mathcal{R} \to \mathcal{D}$ such that $\Psi^{-1}\circ\Psi(x) = x$, for all $x \in \mathcal{D}$, and $\Psi\circ\Psi^{-1}(x) = x$, for all $x \in \mathcal{R}$.

2.2. Dynamical systems and stability notions

In this paper, we study continuous-time dynamical systems with states $(x,\tau) \in \mathbb{R}^n\times\mathbb{R}_{\geq 0}$, and dynamics

$$\dot{x} = f_\varepsilon(x,\tau), \qquad \dot{\tau} = \varepsilon^{-2}, \tag{1}$$

where $f_\varepsilon : \mathbb{R}^n\times\mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is a continuous function parameterized by a small constant $\varepsilon > 0$. Systems of the form (1) can model highly oscillatory systems that showcase fast variations of τ compared to the state x. For completeness, the notion of solutions to systems of the form (1) is reviewed below.

Definition 1. For $(x_0,\tau_0) \in \mathbb{R}^n\times\mathbb{R}_{\geq 0}$, a function $(x,\tau) : \mathrm{dom}(x,\tau) \to \mathbb{R}^n\times\mathbb{R}_{\geq 0}$ is said to be a solution to (1) from the initial condition (x_0,τ_0) if: (i) there exist $t_s \in \mathbb{R}_{>0}\cup\{\infty\}$ such that $\mathrm{dom}(x,\tau) = [0,t_s)$, (ii) $(x(0),\tau(0)) = (x_0,\tau_0)$, and (iii) the function (x,τ) is $\mathcal{C}^1$ on $\mathrm{dom}(x,\tau)$ and satisfies

$$\frac{dx(t)}{dt} = f_\varepsilon(x(t),\tau(t)), \qquad \frac{d\tau(t)}{dt} = \varepsilon^{-2}.$$

for all $t \in \mathrm{dom}(x,\tau)$. A solution (x,τ) to system (1) is said to be complete if $t_s = \infty$. □

To study the (uniform) stability properties of the parameter-dependent system (1), we will use the following standard notions (see, e.g., [37]), which, without loss of generality, are stated with respect to the origin $x = 0$.

Definition 2. System (1) is said to be *uniformly globally practically asymptotically stable (**UGpAS**) as* $\varepsilon \to 0^+$ if there exists a class $\mathcal{KL}$-function β such that, for every $\nu \in \mathbb{R}_{>0}$, there exists $\varepsilon^* > 0$, such that, for all $\varepsilon \in (0,\varepsilon^*)$, each solution (x,τ) to system (1) from any initial condition $(x_0,\tau_0) \in \mathbb{R}^n\times\mathbb{R}_{\geq 0}$ satisfies

$$|x(t)| \leq \beta(|x_0|, t) + \nu, \tag{2}$$

for all $t \geq 0$. When (2) holds with $\upsilon = 0$, system (1) is said to be *uniformly globally asymptotically stable (**UGAS**)*. □

If the residual upper-bound ν in (2) cannot be controlled by the parameter ε, but the validity of the bound still depends on ε, we will study the following property:

Definition 3. System (1) is said to be *Δ-uniformly globally ultimately bounded* (**Δ-UGUB**) if there exists $\Delta > 0$, $\beta \in \mathcal{KL}$, and $\varepsilon^* \in \mathbb{R}_{>0}$, such that for all $\varepsilon \in (0,\varepsilon^*)$, each solution (x,τ) to system (1) from any initial condition $(x_0,\tau_0) \in \mathbb{R}^n\times\mathbb{R}_{\geq 0}$ satisfies

$$|x(t)| \leq \beta(|x_0|, t) + \Delta, \tag{3}$$

for all $t \geq 0$. □

Note that in Definitions 2 and 3 we do not insist on uniqueness of solutions, but rather impose the appropriate bound (and the property of completeness) to every solution of the system.

3. On global stability via second-order averaging

To study the stability properties of (1) using second-order averaging, we consider a sub-class of systems of the form

$$\dot{x} = f_\varepsilon(x,\tau) = \varepsilon^{-1} f_1(x,\tau) + f_2(x,\tau), \qquad \dot{\tau} = \varepsilon^{-2}, \tag{4}$$

where $f_k : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, $k \in \{1,2\}$, are continuous functions, and $\varepsilon > 0$. Such types of systems commonly emerge in ES [18,30] and vibrational control [38], and they are typically studied via averaging theory. A representative example is given by control-affine systems of the form

$$\dot{x} = \varepsilon^{-1}\left(\sum_{i=1}^{r}\sum_{j=1}^{2} b_{i,j} u_{i,j}\left(J(x),\tau\right)\right) + b_0(x), \qquad \dot{\tau} = \varepsilon^{-2}, \tag{5}$$

where $x \in \mathbb{R}^n$, $r \in \mathbb{N}_{\geq \frac{n}{2}}$, J is an application-dependent C^2 cost function to be minimized, $b_{i,j}$ are suitable vectors, $u_{i,j}(\cdot,\cdot)$ is a scalar-valued feedback law to be designed, and $\varepsilon > 0$ is a small tunable parameter, see Fig. 1 for a block representation of these systems. Particular examples of functions $b_0, b_{i,j}, u_{i,j}$ and J will be discussed later in Section 4.

The stability properties of system (4) will be studied using a change of coordinates induced by a suitable diffeomorphism. In particular, under the action of a diffeomorphism $\Psi : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, a solution (x,τ) to system (4) is transformed into a new function $\Psi \circ (x,\tau) : \mathrm{dom}(x,\tau) \to \mathbb{R}^n \times \mathbb{R}_{\geq 0}$ that is a solution (see Lemma 11 in the appendix) to the following system:

$$\dot{x} = \Psi_* f_\varepsilon(x,\tau), \qquad \dot{\tau} = \varepsilon^{-2}, \tag{6a}$$

where, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, the map $\Psi_* f_\varepsilon$ is given by

$$\Psi_* f_\varepsilon = \left(\mathrm{D}_x\left(\pi_1 \circ \Psi\right) \circ \Psi^{-1}\right) f_\varepsilon \circ \Psi^{-1} + \varepsilon^{-2} \mathrm{D}_\tau\left(\pi_1 \circ \Psi\right) \circ \Psi^{-1}, \tag{6b}$$

which is continuous by construction. System (6) is called the *pushforward* of system (4) under the action of Ψ.

3.1. A global practical near-identity transformation

Traditionally, the averaging-based analysis of oscillatory systems relies on the construction of a suitable (first-order) "near-identity" transformation that maps the original dynamics into a perturbed version of the so-called average dynamics, see [5, Ch.10]. Therefore, to study the global stability properties of (4), we first construct a similar "second-order" near-identity transformation of global nature, which we denote by Ψ. Then, we show how to use Ψ as a diffeomorphism such that the pushforward of the ODE (4) under Ψ (cf. the Eqs. (6)) is a perturbed version of the average dynamics of (4).

We begin by imposing some regularity conditions on f_k:

Assumption 1. There exists $\delta_1 \in [0,\infty)$ such that, for all $k \in \{1,2\}$, the following conditions hold

(a) The map f_k is C^0 in $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$, and there exist a positive constant L_k such that

$$|f_k(x_1,\tau) - f_k(x_2,\tau)| \leq L_k |x_1 - x_2|,$$

for all $x_1, x_2 \in \{x \in \mathbb{R}^n : |x| \geq \delta_1\}$ and all $\tau \in \mathbb{R}_{\geq 0}$.

(b) There exists $T \in \mathbb{R}_{>0}$ such that

$$f_k(x,\tau+T) = f_k(x,\tau), \qquad \int_0^T f_1(x,\tau)d\tau = 0,$$

for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$.

(c) The map f_k is C^{3-k} with respect to x in the domain $\{x \in \mathbb{R}^n : |x| \geq \delta_1\}$.

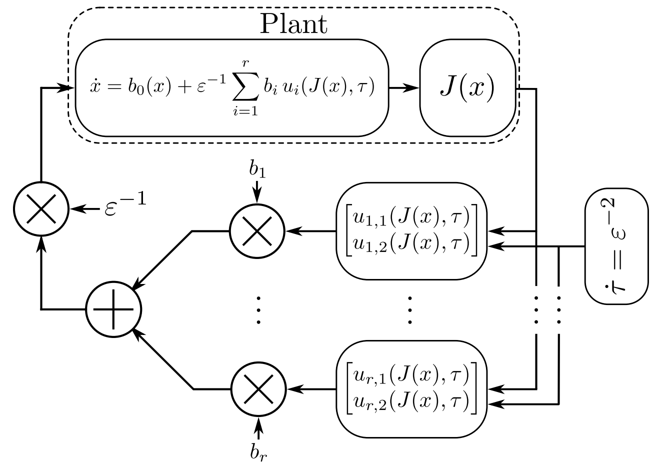

Fig. 1. Block diagram description of system (5). In the diagram, the matrix $b_i = [b_{i,1}, b_{i,2}]$ multiplies the vector $u_i(J(x),\tau) = (u_{i,1}(J(x),\tau), u_{i,2}(J(x),\tau))$.

(d) There exists $L_3 > 0$ such that

$$|\mathrm{D}_x f_1(x_1,\tau_1) f_1(x_1,\tau_2) - \mathrm{D}_x f_1(x_2,\tau_1) f_1(x_2,\tau_2)| \leq L_3 |x_1 - x_2|,$$

for all $x_1, x_2 \in \{x \in \mathbb{R}^n : |x| \geq \delta_1\}$ and all $\tau_1, \tau_2 \in \mathbb{R}_{\geq 0}$. □

In Assumption 1, the case $\delta_1 = 0$ is not excluded. However, by allowing for positive values of δ_1 we can consider maps f_k whose regularity drops from being C^{3-k} to merely C^0, as required by item (a) in Assumption 1, near the origin. This opens the door in our analysis to study certain non-smooth ES dynamics that have been show to exhibit suitable local exponential/asymptotic stability properties [29,39].

Next, for the purpose of analysis, we introduce the auxiliary functions $\chi_j : \mathbb{R} \to \mathbb{R}_{\geq 0}$, where, $j \in \{1,2\}$, given by

$$\chi_1(r) := \begin{cases} \exp\left(-r^{-1}\right) & r > 0, \\ 0 & r \leq 0, \end{cases} \qquad \chi_2(r) := \frac{\chi_1(r)}{\chi_1(r) + \chi_1(1-r)}.$$

Also, let $\delta := (\delta_1, \delta_2, \delta_3) \in \mathbb{R}^3_{\geq 0}$ be a vector of non-negative constants satisfying:

$$\delta_3 > \delta_2, \quad \text{and} \quad \begin{cases} \delta_2 > \delta_1 & \text{if } \delta_1 > 0 \\ \delta_2 = \delta_1 & \text{if } \delta_1 = 0 \end{cases} \tag{7}$$

where the choice of δ_1 will be clear from the context. Using the function χ_2 and the vector δ, we define the smooth "reverse" bump function $\varphi : \mathbb{R}^n \to [0,1]$ as:

$$\varphi(x) = \begin{cases} \chi_2\left(\frac{|x|-\delta_1}{\delta_2-\delta_1}\right) & \delta_2 > \delta_1 \\ 1 & \delta_2 = \delta_1 = 0. \end{cases} \tag{8a}$$

The function φ will be used only for the purpose of analysis, and any similarly defined smooth "reverse" bump function suffices for our purposes. The following Lemma states some useful properties of φ.

Lemma 1. *Let $\delta_2 > \delta_1$. Then, the function φ is C^∞ on $\mathbb{R}^n$, all of its derivatives have the compact support $[\delta_1, \delta_2]$, and it satisfies:*

(a) $\varphi(x) = 1$ for all $x \in \{x' \in \mathbb{R}^n : |x'| \geq \delta_2\}$.
(b) $\varphi(x) = 0$ for all $x \in \{x' \in \mathbb{R}^n : |x'| \leq \delta_1\}$.

Proof. Follows by [40, Lemmas 2.20-2.22] and the construction of the argument of χ_2. □

To state our first result, and using φ, we introduce the auxiliary maps $\hat{f}_k : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, for $k \in \{1,2\}$, defined as

$$\hat{f}_k(x,\tau) := \varphi(x) f_k(x,\tau), \tag{8b}$$

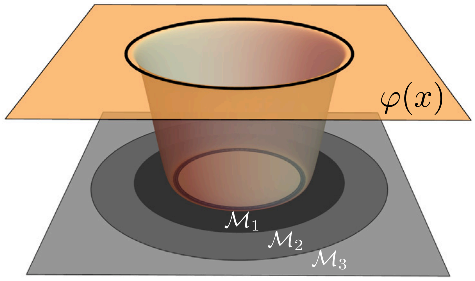

Fig. 2. Visual depiction of φ and the sets $\mathcal{M}_j$ for $j \in \{1,2,3\}$.

as well as the transformation Ψ, defined as

$$\pi_1 \circ \Psi(x,\tau) = \Phi(x,\tau), \qquad \pi_2 \circ \Psi(x,\tau) = \tau, \tag{9a}$$

for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, where the map Φ is defined as follows:

$$\Phi(x,\tau) := x - \varepsilon\, v_1(x,\tau) - \varepsilon^2 v_2(x,\tau), \tag{9b}$$

for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, with

$$v_1(x,\tau) := \int_0^\tau \hat{f}_1(x,s)\,\mathrm{d}s, \tag{9c}$$

$$v_2(x,\tau) := w(x,\tau) - \mathrm{D}_x v_1(x,\tau) v_1(x,\tau), \tag{9d}$$

$$w(x,\tau) := \int_0^\tau \left(\hat{f}_2(x,s) + \mathrm{D}_x \hat{f}_1(x,s) v_1(x,s) - \bar{f}(x)\right) \mathrm{d}s, \tag{9e}$$

and where the *second-order average mapping* $\bar{f}$ is given by

$$\bar{f}(x) := \frac{1}{2T} \int_0^T \left(2\hat{f}_2(x,\tau) + [v_1, \hat{f}_1](x,\tau)\right) \mathrm{d}\tau, \tag{9f}$$

for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, with $T \in \mathbb{R}_{>0}$ being the same constant from Assumption 1, and $[v_1, \hat{f}_1]$ denoting the Lie bracket between the vector v_1 and $\hat{f}_1$, i.e.,

$$[v_1, \hat{f}_1](x,\tau) = \mathrm{D}_x \hat{f}_1(x,\tau) v_1(x,\tau) - \mathrm{D}_x v_1(x,\tau) \hat{f}_1(x,\tau).$$

Remark 1. The map Ψ defined via (9) is an example of a (second-order) *near-identity* transformation [6], which is a standard tool in the averaging literature. The nomenclature stems from the fact that when $\varepsilon = 0$, the transformation Ψ defined by (9) reduces to the identity map on its domain and, by choosing $0 < \varepsilon \ll 1$ sufficiently small, Ψ can be made arbitrarily close to the identity map on bounded subsets of its domain [6, Lemma 2.8.3]. Note that Ψ depends (smoothly) on ε, but we suppress this dependency in the notation for brevity. □

Next, for δ of the form (7), we also consider the closed sets

$$\mathcal{M}_j := \{x \in \mathbb{R}^n : |x| \geq \delta_j\}, \qquad j \in \{1,2,3\}, \tag{10}$$

which satisfy $\mathcal{M}_1 \supseteq \mathcal{M}_2 \supseteq \mathcal{M}_3$. In fact, by construction, the case $\mathcal{M}_1 = \mathcal{M}_2 = \mathbb{R}^n$ can only occur if $\delta_1 = 0$. We illustrate the function φ and the sets $\mathcal{M}_j$, for $j \in \{1,2,3\}$, in Fig. 2.

The following proposition, key for our results, characterizes some useful properties of the map Ψ and the pushforward under Ψ of the vector field (4) (cf. Eqs. (6)).

Proposition 1. *Suppose that Assumption 1 holds, and let δ satisfy (7). Then, there exists $\varepsilon_0, L_\Psi, L_g \in \mathbb{R}_{>0}$, and a C^0 map $g : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times [0,\varepsilon_0] \to \mathbb{R}^n$, such that for all $\varepsilon \in (0,\varepsilon_0)$ the following holds:*

(a) The map $\Psi : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n \times \mathbb{R}_{\geq 0}$ is a diffeomorphism.

(b) The map Ψ and its inverse Ψ^{-1} satisfy:

$$\left|\pi_1 \circ \Psi(0,\tau)\right| \leq L_\Psi \varepsilon,$$
$$\left|\pi_1 \circ \Psi^{-1}(0,\tau)\right| \leq L_\Psi \varepsilon,$$
$$\left|\Psi(x_1,\tau) - \Psi(x_2,\tau)\right| \leq (1 + L_\Psi \varepsilon)|x_1 - x_2|,$$
$$\left|\Psi^{-1}(x_1,\tau) - \Psi^{-1}(x_2,\tau)\right| \leq (1 + L_\Psi \varepsilon)|x_1 - x_2|,$$

for all $x_1, x_2 \in \mathbb{R}^n$, and for all $\tau \in \mathbb{R}_{\geq 0}$.

(c) For all $(x,\tau) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0}$, we have $\Psi^{-1}(x,\tau) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0}$.

(d) The pushforward $\Psi_* f_\varepsilon$ is given by

$$\Psi_* f_\varepsilon(x,\tau) = \bar{f}(x) + \varepsilon\, g(x,\tau,\varepsilon), \tag{11}$$

for all $(x,\tau) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0}$.

(e) The map g satisfies

$$|g(x,\tau,\varepsilon)| \leq L_g(|x| + 1), \tag{12}$$

for all $(x,\tau,\varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times [0,\varepsilon_0]$. □

Proof. See Section 5.1.

Remark 2. Apart from the suitable smoothness and boundedness properties of Ψ, Proposition 1 asserts that the pushforward of system (4) under the action of the diffeomorphism Ψ, i.e., system (6), is given by

$$\dot{x} = \bar{f}(x) + \varepsilon g(x,\tau,\varepsilon), \tag{13}$$

for all $(x,\tau,\varepsilon) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, where $\bar{f}$ is given by (9f). Since Proposition 1 also asserts that the map g is C^0, it is clear that, for $0 < \varepsilon \ll 1$, system (13) can be considered as a perturbed version of the nominal second-order average system

$$\dot{\bar{x}} = \bar{f}(\bar{x}), \qquad \bar{x} \in \mathbb{R}^n, \tag{14}$$

for all $(x,\tau,\varepsilon) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$. By using this relationship, as well as the properties of Ψ, we can inform the stability analysis of system (4) based on the stability properties of the nominal averaged system (14). □

3.2. Global stability via second-order averaging

To study the stability properties of system (4) via averaging, we make the following assumption on the average map $\bar{f}$.

Assumption 2. There exists a vector δ satisfying (7) with the same δ_1 generated by Assumption 1, a C^1 function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, $\alpha_i \in \mathcal{K}_\infty$, $c_i > 0$, for $i \in \{1,2\}$, and a positive definite function $\phi : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, all independent of δ_1, such that the following holds:

(a) For all $x \in \mathbb{R}^n$, we have that

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|), \tag{15a}$$
$$|\nabla V(x)| \leq c_2 \phi(x). \tag{15b}$$

(b) For all $x \in \mathcal{M}_3$, we have that

$$\left\langle \nabla V(x), \bar{f}(x) \right\rangle \leq -c_1 \phi(x)^2. \tag{15c}$$

(c) At least one of the following statements holds:

(i) There exists $\bar{L}_g > 0$, such that

$$|g(x,\tau,\varepsilon)| \leq \bar{L}_g(\phi(x) + 1),$$

for all $(x,\tau,\varepsilon) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0} \times [0,\varepsilon_0]$, where the map g and the constant ε_0 are generated by Proposition 1.

(ii) There exists $\alpha_3 \in \mathcal{K}$, such that $\alpha_3(|x|)|x| \leq \phi(x)$, for all $x \in \mathbb{R}^n$. □

The quadratic-type Lyapunov conditions in items (a) and (b) of Assumption 2 are similar to those studied in the literature of perturbed ODEs [5, Section 9.1]. They imply that the origin is UGUB for the nominal average system (14) [5, Thm. 4.18]. However, without further restrictions on the rate of growth of the norm of the map g in (11)

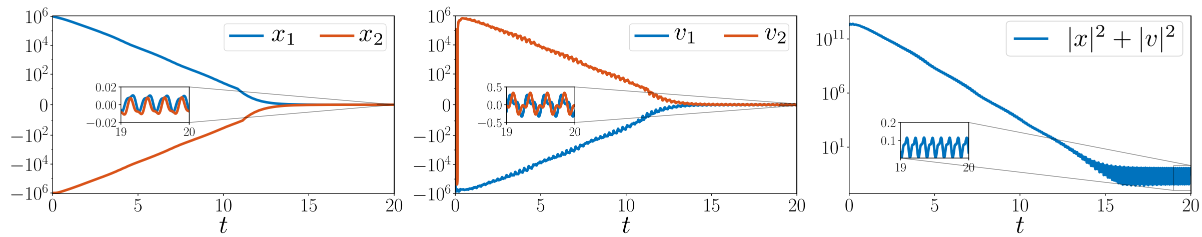

Fig. 3. Numerical results for Example 2. The initial conditions and parameters are: $x(0)=(10^6,-10^6)$, $v(0)=(10^3,-10^3)$, $\varepsilon=\frac{1}{\sqrt{8\pi}}$, $\gamma_1=\frac{3}{4}$, $\gamma_2=1$, and $c=\frac{1}{2}$.

relative to the map ϕ, the perturbation $\varepsilon\, g$ may dominate the average map $\bar{f}$ far from the origin, for any non-zero ε, thereby destroying global stability properties. To preclude this possibility, we impose the additional assumptions in item (c) of Assumption 2, which are discussed in the following remarks:

Remark 3. Item (c)–(i) in Assumption 2 is automatically satisfied whenever the map g is uniformly bounded. As shown later in the proof of Theorem 3, ES systems with bounded vector fields (see, e.g. [38]) satisfy this condition under appropriate assumptions on the cost function. However, note that item (c)–(i) leaves room for unbounded growth of the map g, provided that it can be dominated by the positive definite function ϕ, for all $(x,\tau,\varepsilon)\in\mathcal{M}_3\times\mathbb{R}_{\geq 0}\times[0,\varepsilon_0]$. For example, item (c)–(i) in Assumption 2 automatically holds for the case $\phi(x)=|x|$ thanks to item (e) in Proposition 1.

Remark 4. Item (c)–(ii) in Assumption 2 is automatically satisfied for the case $\phi(x)=|x|$. However, since α_3 is an arbitrary $\mathcal{K}$ function, item (c)–(ii) is a substantial relaxation of the local behavior of the function ϕ on any compact neighborhood of the origin.

By leveraging the previous constructions and Proposition 1, we can now state the first main result of the paper, which applies to the pushforward of system (4). All the proofs are presented in Section 5.

Theorem 1. *Suppose that Assumptions 1–2 hold. Then, there exists $\Delta_\delta>0$ such that system (6) is Δ_δ-UGUB.*

We now provide several useful corollaries of Theorem 1. The first corollary concerns the stability properties of the original system (4).

Corollary 1. *Suppose the assumptions of Theorem 1 hold. Then, there exists $\tilde{\Delta}_\delta>0$ such that system (4) is $\tilde{\Delta}_\delta$-UGUB.*

The following Corollary leverages additional "uniformity" assumptions with respect to the parameters δ (defined as in (7)) to obtain "practical" residual bounds for all the solutions of system (4).

Corollary 2. *Suppose that Assumptions 1 and 2 are satisfied for all δ such that $\delta_1>0$. Then, system (4) is UGpAS as $\varepsilon\to 0^+$.*

Remark 5. Corollary 2 considers the situation in which Assumption 1 is satisfied for each $\delta_1>0$ but might be violated for $\delta_1=0$. Such a situation arises when the vector fields defining system (4) satisfy Assumption 1 on any *closed* subset of the set $\mathbb{R}^n\backslash\{0\}\times\mathbb{R}_{\geq 0}$, but strictly violate the conditions of Assumption 1 on $\mathbb{R}^n\times\mathbb{R}_{\geq 0}$. We illustrate this situation in Example 1 below. □

Example 1. Let $x\in\mathbb{R}$, $\tau\in\mathbb{R}_{\geq 0}$, and consider the dynamical system

$$\dot{x}=-|x|^{\frac{1}{2}}\operatorname{sign}(x)\sin(\tau)^2,\qquad \dot{\tau}=\varepsilon^{-2},\tag{16}$$

which fits the structure of (4) with $f_1=0$ and $f_2(x,\tau)=-|x|^{\frac{1}{2}}\operatorname{sign}(x)\sin(\tau)^2$. For any fixed $\delta_1>0$, there exists a constant $L_{\delta_1}>0$ such that, for all $x_1,x_2\in\mathbb{R}\backslash(-\delta_1,\delta_1)$ and all $\tau\in\mathbb{R}_{\geq 0}$, the function f_2 satisfies $|f_2(x_1,\tau)-f_2(x_2,\tau)|\leq L_{\delta_1}|x_1-x_2|$. Indeed, it can be shown, see Lemma 9 in the Appendix, that the constant may be taken as $L_{\delta_1}=\frac{1}{\sqrt{\delta_1}}+2\sqrt{\delta_1}$. Clearly, the constant $L_{\delta_1}>0$ tends to $+\infty$ in the limit $\delta_1\to 0$. Moreover, it can be shown that there is no constant $L_0>0$ such that, for all $x_1,x_2\in\mathbb{R}$ and all $\tau\in\mathbb{R}_{\geq 0}$, we have $|f_2(x_1,\tau)-f_2(x_2,\tau)|\leq L_0|x_1-x_2|$. Nevertheless, system (16) satisfies Assumption 1 for any $\delta_1>0$ with the constants $L_1=0$ and $L_2=L_{\delta_1}$. Using formula (9f), we obtain that, for any choice of $\delta_1>0$ and δ as in (7), the corresponding nominal averaged system is

$$\dot{\bar{x}}=\bar{f}(x)=-\frac{1}{2}\varphi(\bar{x})|\bar{x}|^{\frac{1}{2}}\operatorname{sign}(\bar{x}),\tag{17}$$

where φ is the function defined in (8a). Using $V(x)=|x|^{\frac{3}{2}}$, which is C^1, and $\phi(x)=|x|^{\frac{1}{2}}$, which is positive definite, we have that V and ϕ satisfy item (a) in Assumption 2 with $c_1=\frac{3}{2}$, and

$$\nabla V(x)=\frac{3}{2}|x|^{\frac{1}{2}}\operatorname{sign}(x),\qquad \left\langle\nabla V(x),\bar{f}(x)\right\rangle=-\frac{3}{4}\varphi(x)\phi(x)^2.\tag{18}$$

Moreover, by construction, for any choice of $\delta_1>0$ and δ satisfying (7), we have $\varphi(x)=1$ for all $|x|\geq\delta_2$. Therefore, V and ϕ satisfy item (b) in Assumption 2 with $c_2=\frac{3}{4}$, for any $\delta_1>0$ and δ satisfying (7). Finally, since $f_1(x,\tau)=0$ and $|f_2(x,\tau)|\leq\phi(x)$, for all $x\in\mathbb{R}$ and all $\tau\in\mathbb{R}_{\geq 0}$, it can be shown (see Lemma 10 in the appendix) that for any $\delta_1>0$ and δ satisfying (7), the function g generated by Proposition 1 satisfies item (c)–(i) with $\bar{L}_g=\sqrt{2}(\bar{L}+L_{v,2})$ where $\bar{L}$ and $L_{v,2}$ are as given in (28). It follows that system (16) satisfies the assumptions of Corollary 2, and we conclude that system (16) is UGpAS. □

The following Corollary considers the situation in which Assumption 1 is also satisfied with $\delta_1=0$ in (7).

Corollary 3. *Suppose that Assumptions 1 and 2 are satisfied for all δ such that $\delta_2=\delta_1=0$. Then, system (4) is UGpAS as $\varepsilon\to 0^+$.*

Below, in Example 2, adapted from [41], we show how Corollary 3 can be used to establish *global* (practical) stabilization via vibrational feedback control in certain systems with unknown control directions.

Example 2. Let $x=(x_1,x_2)\in\mathbb{R}^2$, $v=(v_1,v_2)\in\mathbb{R}^2$, $B\in\mathbb{R}^{2\times 2}$ such that $\operatorname{rank}(B)=2$, and consider the dynamical system

$$\dot{x}=v,\qquad \dot{v}=B\hat{u},\tag{19}$$

where $\hat{u}=(\hat{u}_1,\hat{u}_2)\in\mathbb{R}^2$ is the control input. The goal is to stabilize the equilibrium position $x=v=0$ for system (19) under the assumption that B is unknown. To tackle this problem, we consider a model-free controller inspired by the ES systems studied in [39]. Namely, we let $\varepsilon\in\mathbb{R}_{>0}$, $\tau\in\mathbb{R}_{\geq 0}$, and consider the feedback law:

$$\hat{u}_1=\varepsilon^{-1}u_1(x,v,\tau),\qquad \hat{u}_2=\varepsilon^{-1}u_2(x,v,\tau),\qquad \dot{\tau}=\varepsilon^{-2},\tag{20a}$$

where the functions u_i are given by

$$u_1(x,v,\tau)=\sqrt{2J(x,v)}\cos(\log(J(x,v))+\tau),\tag{20b}$$

$$u_2(x,v,\tau)=\sqrt{4J(x,v)}\cos(\log(J(x,v))+2\tau),\tag{20c}$$

and the function J is taken as

$$J(x,v)=|\gamma_1 x+\gamma_2 v|^2+c,\tag{21}$$

with $c > 0$, where the positive gains γ_1 and γ_2 are tuning parameters. It can be shown that the closed-loop system defined by (19)–(21) satisfies Assumption 1 for $\delta_1 = 0$ (see the proof of Theorem 2 in Section 5.6). Hence, we can use $\delta_1 = \delta_2 = 0$ and let $\delta_3 > 0$ be arbitrary. Using the formula (9f), we obtain that the nominal averaged system is given by

$$\begin{pmatrix} \dot{\bar{x}} \\ \dot{\bar{v}} \end{pmatrix} = A \begin{pmatrix} \bar{x} \\ \bar{v} \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{I} \\ -\gamma_1\gamma_2 BB^\top & -\gamma_2^2 BB^\top \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{v} \end{pmatrix}, \tag{22}$$

for all $(\bar{x}, \bar{v}) \in \mathbb{R}^2 \times \mathbb{R}^2$, which is a linear time-invariant system. If the matrix A in (22) is Hurwitz, then system (22) is UGAS [5, Theorem 4.5] and, by converse Lyapunov theorems [5, Theorem 4.14], it also satisfies Assumption 2 with $\phi(x, v) = |(x, v)|$, and $\alpha_3(r) = \tanh(r)$. Consequently, by invoking Corollary 3 we conclude that the closed-loop system defined by (19)–(21) is UGpAS. Fig. 3 shows the behavior exhibited by the trajectories of the system. In all the simulations, we used $B = (1, 1; 1, -1)$. We remark that, although we treat B as a constant matrix, a similar result can be established when B is time-varying under suitable uniform persistence of excitation conditions, see [41]. □

Remark 6. The proof of Theorem 1 is constructive and provides an explicit form for the upper bound ε^* on the parameter ε (cf. Eq. (44)). However, we remark that, in general, the upper bound ε^* is usually conservative.

4. Applications to extremum seeking systems

In this section, we leverage the averaging results established in Theorem 1 and Corollaries 1–3 to study uniform global practical asymptotic stability (UGpAS) properties in a class of ES systems of the form (5).

4.1. Main assumptions

To guarantee that the (open-loop) amplitudes of the exploration signals in (5) have access to all directions in the parameter space, we impose the following assumption on $b_{i,j}$.

Assumption 3. There exists $\gamma > 0$, such that the vectors $b_{i,j}$ satisfy $\sum_{i=1}^{r} \sum_{j=1}^{2} \left(b_{i,j}^\top v\right)^2 \geq \gamma |v|^2$, for all $v \in \mathbb{R}^n$. □

We also make the following regularity assumption on the cost functions J and the drift term b_0. In all cases, we assume that $J^\star := \inf_{x\in\mathbb{R}^n} J(x) > -\infty$, and that $J^\star = J(0)$. We remark that the assumption $J^\star = J(0)$ is not restrictive since, for the purposes of analysis, if $J^\star = J(x^\star)$ for some $x^\star \in \mathbb{R}^n$, we can always shift the origin of the coordinate system to coincide with the unique minimizer $x^\star$. Similar conditions have been used in the literature [31,39] to analyze ES systems with (local) asymptotic stability properties.

Assumption 4. The following holds:

(a) $J(x) > J(0)$, for all $x \neq 0$.
(b) $\nabla J(x) = 0$ if and only if $x = 0$.
(c) There exists $L_J > 0$ such that $|\nabla^2 J(x)| \leq L_J$, for all $x \in \mathbb{R}^n$.
(d) There exists $\kappa_3 > 0$ such that $|b_0(x)| \leq \kappa_3 |\nabla J(x)|$, for all $x \in \mathbb{R}^n$.
(e) There exists $L_0 > 0$ such that $|b_0(x_1) - b_0(x_2)| \leq L_0 |x_1 - x_2|$, for all $x_1, x_2 \in \mathbb{R}^n$. □

Remark 7. Items (a)–(b) in Assumption 4 are standard in ES problems [7,8]. Similarly, item (c) is equivalent to the assumption that ∇J is L_J-globally Lipschitz [42, Lemma 1.2.2], which is satisfied by, for example, quadratic maps, typical in the study of ES problems [2]. Finally, note that items (d)–(e) are relevant only when the drift term b_0 in (5) is not zero. However, in most ES systems this term is set to zero. □

Next, we characterize two classes of cost functions $J : \mathbb{R}^n \to \mathbb{R}$ that we seek to globally minimize via the dynamics (5).

Assumption 5. The cost J is a radially unbounded C^2-function and there exists $\alpha \in \mathcal{K}$ such that at least one of the following statements holds:

(a) For all $x \in \mathbb{R}^n$, we have

$$\alpha_J(|x|)|x| \leq |\nabla J(x)|.$$

(b) There exists $M_J > 0$ such that

$$\alpha_J(|x|) \leq |\nabla J(x)| \leq M_J,$$

for all $x \in \mathbb{R}^n$. □

Remark 8. As shown in Lemma 7 in the Appendix, item (a) in Assumption 5 is satisfied by any strongly convex C^2-function with a globally Lipschitz gradient. This family of functions includes quadratic cost functions having a positive definite Hessian, which are common in ES. However, as shown in the next example, convexity of the cost function J is not needed to satisfy Assumption 5. □

Example 3. Let $n = 2$, and let the function $J : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$J(x) := |x|^2 + 3\sin(|x|)^2 + 1, \tag{23}$$

which is not convex [43, pp. 4]. However, as shown in Lemma 8 in the Appendix, the function J in (23) satisfies items (a)–(c) of Assumption 4 and item (a) of Assumption 5 with $L_J = 20$ and class-$\mathcal{K}$ function $\alpha_J(s) = 0.5\tanh(s)$. □

The following example considers a cost function J obtained as a regularization of the vector norm function, which satisfies item (b) in Assumption 5.

Example 4. Let $n = 2$, and let the function $J : \mathbb{R}^n \to \mathbb{R}$ be given by

$$J(x) := |x|\tanh(|x|) - 100. \tag{24}$$

It can be directly verified that the function J satisfies items (a)–(c) in Assumption 4 and item (b) in Assumption 5 with $L_J = 3$, $M_J = 2$, and the class-$\mathcal{K}$ function $\alpha_J(s) = \tanh(s)$. □

4.2. ES dynamics with linear growth

We now consider two different algorithms of the form (5) that are able to achieve global ES. The first algorithm that we consider, initially introduced in [39], can be written as (5) with the following functions $u_{i,j}$:

$$u_{i,1}(y,\tau) := \begin{cases} \sqrt{2\omega_i y}\cos(\log(y) + \omega_i\tau) & y > 0 \\ 0 & y \leq 0\,, \end{cases} \tag{25a}$$

$$u_{i,2}(y,\tau) := \begin{cases} \sqrt{2\omega_i y}\sin(\log(y) + \omega_i\tau) & y > 0 \\ 0 & y \leq 0, \end{cases} \tag{25b}$$

where $\omega_i \in \mathbb{Q}_{>0}$, such that $\omega_i \neq \omega_j$ for $i \neq j$. For the sake of convenience, the closed loop system (5) is rewritten here:

$$\dot{x} = b_0(x) + \varepsilon^{-1} \sum_{i=1}^{r} \sum_{j=1}^{2} b_{i,j} u_{i,j}(J(x), \tau), \qquad \dot{\tau} = \varepsilon^{-2}. \tag{26}$$

Clearly, (26) has the same form as (4).

The following theorem is the second main result of this paper.

Theorem 2. *Suppose that Assumptions 3 and 4 hold with $\gamma > \kappa_3$. Then, if item (a) in Assumption 5 holds:*

(a) There exists $\Delta > 0$ such that system (26) is Δ-UGUB under the feedback law (25).
(b) If $J^\star \geq 0$, then system (26) is UGpAS under the feedback law (25). □

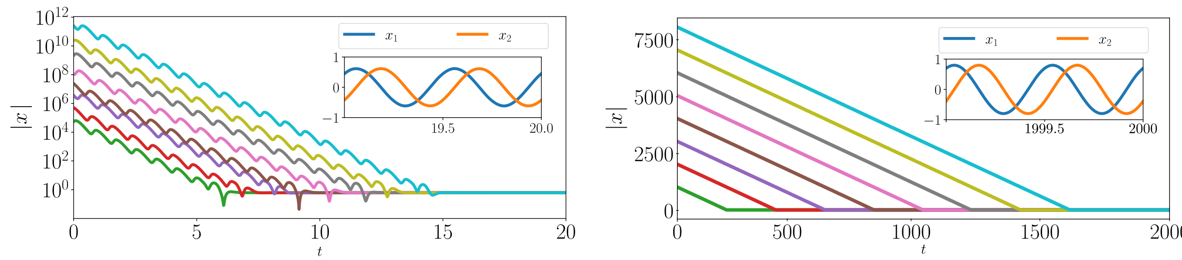

Fig. 4. Numerical results for Example 4 (left) and Example 5 (right). The insets in the top right of the figures depict the quasi-steady state.

The novelty of Theorem 2 compared to existing literature is to establish *uniform global* bounds of the form (2)–(3) for the ES dynamics (5) with feedback law (25). As discussed in Example 3, such bounds can be obtained even when J is not convex.

Example 5 (*Example 3 Continued*). Let $\omega_1 = 1$, $r = 1$, and consider the ES system (5) with destabilizing drift $b_0(x) = \frac{1}{2}x$, and constant vectors $b_{1,1} = (1,0)$, $b_{1,2} = (0,1)$. Notice that in this case $\gamma = 1$, and it can be shown that $\kappa_3 = 0.8$. Therefore, item (d) in Assumption 4 is satisfied. The feedback law is given by (25), with cost function (23). Since all the assumptions of Theorem 2 are satisfied and $J^\star > 0$, we conclude that system (5) is UGpAS. Numerical simulation results are shown in Fig. 4. In the figure, we present simulations obtained from various randomly generated initial conditions and using $\varepsilon = 1/\sqrt{4\pi}$. As shown in the figure, all trajectories converge to a neighborhood of the origin. □

4.3. ES dynamics with bounded control

The second ES algorithm that we consider can also be written as system (5) with the following choice of the functions $u_{i,j}$:

$$u_{i,1}(y,\tau) := \sqrt{2\omega_i}\cos(y + \omega_i\tau), \tag{27a}$$

$$u_{i,2}(y,\tau) := \sqrt{2\omega_i}\sin(y + \omega_i\tau). \tag{27b}$$

The semi-global practical stability properties of these systems have been studied in [28,38]. These algorithms are characterized by uniformly bounded vector fields, which are advantageous for applications with actuator constraints. Note that system (5) with the feedback law (27) can also be written as (26).

The following theorem is the third main result of this paper.

Theorem 3. *Suppose that Assumptions 3 and 4 hold. If $\gamma > \kappa_3$ and item (b) in Assumption 5 is satisfied, then system (26) is UGpAS under the feedback law (27).* □

The novelty of Theorem 3 compared to the results of [28,38], is to establish a global bound of the form (2) for all solutions of the system, albeit under stronger assumptions on the cost functions.

We conclude this section by presenting a numerical example that illustrates the application of Theorem 3.

Example 6 (*Example 4 Continued*). Let $\omega_1 = 1$, $r = 1$, and consider the ES system (5) with $b_0(x) = (0,0)$, $b_{1,1} = (2,0)$, $b_{1,2} = (0,2)$, and the feedback law (27). Notice that in this case $\kappa_3 = 0$ and $\gamma = 2$, so the assumption that $\kappa_3 < \gamma$ holds trivially. We consider the cost function (24), which satisfies the required Assumptions to apply Theorem 3. We simulate the system from randomly generated initial conditions with $\varepsilon = 1/\sqrt{4\pi}$. Numerical simulation results are shown in Fig. 4. Since item (b) in Assumption 5 restricts the gradient to be uniformly bounded, the convergence rate that emerges is slower compared to the convergence rate of the ES dynamics of Example 5. However, Theorem 3 still asserts that system (5) is UGpAS. □

5. Proofs

In this section, we present the proofs of the main results.

5.1. Proof of Proposition 1

First, we introduce several constants which will be used in subsequent steps of the proof. Let δ_1, L_1, L_2, L_3 and T be the constants generated by Assumption 1. For each δ of the form (7) and $k \in \{1,2\}$, define the constants:

$$B_{\varphi,\delta} = \sup_{x\in 3\delta_3\mathbb{B}} |\nabla^2\varphi(x)| \tag{28a}$$

$$B_{k,\delta} := \sup_{(x,\tau)\in 3\delta_3\mathbb{B}\times\mathbb{R}_{\geq 0}} |f_k(x,\tau)|. \tag{28b}$$

$$\hat{L}_k := \begin{cases} L_k + 4B_{k,\delta}(\delta_2-\delta_1)^{-1} & \delta_2 > \delta_1 \\ L_k & \delta_2 = \delta_1 = 0, \end{cases} \tag{28c}$$

$$\hat{L}_3 := \begin{cases} L_3 + 3B_{1,\delta}L_1 + B_{1,\delta}^2 B_{\varphi,\delta} + \frac{4\hat{L}_1 B_{1,\delta}}{\delta_2-\delta_1} & \delta_2 > \delta_1 \\ L_3 & \delta_2 = \delta_1 = 0. \end{cases} \tag{28d}$$

where $B_{\varphi,\delta} < +\infty$ follows from Lemma 1, $B_{k,\delta} < +\infty$ follows from items (a)–(b) in Assumption 1, and the compactness of the closed ball $3\delta_3\mathbb{B}$. Using (28a)–(28d), we also define the following constants:

$$L_{v,1} := T\hat{L}_1, \qquad L_{v,2} := T\left(2\hat{L}_2 + 4T\hat{L}_3\right), \tag{28e}$$

$$\bar{L} := (\hat{L}_2 + 2T\hat{L}_3), \qquad \tilde{L}_\psi := n(L_{v,1} + L_{v,2}), \tag{28f}$$

$$L_\psi := 8\tilde{L}_\psi + 2T\left(B_{1,\delta} + 2B_{2,\delta} + 3T\hat{L}_1 B_{1,\delta}\right), \tag{28g}$$

$$\hat{L}_{v,1} := 2L_{v,1} + TB_{1,\delta}, \tag{28h}$$

$$\hat{L}_{v,2} := 2L_{v,2} + T\left(2B_{2,\delta} + 3T\hat{L}_1 B_{1,\delta}\right). \tag{28i}$$

Next, we define the constant $\varepsilon_0 \in \mathbb{R}_{>0}$ by:

$$\varepsilon_0 := \min\{1, \bar{\varepsilon}_1, \bar{\varepsilon}_2, \bar{\varepsilon}_3, \bar{\varepsilon}_4\}, \tag{29a}$$

where the constants $\bar{\varepsilon}_k$ are given by

$$\bar{\varepsilon}_1 := 4^{-1}n^{-1}(L_{v,1} + L_{v,2})^{-1}, \quad \bar{\varepsilon}_2 := (\hat{L}_{v,1} + \hat{L}_{v,2})^{-1}, \tag{29b}$$

$$\bar{\varepsilon}_3 := \delta_3 L_\psi^{-1}, \quad \bar{\varepsilon}_4 := (\delta_3 - \delta_2)(1+\delta_3)^{-1}(\hat{L}_{v,1} + \hat{L}_{v,2})^{-1}. \tag{29c}$$

Henceforth, and for each δ of the form (7), we shall require that $\varepsilon \in (0, \varepsilon_0)$.

Throughout the proof, we assume that a choice of δ satisfying (7) is fixed. For the sake of clarity, we divide the proof into several key lemmas.

Lemma 2. *Let the assumptions of Proposition 1 be satisfied. Then, for all $k \in \{1,2\}$, the following holds:*

(a) For all $(x,\tau) \in \mathcal{M}_2$, we have that $\hat{f}_k(x,\tau) = f_k(x,\tau)$.
(b) For all $(x,\tau) \in cl\left(\mathbb{R}^n \backslash \mathcal{M}_1\right) \times \mathbb{R}_{\geq 0}$, we have that $\hat{f}_k(x,\tau) = 0$.
(c) For all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, and for the same $T \in \mathbb{R}_{>0}$ from item (b) in Assumption 1, we have that

$$\hat{f}_k(x,\tau+T) = \hat{f}_k(x,\tau), \qquad \int_0^T \hat{f}_1(x,\tau)d\tau = 0.$$

(d) The map $\hat{f}_k$ is C^0 and, for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, the map $\hat{f}_k$ satisfies

$$\left|\hat{f}_k(x_1,\tau) - \hat{f}_k(x_2,\tau)\right| \leq \hat{L}_k|x_1 - x_2|.$$

(e) The map $\hat{f}_k$ is C^{3-k} with respect to x on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$.

(f) For all $(x_1,\tau_1),(x_2,\tau_2) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, the map $D_x\hat{f}_1 \cdot \hat{f}_1$, satisfies

$$|D_x\hat{f}_1(x_1,\tau_1)\cdot \hat{f}_1(x_1,\tau_2) - D_x\hat{f}_1(x_2,\tau_1)\cdot \hat{f}_1(x_2,\tau_2)| \leq \hat{L}_3|x_1 - x_2|.$$

Proof. If $\delta_1 = \delta_2 = 0$, then $\varphi(x) = 1$ and all the conclusions of the proposition follow immediately by (8b) and Assumption 1. Therefore, without loss of generality, we assume that $\delta_2 > \delta_1$.

Proof of items (a), (b), and (c): By Lemma 1, $\varphi(x) = 1$, for all $x \in \mathcal{M}_2$. Hence, by construction, the map $\hat{f}_k$ satisfies $\hat{f}_k(x,\tau) = f_k(x,\tau)$, for all $(x,\tau) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0}$, which proves item (a). Similarly, by Lemma 1, $\varphi(x) = 0$, for all $x \in \delta_1\mathbb{B}$. Hence, by construction, the map $\hat{f}_k$ also satisfies $\hat{f}_k(x,\tau) = 0$, for all $(x,\tau) \in \delta_1\mathbb{B} \times \mathbb{R}_{\geq 0}$, which proves item (b). Item (c) follows directly from the definition of the map $\hat{f}_k$.

Proof of item (d): By Lemma 1, the map $\varphi(x)$ is C^∞. Hence, by item (a) in Assumption 1, the definition of the map $\hat{f}_k$ implies that $\hat{f}_k(\cdot,\cdot)$ is C^0 on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$. Since $\hat{f}_k(x,\tau) = f_k(x,\tau)$, for all $(x,\tau) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0}$, it follows that $\hat{f}_k$ inherits all the properties of f_k in the domain $\mathcal{M}_2 \times \mathbb{R}_{\geq 0}$. In particular, items in Assumption 1 imply that $\mathrm{D}_x f_k$ is well-defined and satisfies the bound $|\mathrm{D}_x f_k(x,\tau)| \leq L_k$, for all $(x,\tau) \in \mathcal{M}_1 \times \mathbb{R}_{\geq 0} \supset \mathcal{M}_2 \times \mathbb{R}_{\geq 0}$. Consequently, $\mathrm{D}_x\hat{f}_k(x,\tau)$ also satisfies the bound $|\mathrm{D}_x\hat{f}_k(x,\tau)| \leq L_k$, for all $(x,\tau) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0}$. Similarly, since $\hat{f}_k(x,\tau) = 0$, for all $(x,\tau) \in \delta_1\mathbb{B} \times \mathbb{R}_{\geq 0}$, it follows that $\mathrm{D}_x\hat{f}_k(x,\tau) = 0$ is well-defined and satisfies $\mathrm{D}_x\hat{f}_k(x,\tau) = 0$, for all $(x,\tau) \in \delta_1\mathbb{B} \times \mathbb{R}_{\geq 0}$.

The definition of the map $\hat{f}_k$ implies that, for all $(x,\tau) \in \mathrm{cl}\left(\mathcal{M}_1\backslash\mathcal{M}_2\right)\times \mathbb{R}_{\geq 0}$, we have

$$|\mathrm{D}_x\hat{f}_k(x,\tau)| \leq |\varphi(x)\mathrm{D}_x f_k(x,\tau)| + |f_k(x,\tau)||\nabla\varphi(x)|.$$

On the other hand, it can be shown that $|\nabla\varphi(x)| \leq 4/(\delta_2 - \delta_1)$, for all $x \in \mathbb{R}^n$. In addition, since $\mathrm{cl}\left(\mathcal{M}_1\backslash\mathcal{M}_2\right) \subset 3\delta_3\mathbb{B} \cap \mathcal{M}_1$, it follows that, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, the Jacobian $\mathrm{D}_x\hat{f}_k$ is well-defined and satisfies

$$|\mathrm{D}_x\hat{f}_k(x,\tau)| \leq \hat{L}_k.$$

Consequently, for all $x_1, x_2 \in \mathbb{R}^n$, and $\forall\tau \in \mathbb{R}_{\geq 0}$, the map $\hat{f}_k$ satisfies

$$|\hat{f}_k(x_1,\tau) - \hat{f}_k(x_2,\tau)| \leq \hat{L}_k|x_1 - x_2|,$$

which concludes the proof of item (d).

Proof of item (e): Since the map φ is C^∞, the definition of the map $\hat{f}_k$ implies that it inherits all the smoothness properties of f_k in the domain $\mathcal{M}_1 \times \mathbb{R}_{\geq 0}$. In particular, item (c) in Assumption 1 implies that $\hat{f}_k(\cdot,\tau)$ is C^{3-k} on the closed set $\mathcal{M}_1$, for all $\tau \in \mathbb{R}_{\geq 0}$. On the other hand, since $\hat{f}_k(x,\tau) = 0$ for all $(x,\tau) \in \delta_1\mathbb{B} \times \mathbb{R}_{\geq 0}$, it follows that $\hat{f}_k(\cdot,\tau)$ is C^∞ on the open set $\mathbb{R}^n\backslash\mathcal{M}_1$, for all $\tau \in \mathbb{R}_{\geq 0}$. Therefore, $\hat{f}_k(\cdot,\tau)$ is C^{3-k} on $\mathbb{R}^n$, for all $\tau \in \mathbb{R}_{\geq 0}$, which proves item (e).

Proof of item (f): For $(x,\tau_1,\tau_2) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, define the maps

$$F(x,\tau_1,\tau_2) := \mathrm{D}_x\hat{f}_1(x,\tau_1)f_1(x,\tau_2),$$
$$\hat{F}(x,\tau_1,\tau_2) := \varphi(x)F(x,\tau_1,\tau_2) = \mathrm{D}_x\hat{f}_1(x,\tau_1)\hat{f}_1(x,\tau_2).$$

Since $\hat{f}_1(\cdot,\tau_1)$ is C^2 on $\mathbb{R}^n$, for all $\tau_1 \in \mathbb{R}_{\geq 0}$, and $f_1(\cdot,\tau_1)$ is C^2 on $\mathcal{M}_1$, for all $\tau_1 \in \mathbb{R}_{\geq 0}$, it follows that the map $F(\cdot,\tau_1,\tau_2)$ is C^1 on $\mathcal{M}_1$, for all $(\tau_1,\tau_2) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$. In addition, since the map φ is C^∞, it follows that the map $\hat{F}(\cdot,\tau_1,\tau_2)$ is also C^1 on $\mathcal{M}_1$, for all $(\tau_1,\tau_2) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$.

From Lemma 1, $\varphi(x) = 1$, for all $x \in \mathcal{M}_2$. Hence, by definition, the map $\hat{F}$ satisfies $\hat{F}(x,\tau_1,\tau_2) = F(x,\tau_1,\tau_2)$, for all $(x,\tau_1,\tau_2) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, which means that $\hat{F}$ inherits all the properties of the map F in the domain $\mathcal{M}_2 \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$. In particular, from items (c) and (d) in Assumption 1, $\mathrm{D}_x F$ is well-defined and satisfies the bound $|\mathrm{D}_x F(x,\tau_1,\tau_2)| \leq L_3$, for all $(x,\tau_1,\tau_2) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, which implies that $\mathrm{D}_x\hat{F}$ also satisfies the bound $|\mathrm{D}_x\hat{F}(x,\tau_1,\tau_2)| \leq L_3$, for all $(x,\tau_1,\tau_2) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$.

From Lemma 1, $\varphi(x) = 0$, for all $x \in \delta_1\mathbb{B}$. Hence, by definition, the map $\hat{F}$ satisfies $\hat{F}(x,\tau_1,\tau_2) = 0$, for all $(x,\tau_1,\tau_2) \in \delta_1\mathbb{B}\times\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}$, which implies that $\hat{F}(\cdot,\tau_1,\tau_2)$ is C^1 on $\mathbb{R}^n$, for all $(\tau_1,\tau_2) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, and that $\mathrm{D}_x\hat{F}$ satisfies $\mathrm{D}_x\hat{F}(x,\tau_1,\tau_2) = 0$, for all $(x,\tau_1,\tau_2) \in \delta_1\mathbb{B} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$.

The definition of the map $\hat{F}$ implies that, for all $(x,\tau) \in \mathrm{cl}\left(\mathcal{M}_1\backslash\mathcal{M}_2\right) \times \mathbb{R}_{\geq 0}$, we have

$$|\mathrm{D}_x\hat{F}(x,\tau_1,\tau_2)| \leq |\varphi(x)\mathrm{D}_x F(x,\tau_1,\tau_2)| + |F(x,\tau_1,\tau_2)||\nabla\varphi(x)^\top|.$$

Recalling the definition of the map F, we obtain that

$$\begin{aligned}\mathrm{D}_x F(x,\tau_1,\tau_2) &= \mathrm{D}_x f_1(x,\tau_1)f_1(x,\tau_2)\nabla\varphi(x)^\intercal + \varphi(x)\mathrm{D}_x(\mathrm{D}_x f_1(x,\tau_1)f_1(x,\tau_2))\\ &+ \mathrm{D}_x f_1(x,\tau_1)\nabla\varphi(x)^\intercal f_1(x,\tau_2) + f_1(x,\tau_1)\nabla\varphi(x)^\intercal\mathrm{D}_x f_1(x,\tau_2)\\ &+ f_1(x,\tau_1)f_1(x,\tau_2)^\intercal\nabla^2\varphi(x),\end{aligned}$$

which leads to the upper bounds

$$|F(x,\tau_1,\tau_2)| \leq \hat{L}_1 B_{1,\delta},$$
$$|\mathrm{D}_x F(x,\tau_1,\tau_2)| \leq L_3 + 3L_1 B_{1,\delta} + B_{1,\delta}^2 B_{\varphi,\delta},$$

for all $(x,\tau_1,\tau_2) \in 3\delta_3\mathbb{B} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$. Therefore, we obtain that

$$|\mathrm{D}_x\hat{F}(x,\tau_1,\tau_2)| \leq L_3 + 3L_1 B_{1,\delta} + B_{1,\delta}^2 B_{\varphi,\delta} + \frac{4\hat{L}_1 B_{1,\delta}}{\delta_2 - \delta_1},$$

for all $(x,\tau_1,\tau_2) \in 3\delta_3\mathbb{B} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, which concludes the proof of the Lemma. ■

Lemma 3. *Let the assumptions of Proposition 1 be satisfied. Then, the following holds:*

(a) The maps $\bar{f}$ and v_k, for $k \in \{1,2\}$, are C^1 on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$.

(b) For all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, the following holds:

$$|\bar{f}(x_1) - \bar{f}(x_2)| \leq \bar{L}|x_1 - x_2|,$$
$$|v_k(x_1,\tau) - v_k(x_2,\tau)| \leq L_{v,k}|x_1 - x_2|, \quad \forall\ k \in \{1,2\}.$$

Proof. We prove each item separately.

Proof of item (a): Since v_1 is the integral of $\hat{f}_1$ with respect to τ, it follows from item (e) in Lemma 2 that the map v_1 is C^2 in x, and C^1 in τ. In addition, since $\bar{f}$ is obtained as the definite integral with respect to τ of the terms $\mathrm{D}_x v_1\,\hat{f}_1$, $\mathrm{D}_x\hat{f}_1\,v_1$, and $\hat{f}_2$, and, from item (e) in Lemma 2, all those of terms are C^1 in x, it follows that $\bar{f}$ is C^1. Moreover, since v_2 is the sum of the term $\mathrm{D}_x v_1 v_1$, which is C^1 in all arguments, and the integral with respect to τ of the terms $\mathrm{D}_x v_1\,\hat{f}_1$, $\mathrm{D}_x\hat{f}_1\,v_1$, and $\hat{f}_2$, which are all, from item (e) in Lemma 2, C^1 in x and C^0 in τ, it follows that v_2 is C^1 in all arguments.

Proof of item (b): From the definition of the map v_1 and item (d) in Lemma 2, we have that

$$\begin{aligned}|v_1(x_1,\tau) - v_1(x_2,\tau)| &= \left|\int_0^\tau \left(\hat{f}_1(x_1,s) - \hat{f}_1(x_2,s)\right)\mathrm{d}s\right|\\ &\leq \hat{L}_1\tau|x_1 - x_2| \leq T\hat{L}_1|x_1 - x_2|,\end{aligned}$$

for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in [0,T]$. In addition, from item (c) in Lemma 2, v_1 is periodic in τ. It follows that, for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have

$$|v_1(x_1,\tau) - v_1(x_2,\tau)| \leq L_{v,1}|x_1 - x_2|.$$

From the definition of v_1, and by interchanging matrix multiplication with the integral, we have that

$$\mathrm{D}_x\hat{f}_1(x,\tau)v_1(x,\tau) = \int_0^\tau \mathrm{D}_x\hat{f}_1(x,\tau)\hat{f}_1(x,s)\mathrm{d}s.$$

From item (f) in Lemma 2, we have that

$$\begin{aligned}&|\mathrm{D}_x\hat{f}_1(x_1,\tau)v_1(x_1,\tau) - \mathrm{D}_x\hat{f}_1(x_2,\tau)v_1(x_2,\tau)|\\ &= \left|\int_0^\tau \left(\mathrm{D}_x\hat{f}_1(x_1,\tau)\hat{f}_1(x_1,s) - \mathrm{D}_x\hat{f}_1(x_2,\tau)\hat{f}_1(x_2,s)\right)\mathrm{d}s\right|\\ &\leq \hat{L}_3\tau|x_1 - x_2| \leq T\hat{L}_3|x_1 - x_2|.\end{aligned}$$

for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in [0, T]$. In addition, from item (c) in Lemma 2, v_1 and $\hat{f}_1$ are periodic in τ. It follows that, for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have that

$$|\mathrm{D}_x \hat{f}_1(x_1,\tau) v_1(x_1,\tau) - \mathrm{D}_x \hat{f}_1(x_2,\tau) v_1(x_2,\tau)| \leq T \hat{L}_3 |x_1 - x_2|.$$

From the definition of v_1, using Leibniz's rule, and by interchanging matrix multiplication with the integral, we have that

$$\mathrm{D}_x v_1(x,\tau) \hat{f}_1(x,\tau) = \int_0^\tau \mathrm{D}_x \hat{f}_1(x,s) \hat{f}_1(x,\tau) \mathrm{d}s.$$

From item (f) in Lemma 2, we have that

$$\begin{aligned} &|\mathrm{D}_x v_1(x_1,\tau) \hat{f}_1(x_1,\tau) - \mathrm{D}_x v_1(x_2,\tau) \hat{f}_1(x_2,\tau)| \\ &\quad = \left| \int_0^\tau \left(\mathrm{D}_x \hat{f}_1(x_1,s) \hat{f}_1(x_1,\tau) - \mathrm{D}_x \hat{f}_1(x_2,s) \hat{f}_1(x_2,\tau) \right) \mathrm{d}s \right| \\ &\quad \leq \hat{L}_3 \tau |x_1 - x_2| \leq T \hat{L}_3 |x_1 - x_2|. \end{aligned}$$

for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in [0, T]$. In addition, from item (c) in Lemma 2, v_1 and $\hat{f}_1$ are periodic in τ. It follows that, for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have that

$$|\mathrm{D}_x v_1(x_1,\tau) \hat{f}_1(x_1,\tau) - \mathrm{D}_x v_1(x_2,\tau) \hat{f}_1(x_2,\tau)| \leq T \hat{L}_3 |x_1 - x_2|.$$

From the definition of v_1, using Leibniz's rule, and interchanging matrix multiplication with the integral, we have that

$$\mathrm{D}_x v_1(x,\tau) v_1(x,\tau) = \int_0^\tau \int_0^\tau \mathrm{D}_x \hat{f}_1(x,s) \hat{f}_1(x,\sigma) \mathrm{d}s \, \mathrm{d}\sigma.$$

From item (f) in Lemma 2, we have that

$$\begin{aligned} &|\mathrm{D}_x v_1(x_1,\tau) v_1(x_1,\tau) - \mathrm{D}_x v_1(x_2,\tau) v_1(x_2,\tau)| \\ &= \left| \int_0^\tau \int_0^\tau \left(\mathrm{D}_x \hat{f}_1(x_1,s) \hat{f}_1(x_1,\sigma) - \mathrm{D}_x \hat{f}_1(x_2,s) \hat{f}_1(x_2,\sigma) \right) \mathrm{d}s \, \mathrm{d}\sigma \right| \\ &\leq \hat{L}_3 \tau^2 |x_1 - x_2| \leq T^2 \hat{L}_3 |x_1 - x_2|, \end{aligned}$$

for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in [0, T]$. In addition, from item (c) in Lemma 2, v_1 is periodic in τ. It follows that, for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have that

$$|\mathrm{D}_x v_1(x_1,\tau) v_1(x_1,\tau) - \mathrm{D}_x v_1(x_2,\tau) v_1(x_2,\tau)| \leq T^2 \hat{L}_3 |x_1 - x_2|.$$

Finally, note that, for all $x_1, x_2 \in \mathbb{R}^n$, we have that

$$\begin{aligned} T|\bar{f}(x_1) - \bar{f}(x_2)| &\leq \int_0^T |\hat{f}_2(x_1,\tau) - \hat{f}_2(x_2,\tau)| \mathrm{d}\tau \\ &\quad + \int_0^T |\mathrm{D}_x v_1(x_1,\tau) \hat{f}_1(x_1,\tau) - \mathrm{D}_x v_1(x_2,\tau) \hat{f}_1(x_2,\tau)| \mathrm{d}\tau \\ &\quad + \int_0^T |\mathrm{D}_x \hat{f}_1(x_1,\tau) v_1(x_1,\tau) - \mathrm{D}_x \hat{f}_1(x_2,\tau) v_1(x_2,\tau)| \mathrm{d}\tau \\ &\quad \leq T \left(\hat{L}_2 + 2T \hat{L}_3 \right) |x_1 - x_2|, \end{aligned}$$

and, for all $x_1, x_2 \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have that

$$\begin{aligned} |v_2(x_1,\tau) - v_2(x_2,\tau)| &\leq \int_0^\tau |\hat{f}_2(x_1,\tau) - \hat{f}_2(x_2,\tau)| \mathrm{d}\tau \\ &\quad + \int_0^\tau |\bar{f}(x_1) - \bar{f}(x_2)| \mathrm{d}\tau \\ &\quad + \int_0^T |\mathrm{D}_x \hat{f}_1(x_1,\tau) v_1(x_1,\tau) - \mathrm{D}_x \hat{f}_1(x_2,\tau) v_1(x_2,\tau)| \mathrm{d}\tau \\ &\quad + |\mathrm{D}_x v_1(x_1,\tau) v_1(x_1,\tau) - \mathrm{D}_x v_1(x_2,\tau) v_1(x_2,\tau)| \\ &\quad \leq T \left(2\hat{L}_2 + 4T \hat{L}_3 \right) |x_1 - x_2|, \end{aligned}$$

The proof of the Lemma is concluded by noting that $\bar{L} = T(\hat{L}_2 + 2\hat{L}_3)$ and $L_{v,2} = T \left(2\hat{L}_2 + 4T \hat{L}_3 \right)$. ■

Lemma 4. *Let the assumptions of Proposition 1 be satisfied. Then, for all $\varepsilon \in [0, \varepsilon_0]$, the following holds:*

(a) Ψ is a diffeomorphism on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$.

(b) For all $\tau \in \mathbb{R}_{\geq 0}$, Ψ and its inverse Ψ^{-1} satisfy

$$\left| \pi_1 \circ \Psi(0,\tau) \right| \leq L_\Psi \varepsilon, \qquad \left| \pi_1 \circ \Psi^{-1}(0,\tau) \right| \leq L_\Psi \varepsilon.$$

(c) For all $x_1, x_2 \in \mathbb{R}^n$ and for all $\tau \in \mathbb{R}_{\geq 0}$, the map Ψ and its inverse Ψ^{-1} satisfy

$$\begin{aligned} \left| \Psi(x_1,\tau) - \Psi(x_2,\tau) \right| &\leq (1 + L_\Psi \varepsilon) |x_1 - x_2|, \\ \left| \Psi^{-1}(x_1,\tau) - \Psi^{-1}(x_2,\tau) \right| &\leq (1 + L_\Psi \varepsilon) |x_1 - x_2|. \end{aligned}$$

(d) For all $(x,\tau) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0}$, $\Psi^{-1}(x,\tau) \in \mathcal{M}_2$.

Proof (*Proof of Item (a)*)**.** Let (x,τ) and $(\tilde{x},\tilde{\tau})$ be any two points in $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$ and suppose that $\Psi(x,\tau) = \Psi(\tilde{x},\tilde{\tau})$. Then, by construction, we have $\tau = \tilde{\tau}$, and

$$|x - \tilde{x}| \leq \sum_{i=1}^{2} \varepsilon^i |v_i(x,\tau) - v_i(\tilde{x},\tau)|.$$

From item (b) in Lemma 3, we obtain that

$$|x - \tilde{x}| \leq \varepsilon (L_{v,1} + L_{v,2}\varepsilon) |x - \tilde{x}|.$$

We note that for all $\varepsilon \in (0, \varepsilon_0)$ we have that $|x - \tilde{x}| \leq \frac{1}{2}|x - \tilde{x}|$, which can only happen if $|x - \tilde{x}| = 0$. Therefore, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, for all $(\tilde{x},\tilde{\tau}) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, and for all $\varepsilon \in (0, \varepsilon_0)$, $\Psi(x,\tau) = \Psi(\tilde{x},\tilde{\tau}) \implies x = \tilde{x}$, and $\tau = \tilde{\tau}$, which in turn implies that the map Ψ is injective on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$. Next, for each $(\tilde{x},\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, define the map $\tilde{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$:

$$\tilde{\Phi}(x) = \tilde{x} + x - \Phi(x,\tau).$$

By direct computation

$$\tilde{\Phi}(x) = \tilde{x} + \sum_{k=1} \varepsilon^k v_k(x,\tau).$$

Now let $x_1, x_2 \in \mathbb{R}^n$ be any two points. It follows that

$$\left| \tilde{\Phi}(x_1) - \tilde{\Phi}(x_2) \right| \leq \sum_{k=1} \varepsilon^k |v_k(x_1,\tau) - v_k(x_2,\tau)|.$$

From item (b) in Lemma 3, we obtain that

$$\left| \tilde{\Phi}(x_1) - \tilde{\Phi}(x_2) \right| \leq \varepsilon (L_{v,1} + L_{v,2}\varepsilon) |x_1 - x_2|.$$

Since $\varepsilon \in (0, \varepsilon_0)$, it follows that $\left| \tilde{\Phi}(x_1) - \tilde{\Phi}(x_2) \right| \leq \frac{1}{2}|x_1 - x_2|$ for all $x_1, x_2 \in \mathbb{R}^n$, which implies that $\tilde{\Phi}$ is a contraction. Thus, $\tilde{\Phi}$ has a unique fixed point [40, Lemma C.35], which implies that for all $\varepsilon \in (0, \varepsilon_0)$, for all $\tilde{x} \in \mathbb{R}^n$ and for all $\tau \in \mathbb{R}_{\geq 0}$, there exists a unique point $x \in \mathbb{R}^n$ such that

$$\tilde{x} = \pi_1 \circ \Psi(x,\tau) = \Phi(x,\tau). \tag{30}$$

In other words, Ψ is onto, and therefore a bijection on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$.

From item (a) in Lemma 3, we know that the map Ψ is C^1 on $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$, and so its Jacobian $\mathrm{D}\Psi$ is well-defined, and given by

$$\mathrm{D}\Psi = \begin{pmatrix} \mathrm{D}_x \Phi & \mathrm{D}_\tau \Phi \\ 0 & 1 \end{pmatrix},$$

where the Jacobians $\mathrm{D}_x \Phi$ and $\mathrm{D}_\tau \Phi$ are given by

$$\mathrm{D}_x \Phi = \mathrm{I} - \sum_{k=1}^{2} \varepsilon^k \mathrm{D}_x v_k, \qquad \mathrm{D}_\tau \Phi = -\sum_{k=1}^{2} \varepsilon^k \mathrm{D}_\tau v_k.$$

For each $i \in \{1, \ldots, n\}$ and $k \in \{1, 2\}$, let $R^i_{v,k} : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be given by

$$R^i_{v,k}(x,\tau) = \sum_{j=1, j \neq i}^{n} \left| \left(D_x v_k(x,\tau) \right)_{ij} \right|,$$

where $\left(D_x v_k(x,\tau) \right)_{ij}$ are the entries of the matrix $D_x v_k(x,\tau)$. From items in Lemma 3, the maps v_k are C^1 and satisfy $\left| D_x v_k(x,\tau) \right| \leq L_{v,k}$, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$ and $k \in \{1, 2\}$. In particular, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, for all $i, j \in \{1, \ldots, n\}$, and for all $k \in \{1, 2\}$, we have that $\left| \left(D_x v_k(x,\tau) \right)_{ij} \right| \leq$

$L_{v,k}$. As such, for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, for all $i\in\{1,\dots,n\}$, and for all $k\in\{1,2\}$, we have that $0\leq R^i_{v,k}(x,\tau)\leq(n-1)L_{v,k}$. Next, note that the entries of $\mathrm{D}_x\Phi$ are given by:

$$\left(\mathrm{D}_x\Phi(x,\tau)\right)_{ii}=1-\sum_{k=1}^{2}\varepsilon^k\left(\mathrm{D}_x v_k(x,\tau)\right)_{ii}, \qquad \forall i\in\{1,\dots,n\}$$

$$\left(\mathrm{D}_x\Phi(x,\tau)\right)_{ij}=-\sum_{k=1}^{2}\varepsilon^k\left(\mathrm{D}_x v_k(x,\tau)\right)_{ij}, \qquad \forall i\neq j\in\{1,\dots,n\},$$

Consequently, for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, for all $\varepsilon\in\mathbb{R}_{\geq 0}$, and for all $i\in\{1,\dots,n\}$, we have

$$\begin{aligned}R^i(x,\tau) &:= \sum_{j\neq i=1}^{n}\left|\left(\mathrm{D}_x\Phi(x,\tau)\right)_{ij}\right|\leq\sum_{k=1}^{2}\varepsilon^k\sum_{j\neq i=1}^{n}\left|\left(\mathrm{D}_x v_k(x,\tau)\right)_{ij}\right|\\ &=\sum_{k=1}^{2}\varepsilon^k R^i_{v,k}(x,\tau)\leq(n-1)\sum_{k=1}^{2}\varepsilon^k L_{v,k}.\end{aligned}$$

Similarly, we have

$$1-\sum_{k=1}^{2}\varepsilon^k L_{v,k}\leq\left(\mathrm{D}_x\Phi(x,\tau)\right)_{ii}\leq 1+\sum_{k=1}^{2}\varepsilon^k L_{v,k}.$$

Therefore, for all $\varepsilon\in(0,\varepsilon_0)$, all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, and all $i\in\{1,\dots,n\}$:

$$\frac{3}{4}\leq 1-\tilde{L}_\Psi\varepsilon\leq\left(\mathrm{D}_x\Phi(x,\tau)\right)_{ii}\leq 1+\tilde{L}_\Psi\varepsilon\leq\frac{5}{4},$$
$$0\leq R^i(x,\tau)\leq\tilde{L}_\Psi\varepsilon\leq\frac{1}{4}.$$

By applying the Geršhgorin circle theorem [44, p. 269], we obtain that for all $\varepsilon\in(0,\varepsilon_0)$, for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, the eigenvalues of the Jacobian matrix $\mathrm{D}_x\Phi(x,\tau)$ are contained in the compact interval $[1-2\tilde{L}_\Psi\varepsilon,1+2\tilde{L}_\Psi\varepsilon]\subset[1/2,3/2]$.

Then, we have the following claim, proved in Appendix B.1.

Claim 1. *For all $\varepsilon\in(0,\varepsilon_0)$, and all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, the Jacobian matrix $D_x\Phi(x,\tau)$ is invertible,*

$$\left|D_x\Phi(x,\tau)\right|\leq 1+4\tilde{L}_\Psi\varepsilon\leq 2,\qquad\left|D_x\Phi(x,\tau)^{-1}\right|\leq 1+4\tilde{L}_\Psi\varepsilon\leq 2,$$

and the Jacobian $D\Psi(x,\tau)^{-1}$ is well-defined and given by

$$D\Psi(x,\tau)^{-1}=\begin{pmatrix}D_x\Phi(x,\tau)^{-1} & -D_x\Phi(x,\tau)^{-1}D_\tau\Phi(x,\tau)\\ 0 & 1\end{pmatrix}.\quad\square \tag{31}$$

For all $\varepsilon\in(0,\varepsilon_0)$, the map Ψ is bijective and, for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, the Jacobian $\mathrm{D}\Psi(x,\tau)$ is invertible and its inverse is continuous. Thus, by invoking the global rank theorem [40, Theorem 4.14], we conclude that, for all $\varepsilon\in(0,\varepsilon_0)$, the map Ψ is a diffeomorphism.

Proof of item (b): Note that by [40, Proposition C.4] the Jacobian of Ψ^{-1} is given by

$$\mathrm{D}\Psi^{-1}(x,\tau)=\left(\mathrm{D}\Psi\circ\Psi^{-1}(x,\tau)\right)^{-1}.$$

From the definition of Ψ, for all $(\tau,\varepsilon)\in\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$, we have that $\pi_1\circ\Psi(0,\tau)=-\sum_{k=1}^{2}\varepsilon^k v_k(0,\tau)$. Since $0\in 3\delta_3\mathbb{B}$, the definition of v_k implies that, for all $\tau\in\mathbb{R}_{\geq 0}$, we have that

$$|v_1(0,\tau)|\leq TB_{1,\delta},\qquad |v_2(0,\tau)|\leq T\left(2B_{2,\delta}+3T\hat{L}_1B_{1,\delta}\right).$$

Define the constant $L_{\Psi,1}$ by

$$L_{\Psi,1}:=4\tilde{L}_\Psi+T\left(B_{1,\delta}+2B_{2,\delta}+3T\hat{L}_1B_{1,\delta}\right). \tag{32}$$

It follows that, for all $(\tau,\varepsilon)\in\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$, we have $|\pi_1\circ\Psi(0,\tau)|\leq L_{\Psi,1}\varepsilon$. Since Ψ is a diffeomorphism, we have

$$0=\pi_1\circ\Psi\circ\Psi^{-1}(0,\tau)=\pi_1\circ\Psi^{-1}(0,\tau)-\sum_{k=1}^{2}\varepsilon^k v_k\circ\Psi^{-1}(0,\tau).$$

By adding and subtracting the term $\sum_{k=1}^{2}\varepsilon^k v_k(0,\tau)$ and invoking the Lipschitz continuity of the maps v_k, we have that

$$\begin{aligned}|\pi_1\circ\Psi^{-1}(0,\tau)| &\leq \left|\sum_{k=1}^{2}\varepsilon^k(v_k\circ\Psi^{-1}(0,\tau)-v_k(0,\tau))\right|+\left|\sum_{k=1}^{2}\varepsilon^k v_k(0,\tau)\right|\\ &\leq\sum_{k=1}^{2}\varepsilon^k L_{v,k}|\pi_1\circ\Psi^{-1}(0,\tau)|+\sum_{k=1}^{2}\varepsilon^k|v_k(0,\tau)|,\end{aligned}$$

which implies that

$$\left(1-\sum_{k=1}^{2}\varepsilon^k L_{v,k}\right)|\pi_1\circ\Psi^{-1}(0,\tau)|\leq\sum_{k=1}^{2}\varepsilon^k|v_k(0,\tau)|.$$

Since $\varepsilon\in(0,\varepsilon_0)$, it follows that

$$|\pi_1\circ\Psi^{-1}(0,\tau)|\leq 2\sum_{k=1}^{2}\varepsilon^k|v_k(0,\tau)|.$$

Recalling the definition of L_Ψ, we obtain that, for all $\tau\in\mathbb{R}_{\geq 0}$, we have

$$|\pi_1\circ\Psi(0,\tau)|\leq L_\Psi\varepsilon,\qquad |\pi_1\circ\Psi^{-1}(0,\tau)|\leq L_\Psi\varepsilon. \tag{33}$$

Proof of item (c): Note that, for all $x_1,x_2\in\mathbb{R}^n$, for all $\tau\in\mathbb{R}_{\geq 0}$, and for all $\varepsilon\in(0,\varepsilon_0)$, we obtain via Hadamard's Lemma [45, Lemma 2.8]

$$\left|\Psi(x_1,\tau)-\Psi(x_2,\tau)\right|\leq\left|\mathrm{J}_\Psi(x_1,x_2,\tau)\right||x_1-x_2|,$$
$$\left|\Psi^{-1}(x_1,\tau)-\Psi^{-1}(x_2,\tau)\right|\leq\left|\mathrm{J}_{\Psi^{-1}}(x_1,x_2,\tau)\right||x_1-x_2|,$$

where the matrix-valued maps J_Ψ and $\mathrm{J}_{\Psi^{-1}}$ are given by

$$\mathrm{J}_\Psi(x_1,x_2,\tau):=\int_0^1\mathrm{D}_x\Phi(x_2+\lambda(x_1-x_2),\tau)\mathrm{d}\lambda,$$
$$\mathrm{J}_{\Psi^{-1}}(x_1,x_2,\tau):=\int_0^1\mathrm{D}_x\Phi^{-1}(x_2+\lambda(x_1-x_2),\tau)\mathrm{d}\lambda,$$

and where we used the shorthand notation

$$\mathrm{D}_x\Phi^{-1}(x,\tau):=\left(\mathrm{D}_x\Phi\circ\Psi^{-1}(x,\tau)\right)^{-1}.$$

It follows that

$$\left|\Psi(x_1,\tau)-\Psi(x_2,\tau)\right|\leq(1+L_\Psi\varepsilon)|x_1-x_2|, \tag{34a}$$
$$\left|\Psi^{-1}(x_1,\tau)-\Psi^{-1}(x_2,\tau)\right|\leq(1+L_\Psi\varepsilon)|x_1-x_2|. \tag{34b}$$

for all $x_1,x_2\in\mathbb{R}^n$, for all $\tau\in\mathbb{R}_{\geq 0}$, and for all $\varepsilon\in(0,\varepsilon_0)$.

Proof of item (d): Since Ψ is a diffeomorphism, we have that

$$x=\pi_1\circ\Psi\circ\Psi^{-1}(x,\tau)=\pi_1\circ\Psi^{-1}(x,\tau)-\sum_{k=1}^{2}\varepsilon^k v_k\circ\Psi^{-1}(x,\tau).$$

Therefore, for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$:

$$\left|\pi_1\circ\Psi^{-1}(x,\tau)-x\right|=\left|\sum_{k=1}^{2}\varepsilon^k v_k\circ\Psi^{-1}(x,\tau)\right|. \tag{35}$$

From item (b) in Lemma 3 and the inequality (34), we know that, for all $x_1,x_2\in\mathbb{R}^n$ and all $\tau\in\mathbb{R}_{\geq 0}$, the maps v_k satisfy:

$$\left|v_k\circ\Psi^{-1}(x_1,\tau)-v_k\circ\Psi^{-1}(x_2,\tau)\right|\leq 2L_{v,k}|x_1-x_2|.$$

Moreover, since $\varepsilon_0\leq\delta_3/L_\Psi$, it follows from inequality (33) that $\pi_1\circ\Psi^{-1}(0,\tau)\in 3\delta_3\mathbb{B}$, which implies that, for all $\tau\in\mathbb{R}_{\geq 0}$, we have that

$$\left|v_1\circ\Psi^{-1}(0,\tau)\right|\leq TB_{1,\delta},$$
$$\left|v_2\circ\Psi^{-1}(0,\tau)\right|\leq T\left(2B_{2,\delta}+3T\hat{L}_1B_{1,\delta}\right).$$

Therefore, by adding and subtracting terms to (35) and using the triangle inequality and the previous bounds, we obtain that

$$\left|\pi_1\circ\Psi^{-1}(x,\tau)-x\right|\leq\sum_{k=1}^{2}\varepsilon^k\hat{L}_{v,k}(|x|+1),$$

for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$ and all $\varepsilon\in(0,\varepsilon_0)$. Invoking the reverse triangle inequality, we get

$$|x|-\varepsilon(|x|+1)\sum_{k=1}^{2}\hat{L}_{v,k}\leq\left|\pi_1\circ\Psi^{-1}(x,\tau)\right|,$$

for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$. Next, for all $\varepsilon\in(0,\varepsilon_0)$, we have that $0\leq 1-\varepsilon\sum_{k=1}^{2}\hat{L}_{v,k}$, for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$. Hence, we obtain that

$$\left(1-\varepsilon\sum_{k=1}^{2}\hat{L}_{v,k}\right)|x|-\varepsilon\sum_{k=1}^{2}\hat{L}_{v,k}\leq\left|\pi_1\circ\Psi^{-1}(x,\tau)\right|,$$

for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$, which implies that

$$\delta_2\leq\left(1-\varepsilon\sum_{k=1}^{2}\hat{L}_{v,k}\right)\delta_3-\varepsilon\sum_{k=1}^{2}\hat{L}_{v,k}\leq\left|\pi_1\circ\Psi^{-1}(x,\tau)\right|,$$

for all $(x,\tau,\varepsilon)\in\mathcal{M}_3\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$. ■

Lemma 5. *There exists a C^0 map $g:\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)\to\mathbb{R}^n$ such that, for all $(x,\tau)\in\mathcal{M}_3\times\mathbb{R}_{\geq 0}$ and all $\varepsilon\in(0,\varepsilon_0)$, the map Ψ_*f_ε, given by (6b), satisfies*

$$\Psi_*f_\varepsilon(x,\tau)=\bar{f}(x)+\varepsilon\, g(x,\tau,\varepsilon),$$

where $\bar{f}$ is given by (9f).

Proof. By direct computation, we obtain:

$$\begin{aligned}\Psi_*f_\varepsilon(x,\tau)&=\left(\mathrm{D}_x\Phi\circ\Psi^{-1}(x,\tau)\right)f_2\circ\Psi^{-1}(x,\tau)\\&\quad+\left(\mathrm{D}_x\Phi\circ\Psi^{-1}(x,\tau)\right)f_1\circ\Psi^{-1}(x,\tau)\,\varepsilon^{-1}\\&\quad+\mathrm{D}_\tau\Phi\circ\Psi^{-1}(x,\tau)\,\varepsilon^{-2},\end{aligned}$$

$$\mathrm{D}_\tau\Phi\circ\Psi^{-1}(x,\tau)=-\sum_{k=1}^{2}\varepsilon^k\mathrm{D}_\tau v_k\circ\Psi^{-1}(x,\tau)$$

$$\mathrm{D}_x\Phi\circ\Psi^{-1}(x,\tau)=\left(I-\sum_{k=1}^{2}\varepsilon^k\mathrm{D}_x v_k\right)\circ\Psi^{-1}(x,\tau).$$

Moreover, note that

$$\mathrm{D}_\tau v_1\circ\Psi^{-1}(x,\tau)=\hat{f}_1\circ\Psi^{-1}(x,\tau),$$

and also that

$$\begin{aligned}\mathrm{D}_\tau v_2(x,\tau)&=\underbrace{\hat{f}_2(x,\tau)+\mathrm{D}_x\hat{f}_1(x,\tau)v_1(x,\tau)-\bar{f}(x)}_{=\mathrm{D}_\tau w(x,\tau)}\\&\quad\underbrace{-\mathrm{D}_x\hat{f}_1(x,\tau)v_1(x,\tau)-\mathrm{D}_x v_1(x,\tau)\hat{f}_1(x,\tau)}_{=-\mathrm{D}_\tau(\mathrm{D}_x v_1(x,\tau)v_1(x,\tau))}\\&=\hat{f}_2(x,\tau)-\mathrm{D}_x v_1(x,\tau)\hat{f}_1(x,\tau)-\bar{f}(x)\end{aligned}$$

which implies that

$$\begin{aligned}\mathrm{D}_\tau v_2\circ\Psi^{-1}(x,\tau)&=\hat{f}_2\circ\Psi^{-1}(x,\tau)-\bar{f}\circ\pi_1\circ\Psi^{-1}(x,\tau)\\&\quad-\left(\mathrm{D}_x v_1\circ\Psi^{-1}(x,\tau)\right)\hat{f}_1\circ\Psi^{-1}(x,\tau).\end{aligned}$$

Therefore, another direct computation shows that

$$\begin{aligned}\Psi_*f_\varepsilon(x,\tau)&=\left(f_1\circ\Psi^{-1}(x,\tau)-\hat{f}_1\circ\Psi^{-1}(x,\tau)\right)\varepsilon^{-1}\\&\quad+\left(f_2\circ\Psi^{-1}(x,\tau)-\hat{f}_2\circ\Psi^{-1}(x,\tau)\right)\\&\quad-\mathrm{D}_x v_1\circ\Psi^{-1}(x,\tau)f_1\circ\Psi^{-1}(x,\tau)\\&\quad+\mathrm{D}_x v_1\circ\Psi^{-1}(x,\tau)\hat{f}_1\circ\Psi^{-1}(x,\tau)+\bar{f}(x)+\hat{f}(x,\tau,\varepsilon),\end{aligned}\tag{36}$$

where the map $\hat{f}$ is given by

$$\begin{aligned}\hat{f}(x,\tau,\varepsilon)&=\bar{f}\circ\pi_1\circ\Psi^{-1}(x,\tau)-\bar{f}(x)-\varepsilon\mathrm{D}_x v_1\circ\Psi^{-1}(x,\tau)\,f_2\circ\Psi^{-1}(x,\tau)\\&\quad-\varepsilon\mathrm{D}_x v_2\circ\Psi^{-1}(x,\tau)\,f_1\circ\Psi^{-1}(x,\tau)\\&\quad-\varepsilon^2\mathrm{D}_x v_2\circ\Psi^{-1}(x,\tau)\,f_2\circ\Psi^{-1}(x,\tau).\end{aligned}$$

Using Hadamard's Lemma [45, Lemma 2.8] and the fact that $\bar{f}$ is C^1, we obtain:

$$\bar{f}(x_1)-\bar{f}(x_2)=\bar{F}(x_1,x_2)(x_1-x_2),$$

for all $x_1,x_2\in\mathbb{R}^n$, where $\bar{F}$ is given by

$$\bar{F}(x_1,x_2):=\int_0^1\mathrm{D}_x\bar{f}(\lambda x_1+(1-\lambda)x_2)\,d\lambda.$$

Hence, using the fact that Ψ^{-1} is a bijection:

$$\bar{f}\circ\pi_1\circ\Psi^{-1}(x,\tau)-\bar{f}(x)=\tilde{F}(x,\tau)\left(\pi_1\circ\Psi^{-1}(x,\tau)-x\right),$$

for all $x\in\mathbb{R}^n$, where $\tilde{F}$ is given by

$$\tilde{F}(x,\tau):=\bar{F}\left(\pi_1\circ\Psi^{-1}(x,\tau),x\right).\tag{37}$$

However, since Ψ is a diffeomorphism, we have

$$x=\pi_1\circ\Psi\circ\Psi^{-1}(x,\tau)=\pi_1\circ\Psi^{-1}(x,\tau)-\sum_{k=1}^{2}\varepsilon^k v_k\circ\Psi^{-1}(x,\tau),$$

which implies that

$$\begin{aligned}\bar{f}\circ\pi_1\circ\Psi^{-1}(x,\tau)-\bar{f}(x)&=\varepsilon\ \tilde{F}(x,\tau)\,v_1\circ\Psi^{-1}(x,\tau)\\&\quad+\varepsilon^2\tilde{F}(x,\tau)\,v_2\circ\Psi^{-1}(x,\tau),\end{aligned}$$

and that $\hat{f}$ can be written as

$$\hat{f}(x,\tau,\varepsilon)=\varepsilon\, g(x,\tau,\varepsilon),$$

where g can be written in compact form as:

$$\begin{aligned}g&=\tilde{F}\,v_1\circ\Psi^{-1}-\mathrm{D}_x v_1\circ\Psi^{-1}\,f_2\circ\Psi^{-1}-\mathrm{D}_x v_2\circ\Psi^{-1}\,f_1\circ\Psi^{-1}\\&\quad+\varepsilon\left(\tilde{F}\,v_2\circ\Psi^{-1}-\mathrm{D}_x v_2\circ\Psi^{-1}\,f_2\circ\Psi^{-1}\right).\end{aligned}\tag{38}$$

Since Ψ is a diffeomorphism, and g is a combination of C^0 maps composed with Ψ, it follows that g is C^0 in all arguments.

Finally, by item (d) in Lemma 4, for all $(x,\tau,\varepsilon)\in\mathcal{M}_3\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$, we have that $\Psi^{-1}(x,\tau)\in\mathcal{M}_2\times\mathbb{R}_{\geq 0}$. Also, by item (a) in Lemma 2, for all $(x,\tau)\in\mathcal{M}_2\times\mathbb{R}_{\geq 0}$, we have that $\hat{f}_k(x,\tau)=f_k(x,\tau)$. Therefore,

$$\hat{f}_k\circ\Psi^{-1}(x,\tau)=f_k\circ\Psi^{-1}(x,\tau),$$

for all $(x,\tau,\varepsilon)\in\mathcal{M}_3\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$. Hence, in this set the first four terms in (36) cancel, and we obtain that the pushforward map Ψ_*f_ε satisfies

$$\Psi_*f_\varepsilon(x,\tau)=\bar{f}(x)+\varepsilon\, g(x,\tau,\varepsilon).$$

for all $(x,\tau,\varepsilon)\in\mathcal{M}_3\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$. ■

Lemma 6. *There exists a positive constant $L_g>0$ such that the map g, defined in (38), satisfies*

$$|g(x,\tau,\varepsilon)|\leq L_g(|x|+1),\tag{39}$$

for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$.

Proof. The map g can be written in compact form as

$$g(x,\tau,\varepsilon)=\sum_{k=1}^{5}G_i(x,\tau,\varepsilon)\,g_i(x,\tau,\varepsilon),\tag{40}$$

with the matrix-valued maps G_i given by

$$\begin{aligned}&G_1(x,\tau)=\tilde{F}(x,\tau), && G_2(x,\tau)=\mathrm{D}_x v_1\circ\Psi^{-1}(x,\tau),\\&G_3(x,\tau)=\mathrm{D}_x v_2\circ\Psi^{-1}(x,\tau), && G_4(x,\tau)=\varepsilon\tilde{F}(x,\tau),\\&G_5(x,\tau)=\varepsilon\mathrm{D}_x v_2\circ\Psi^{-1}(x,\tau),\end{aligned}$$

and the maps g_i given by

$$\begin{aligned}&g_1(x,\tau)=v_1\circ\Psi^{-1}(x,\tau), && g_2(x,\tau)=-f_2\circ\Psi^{-1}(x,\tau),\\&g_3(x,\tau)=-f_1\circ\Psi^{-1}(x,\tau), && g_4(x,\tau)=v_2\circ\Psi^{-1}(x,\tau),\\&g_5(x,\tau)=-f_2\circ\Psi^{-1}(x,\tau).\end{aligned}$$

where the explicit (smooth) dependence on ε is omitted to simplify notation. By Lemma 3, the maps $\bar{f}$ and v_k are C^1 and satisfy

$$|\mathrm{D}_x\bar{f}(x)|\leq\bar{L},\qquad|\mathrm{D}_x v_k(x,\tau)|\leq L_{v,k},$$

for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times(0,\varepsilon_0)$. Thus, for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times[0,\varepsilon_0]$, we have that $|G_i(x,\tau,\varepsilon)|\leq M_{g,i}$ for all i, where the constants $M_{g,i}$ are given by

$$M_{g,1}:=2\bar{L},\qquad M_{g,2}:=2L_{v,1},\qquad M_{g,3}:=2L_{v,2},$$

$$M_{g,4} := 2\bar{L}, \qquad M_{g,5} := 2L_{v,2},$$

By Lemma 4, the diffeomorphism Ψ and its inverse Ψ^{-1} are globally Lipschitz in x. In addition, from items (a) in Assumption 1 and Lemma 3, the maps f_k are Lipschitz in x for all $(x,\tau) \in \mathcal{M}_1 \times \mathbb{R}_{\geq 0}$ and are C^0 for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$. Furthermore, $\Psi^{-1}(x,\tau) \in \mathcal{M}_2 \times \mathbb{R}_{\geq 0} \subset \mathcal{M}_1 \times \mathbb{R}_{\geq 0}$, for all $(x,\tau) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0}$. Therefore, for all $x_1, x_2 \in \mathcal{M}_3$, all $(\tau,\varepsilon) \in \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, and all $i \in \{1,2,\ldots,5\}$:

$$|g_i(x_1,\tau,\varepsilon) - g_i(x_2,\tau,\varepsilon)| \leq L_{g,i}|x_1 - x_2|,$$

where the constants $L_{g,i}$ are

$$L_{g,1} := 2L_{v,1}, \qquad L_{g,2} := 2L_2, \qquad L_{g,3} := 2L_1,$$
$$L_{g,4} := 2L_{v,2}, \qquad L_{g,5} := 2L_2.$$

Substituting with $x_1 = x \in \delta_3\mathbb{B}$ and $x_2 = \pi_1 \circ \Psi(0,\tau)$ in item (c) of Lemma 4, we have that

$$\left|\pi_1 \circ \Psi^{-1}(x,\tau)\right| \leq (1 + L_\Psi \varepsilon)|x| + L_\Psi \varepsilon < 2|x| + \delta_3 < 3\delta_3,$$

for all $x \in \delta_3\mathbb{B}$. It follows that $\Psi^{-1}(x,\tau) \in 3\delta_3\mathbb{B} \times \mathbb{R}_{\geq 0}$, for all $(x,\tau) \in \delta_3\mathbb{B} \times \mathbb{R}_{\geq 0}$. Consequently, we have that

$$\sup_{(x,\tau)\in\delta_3\mathbb{B}\times\mathbb{R}_{\geq 0}} |f_k \circ \Psi^{-1}(x,\tau)| \leq \sup_{(x,\tau)\in 3\delta_3\mathbb{B}\times\mathbb{R}_{\geq 0}} |f_k(x,\tau)| = B_{k,\delta}.$$

Recall that

$$g_1(x,\tau,\varepsilon) = v_1 \circ \Psi^{-1}(x,\tau),$$

Hence, we have that

$$\sup_{(x,\tau)\in\delta_3\mathbb{B}\times\mathbb{R}_{\geq 0}} |g_1(x,\tau)| \leq \sup_{(x,\tau)\in 3\delta_3\mathbb{B}\times\mathbb{R}_{\geq 0}} |v_1(x,\tau)| \leq TB_{1,\delta} =: M_{0,1}.$$

Similarly, it can be shown that, there exist constants $M_{0,i} \in \mathbb{R}_{>0}$ such that $|g_i(x,\tau,\varepsilon)| \leq M_{0,i}$, for all $(x,\tau,\varepsilon) \in \delta_3\mathbb{B} \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$. In particular, the constants $M_{0,i}$ are given by

$$M_{0,1} := TB_{1,\delta}, \qquad M_{0,2} := B_{2,\delta}, \qquad M_{0,3} := B_{1,\delta},$$
$$M_{0,4} := T\left(2B_{2,\delta} + 3T\hat{L}_1 B_{1,\delta}\right), \qquad M_{0,5} := B_{2,\delta}.$$

Let $x_m \in \{x : \mathbb{R}^n : |x| = \delta_3\}$ be an arbitrary point, and note that $x_m \in \mathcal{M}_3 \cap \delta_3\mathbb{B}$, and that, for all $(x,\tau,\varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, we have that each term in (40) can be written as:

$$\begin{aligned} G_i(x,\tau,\varepsilon)g_i(x,\tau,\varepsilon) &= G_i(x,\tau,\varepsilon)(g_i(x,\tau,\varepsilon) - g_i(x_m,\tau,\varepsilon)) \\ &\quad + G_i(x,\tau,\varepsilon)g_i(x_m,\tau,\varepsilon). \end{aligned}$$

For all $x \in \mathbb{R}^n$, either $x \in \mathcal{M}_3$ or $x \in (\mathbb{R}^n \backslash \mathcal{M}_3) \subset \delta_3\mathbb{B}$. If $x \in \mathcal{M}_3$, then we have that

$$\begin{aligned} \left|G_i(x,\tau,\varepsilon)g_i(x,\tau,\varepsilon)\right| &\leq M_{g,i}L_{g,i}|x - x_m| + M_{g,i}M_{g,0} \\ &\leq M_{g,i}L_{g,i}|x| + M_{g,i}(L_{g,i}\delta_3 + M_{g,0}). \end{aligned}$$

Alternatively, if $x \in (\mathbb{R}^n \backslash \mathcal{M}_3)$, then we have that

$$\left|G_i(x,\tau,\varepsilon)g_i(x,\tau,\varepsilon)\right| \leq M_{g,i}M_{g,0}.$$

Combining all of the above, we obtain that, for all $(x,\tau,\varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, the map g satisfies the inequality

$$|g(x,\tau,\varepsilon)| \leq L_g(|x| + 1),$$

where $L_g := \max\left\{\sum_{i=1}^5 M_{g,i}L_{g,i}, \sum_{i=1}^5 M_{g,i}(L_{g,i}\delta_3 + M_{g,0})\right\}$. ■

All the claims of Proposition 1 follow now directly by Lemmas 2–6. ■

5.2. Proof of Theorem 1

First, we introduce several definitions. Let $c_1, c_2, \alpha_1, \alpha_2, \alpha_3, \bar{L}_g$, and ϕ be generated by Assumption 2, and let ε_0 and L_g be the constants generated by Proposition 1. Let $\alpha_4 \in \mathcal{K}$ be such that $\alpha_4(|x|) \leq \phi(x)$ for all $x \in \mathbb{R}^n$. Such a function exists because ϕ is positive definite [5, Lemma 4.3]. Let

$$a_3 := \sup_{r\in\mathbb{R}_{\geq 0}} \alpha_3(r), \qquad a_4 := \sup_{r\in\mathbb{R}_{\geq 0}} \alpha_4(r), \tag{41}$$
$$\eta_1^* := c_1/(4c_2\bar{L}_g), \qquad \varepsilon_1^* := \min\{\varepsilon_0, \eta_1^*, a_4\eta_1^*/2\}, \tag{42}$$
$$\eta_2^* := c_1/(4c_2L_g), \qquad \varepsilon_2^* := \min\{\varepsilon_0, a_3\eta_2^*/2\}, \tag{43}$$

and let $\alpha_5(r) := \alpha_3(r)r \in \mathcal{K}_\infty$. Note that, in general, $a_3, a_4 \in \mathbb{R}_{>0} \cup \{\infty\}$. By Assumption 2, at least one of the following cases holds:

(C1) Item (c)–(i) in Assumption 2 holds; or
(C2) Item (c)–(ii) in Assumption 2 holds.

Therefore, the constants:

$$(\varepsilon^*, \eta^*) := \begin{cases} (\varepsilon_1^*, \eta_1^*) & \text{(C1) is true \& (C2) is false} \\ (\varepsilon_2^*, \eta_2^*) & \text{(C2) is true,} \end{cases} \tag{44}$$

and the function:

$$\rho(\delta,\varepsilon) := \begin{cases} \alpha_4^{-1}\left(\frac{\varepsilon}{\eta_1^*}\right) & \text{(C1) is true \& (C2) is false} \\ \alpha_3^{-1}\left(\frac{\varepsilon}{\eta_2^*}\right) + \alpha_5^{-1}\left(\frac{\varepsilon}{\eta_2^*}\right) & \text{(C2) is true,} \end{cases} \tag{45}$$

are always well-defined whenever Assumption 2 holds, where ρ depends on δ through L_g or $\bar{L}_g$, appearing in the definition of the constants η_1^* and η_2^*, which, in general, are fixed only *after* a choice of δ is fixed. Also, for any fixed choice of δ satisfying (7), the function $\rho(\delta,\cdot)$ is a class $\mathcal{K}$ function [5, Lemma 4.2].

Next, note that from Assumption 1 and Proposition 1, and for all $\varepsilon \in (0,\varepsilon_0)$, the map $\Psi_* f_\varepsilon$ is continuous. Hence, for all $(x_0,\tau_0,\varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, a solution to system (6a) starting at the initial condition (x_0,τ_0) exists.

Let V be given by Assumption 2. Its time derivative along the trajectories of (6a) satisfies:

$$\dot{V} = \nabla V(x)^\top \Psi_* f_\varepsilon(x,\tau,\varepsilon).$$

Using item (d) in Proposition 1 and the bounds from Assumption 2, we obtain that for all $(x,\tau,\varepsilon) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$:

$$\begin{aligned} \dot{V} &\leq \nabla V(x)^\top \bar{f}(x) + \varepsilon \nabla V(x)^\top g(x,\tau,\varepsilon) \\ &\leq -c_1\phi(x)^2 + \varepsilon c_2\phi(x)|g(x,\tau,\varepsilon)|. \end{aligned}$$

We consider two possible cases:

(C1) Item (c)–(i) in Assumption 2 holds, then:

$$\begin{aligned} \dot{V} \leq &-\frac{c_1}{2}\phi(x)^2 - \left(\frac{c_1}{4} - \varepsilon c_2\bar{L}_g\right)\phi(x)^2 \\ &- \phi(x)\left(\frac{c_1}{4}\phi(x) - \varepsilon c_2\bar{L}_g\right), \end{aligned}$$

for all $(x,\tau,\varepsilon) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, which implies that:

$$\dot{V} \leq -\frac{c_1}{2}\phi(x)^2, \qquad \forall |x| \geq \delta_3 + \alpha_4^{-1}\left(\frac{\varepsilon}{\eta_1^*}\right).$$

(C2) Item (c)–(ii) in Assumption 2 holds. Then:

$$\begin{aligned} \dot{V} \leq &-\frac{c_1}{2}\phi(x)^2 - \phi(x)\left(\frac{c_1}{4}\alpha_3(|x|) - \varepsilon c_2L_g\right)|x| \\ &- \phi(x)\left(\frac{c_1}{4}\alpha_3(|x|)|x| - \varepsilon c_2L_g\right), \end{aligned}$$

for all $(x,\tau,\varepsilon) \in \mathcal{M}_3 \times \mathbb{R}_{\geq 0} \times (0,\varepsilon_0)$, where $\alpha_3 \in \mathcal{K}$, and where we used Proposition 1-(e). It follows that:

$$\dot{V} \leq -\frac{c_1}{2}\phi(x)^2, \quad \forall |x| \geq \delta_3 + \alpha_3^{-1}\left(\frac{\varepsilon}{\eta_2^*}\right) + \alpha_5^{-1}\left(\frac{\varepsilon}{\eta_2^*}\right).$$

Combining all of the above, we obtain that $\dot{V} \leq -c\phi(x)^2$, for all $|x| \geq 2\delta_3 + \rho(\delta,\varepsilon)$ and all $\varepsilon \in (0,\varepsilon^*)$, with $c := c_1/2$. Then, following similar steps as in [5, Appendix C.9] and the proof of [46, Appendix C.], there exist functions $\beta \in \mathcal{KL}$ and $\kappa \in \mathcal{K}_\infty$ such that, for all $(x_0, \tau_0, \varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times (0, \varepsilon^*)$, any solution to system (6a) starting at (x_0, τ_0), satisfies

$$|x(t)| \leq \beta(|x(t)|, t) + \Delta_{\delta,\varepsilon}, \quad \forall t \geq 0, \tag{46}$$

where $\Delta_{\delta,\varepsilon} := \kappa(2\delta_3 + \rho(\delta, \varepsilon))$. ■

5.3. Proof of Corollary 1

By Theorem 1, there exists an $\varepsilon^* \in (0, \varepsilon_0)$ and $\beta \in \mathcal{KL}$ such that, for all $(\bar{x}_0, \tau_0, \varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times (0, \varepsilon^*)$, any solution $(\bar{x}, \tau)$ of system (6a) starting from $(\bar{x}_0, \tau_0)$ satisfies the $\mathcal{KL}$ bound (46). Let (x, τ) be a solution of system (4) starting from an initial condition (x_0, τ_0). Since Ψ is a diffeomorphism, and system (6a) is the pushforward of system (4) under Ψ, it follows that $(x(t), \tau(t)) = \Psi^{-1}(\bar{x}(t), \tau(t))$, for all $t > 0$, where $(\bar{x}, \tau)$ is some solution of system (6a) with initial condition $\Psi(x_0, \tau_0)$. Therefore, from the triangle inequality, we have that for all $t \geq 0$:

$$|x(t)| \leq \left|\pi_1 \circ \Psi^{-1}(\bar{x}(t), \tau(t)) - \pi_1 \circ \Psi^{-1}(0, \tau(t))\right| + \left|\pi_1 \circ \Psi^{-1}(0, \tau(t))\right|.$$

From item (b) in Proposition 1, we obtain that

$$\left|\pi_1 \circ \Psi^{-1}(\bar{x}(t), \tau(t)) - \pi_1 \circ \Psi^{-1}(0, \tau(t))\right| \leq (1 + L_\Psi \varepsilon)|\bar{x}(t)|,$$

and also $|\pi_1 \circ \Psi^{-1}(0, \tau(t))| \leq L_\Psi \varepsilon$. Therefore, it follows that

$$|x(t)| \leq (1 + L_\Psi \varepsilon)\beta(|\pi_1 \circ \Psi(x_0, \tau_0)|, t) + (1 + L_\Psi \varepsilon)\Delta_{\delta,\varepsilon} + L_\Psi \varepsilon.$$

Similarly, we have

$$\begin{aligned}|\pi_1 \circ \Psi(x_0, \tau_0)| &\leq |\pi_1 \circ \Psi(x_0, \tau_0) - \pi_1 \circ \Psi(0, \tau_0)| + |\pi_1 \circ \Psi(0, \tau_0)| \\ &\leq (1 + L_\Psi \varepsilon)|x_0| + L_\Psi \varepsilon.\end{aligned}$$

Since $\beta(\cdot, t) \in \mathcal{K}_\infty$, it is strictly increasing and satisfies

$$\beta(|\pi_1 \circ \Psi(x_0, \tau_0)|, t) \leq \beta((1 + L_\Psi \varepsilon)|x_0| + L_\Psi \varepsilon, t).$$

for all $(x_0, \tau_0, t, \varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times [0, \varepsilon^*]$. We then have two possible cases:

(C1) If $|x_0| \leq L_\Psi \varepsilon$, then

$$\beta(|\pi_1 \circ \Psi(x_0, \tau_0)|, t) \leq \beta(2L_\Psi \varepsilon + L_\Psi^2 \varepsilon^2, t).$$

(C2) If $|x_0| > L_\Psi \varepsilon$, then

$$\beta(|\pi_1 \circ \Psi(x_0, \tau_0)|, t) \leq \beta((2 + L_\Psi \varepsilon)|x_0|, t).$$

Therefore, for all $(x_0, \tau_0, t, \varepsilon) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times [0, \varepsilon^*]$:

$$\begin{aligned}\beta(|\pi_1 \circ \Psi(x_0, \tau_0)|, t) &\leq \beta((2 + L_\Psi \varepsilon)|x_0|, t) + \beta(2L_\Psi \varepsilon + L_\Psi^2 \varepsilon^2, t) \\ &\leq \beta((2 + L_\Psi \varepsilon)|x_0|, t) + \beta(2L_\Psi \varepsilon + L_\Psi^2 \varepsilon^2, 0).\end{aligned}$$

However, from Claim 1, we have that $L_\Psi \varepsilon \leq 1$, for all $\varepsilon \in [0, \varepsilon^*]$. Therefore, we have that

$$\beta(|\pi_1 \circ \Psi(x_0, \tau_0)|, t) \leq \beta(3|x_0|, t) + \beta(3L_\Psi \varepsilon, 0).$$

The result of the corollary follows by defining

$$\tilde{\Delta}(\delta, \varepsilon) := (1 + L_\Psi \varepsilon)\Delta_{\delta,\varepsilon} + L_\Psi \varepsilon + (1 + L_\Psi \varepsilon)\beta(3L_\Psi \varepsilon, 0) \tag{47}$$

and

$$\tilde{\beta}(r, s) := 2\beta(3r, s), \tag{48}$$

for all $r, s \geq 0$. ■

5.4. Proof of Corollary 2

It follows from Assumption 2 that $\kappa \in \mathcal{K}$ and $\beta \in \mathcal{KL}$ generated by the proof of Theorem 1 are independent of the choice of $\delta_1 > 0$ and δ satisfying (7). Let $\nu > 0$ be given. Hence, there exists $r > 0$ sufficiently small such that $\kappa(r) < \frac{\nu}{4}$. Let $\delta_1 < r/16$, and choose, $\delta_2 = \sqrt{3}\delta_1$, $\delta_3 = \frac{4}{3}\sqrt{3}\delta_2 = 4\delta_1$, which satisfy (7). It follows that $2\delta_3 = 8\delta_1 < \frac{r}{2}$. Let $\varepsilon_a > 0$ be such that $\rho(\delta, \varepsilon) < \frac{r}{2}$ for all $\varepsilon \in (0, \varepsilon_a)$. Such ε_a always exist because $\rho(\delta, \cdot) \in \mathcal{K}$. It follows that $\Delta_{\delta,\varepsilon} = \kappa(2\delta_3 + \rho(\delta, \varepsilon)) < \frac{\nu}{4}$. Let $\varepsilon_b > 0$ be such that $L_\Psi \varepsilon \leq \min\{\nu/3, 1/3\}$ and $\beta(3L_\Psi \varepsilon, 0) \leq \nu/4$ for all $\varepsilon \in (0, \varepsilon_b)$, where β comes from (46). Such ε_b always exist because $\beta(\cdot, s) \in \mathcal{K}$. Let $\varepsilon^* > 0$ be generated by Corollary 1, and define $\varepsilon^{**} = \min\{\varepsilon_a, \varepsilon_b, \varepsilon^*\}$. Then, every solution of system (4) starting at (x_0, τ_0) satisfies the bound (3) with $\mathcal{KL}$ function given by (48) and ultimate bound $\tilde{\Delta}_{\delta,\varepsilon}$ given by (47). However, by the choice of δ and ε, we have that $\tilde{\Delta}_{\delta,\varepsilon} \leq \nu$, which establishes the desired bound. ■

5.5. Proof of Corollary 3

Since Assumptions 1 and 2 are satisfied for all $\delta_3 > \delta_2 = \delta_1 = 0$, we may pick $\delta_3 \in (0, \infty)$ arbitrarily small. Following similar steps to the proof of Corollary 2 yields the desired result. ■

5.6. Proof of Theorem 2

We consider the case when $J^\star \in \mathbb{R}$ is arbitrary, and we verify that the maps defining system (26) satisfy Assumption 1. Clearly, the right hand side in (26) is C^0 and satisfies item (b) in Assumption 1. Let $\bar{J} \in \mathbb{R}_{>0}$, and let $\delta_1 \in [0, \infty)$ be such that $J(x) \geq \bar{J}$, for all $|x| \geq \delta_1$. Such δ_1 always exists because J is radially unbounded. It follows that the feedback law

$$\begin{aligned}u_{i,1}(J(x), \tau) &= \sqrt{2\omega_i J(x)}\cos(\log(J(x)) + \omega_i \tau) \\ u_{i,2}(J(x), \tau) &= \sqrt{2\omega_i J(x)}\sin(\log(J(x)) + \omega_i \tau),\end{aligned}$$

is C^2 for all $|x| \geq \delta_1$. We recall that system (26) has the form of system (4) with

$$f_1(x, \tau) = \sum_{i=1}^{r}\sum_{j=1}^{2} b_{i,j}u_{i,j}(J(x), \tau), \qquad f_2(x, \tau) = b_0(x). \tag{49}$$

Since the vectors $b_{i,j}$ are constant and the functions $u_{i,j}(J(x), \tau)$ are C^2 for all $|x| \geq \delta_1$, it follows that system (26) satisfies item (c) in Assumption 1. Next, direct computation gives

$$|\mathrm{D}_x f_1(x, \tau)| \leq |\nabla J(x)| \sum_{i=1}^{r}\sum_{j=1}^{2} |\mathrm{D}_y u_{i,j}(J(x), \tau)||b_{i,j}|.$$

$$|\mathrm{D}_x f_1(x, \tau_1) f_1(x, \tau_2)| \leq \sum_{i,k=1}^{r}\sum_{j,l=1}^{2} |b_{i,j}||b_{k,l}||U_{i,j,k,l}(x, \tau_1, \tau_2)|,$$

$$|\mathrm{D}_x(\mathrm{D}_x f_1(x, \tau_1) f_1(x, \tau_2))| \leq \sum_{i,k=1}^{r}\sum_{j,l=1}^{2} |b_{i,j}||b_{k,l}||\mathrm{D}_x U_{i,j,k,l}(x, \tau_1, \tau_2)|,$$

where the maps $U_{i,j,k,l}$ and $\mathrm{D}_x U_{i,j,k,l}$ are given by

$$\begin{aligned}U_{i,j,k,l}(x, \tau_1, \tau_2) &= \mathrm{D}_y u_{i,j}(J(x), \tau_1) u_{k,l}(J(x), \tau_2)\nabla J(x), \\ \mathrm{D}_x U_{i,j,k,l}(x, \tau_1, \tau_2) &= \mathrm{D}_y u_{i,j}(J(x), \tau_1) u_{k,l}(J(x), \tau_2)\nabla^2 J(x) \\ &\quad + \mathrm{D}_y u_{i,j}(J(x), \tau_1)\mathrm{D}_y u_{k,l}(J(x), \tau_2)\nabla J(x)\nabla J(x)^\top \\ &\quad + \mathrm{D}_y^2 u_{i,j}(J(x), \tau_1) u_{k,l}(J(x), \tau_2)\nabla J(x)\nabla J(x)^\top,\end{aligned}$$

and the maps $\mathrm{D}_y u_{i,j}$ and $\mathrm{D}_y^2 u_{i,j}$ are given by

$$\mathrm{D}_y u_{i,1}(y, \tau) = \frac{\sqrt{\omega_i}}{\sqrt{2y}}\left(\cos\left(\tau\omega_i + \log(y)\right) - 2\sin\left(\tau\omega_i + \log(y)\right)\right),$$

$$\mathrm{D}_y u_{i,2}(y, \tau) = \frac{\sqrt{\omega_i}}{\sqrt{2y}}\left(2\cos\left(\tau\omega_i + \log(y)\right) + \sin\left(\tau\omega_i + \log(y)\right)\right),$$

$$\mathrm{D}_y^2 u_{i,1}(y,\tau) = -\frac{5\sqrt{\omega_i}}{2\sqrt{2}y^{\frac{3}{2}}}\cos\left(\tau\omega_i + \log(y)\right),$$
$$\mathrm{D}_y^2 u_{i,2}(y,\tau) = -\frac{5\sqrt{\omega_i}}{2\sqrt{2}y^{\frac{3}{2}}}\sin\left(\tau\omega_i + \log(y)\right),$$

for all $(y,\tau) \in \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0}$. We observe that, for all $|x| \geq \delta_1$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have that

$$|u_{i,j}(J(x),\tau)| \leq \sqrt{2\omega_i}J(x)^{\frac{1}{2}}, \qquad |\mathrm{D}_y u_{i,j}(J(x),\tau)| \leq \sqrt{\frac{5\omega_i}{2}}J(x)^{-\frac{1}{2}},$$
$$|\mathrm{D}_y^2 u_{i,j}(J(x),\tau)| \leq \frac{5\sqrt{\omega_i}}{2\sqrt{2}}J(x)^{-\frac{3}{2}}.$$

Therefore, we obtain that

$$|U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq \sqrt{5\omega_i\omega_k}J(x)^{-\frac{1}{2}}|\nabla J(x)|,$$
$$|\mathrm{D}_x U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq \sqrt{\omega_i\omega_k}(5J(x)^{-1}|\nabla J(x)|^2 + \sqrt{5}|\nabla^2 J(x)|).$$

From Assumption 4-(c), we have that ∇J is L_J-globally Lipschitz, which implies that $|\nabla J(x)|^2 \leq 2L_J(J(x)-J^\star)$, for all $x \in \mathbb{R}^n$ [47, Lemma 1, pp.23]. It follows that

$$|U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq \sqrt{10L_J\omega_i\omega_k}\left(1 + J^\star \bar{J}\right)^{\frac{1}{2}}. \tag{50}$$

Similarly, from Assumption 4-(c), we obtain that

$$|\mathrm{D}_x U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq 2L_J\sqrt{5\omega_i\omega_k}\left(\sqrt{5}\left(1 + J^\star \bar{J}\right) + 1\right).$$

for all $|x| \geq \delta_1$ and all $\tau \in \mathbb{R}_{\geq 0}$. Combining all of the above, we arrive at the upper bounds

$$|\mathrm{D}_x f_1(x,\tau)| \leq L_1, |\mathrm{D}_x(\mathrm{D}_x f_1(x,\tau_1)f_1(x,\tau_2))| \leq L_3,$$

for all $|x| \geq \delta_1$ and all $\tau \in \mathbb{R}_{\geq 0}$, where the constants L_1 and L_3 are given by

$$L_1 = \sqrt{L_J}\left(1 + J^\star \bar{J}\right)^{\frac{1}{2}}\sum_{i=1}^{r}\sum_{j=1}^{2}\sqrt{5\omega_i}|b_{i,j}|, \tag{51a}$$

$$L_3 = 2L_J\left(\sqrt{5}\left(1 + J^\star \bar{J}\right) + 1\right)\sum_{i,k=1}^{r}\sum_{j,l=1}^{2}\sqrt{5\omega_i\omega_k}|b_{i,j}||b_{k,l}|. \tag{51b}$$

Therefore, system (26) satisfies items (a) in Assumption 1 for $k = 1$ and (d) in Assumption 1 with the Lipschitz constants L_1 and L_3 given by (51). Finally, we note that, since $f_2(x,\tau) = b_0(x)$, it follows that system (26) satisfies items (a) for $k = 2$ with the Lipschitz constant $L_2 = L_0$ where L_0 is the Lipschitz constant from item (e) in Assumption 4. Hence, we have shown that system (26) satisfies all of the items in Assumption 1.

Next, let $\delta_2 \in (\delta_1, \infty)$, $\delta_3 \in (\delta_2, \infty)$, and let $\mathcal{M}_j$, for $j \in \{1,2,3\}$, be the corresponding nested subsets defined in (10). Using the formula (9f), the nominal average system (14) corresponding to system (26) on $\mathcal{M}_3$, is given by

$$\dot{x} = \bar{f}(x) = \varphi(x)(b_0(x) + \sum_{i=1}^{r}\sum_{j=1}^{2}b_{i,j}b_{i,j}^\top \nabla J(x)). \tag{52}$$

Consider the Lyapunov function candidate V and the positive definite function ϕ defined by

$$V(x) = J(x) - J^\star, \qquad \phi(x) = |\nabla J(x)|, \tag{53}$$

which satisfy the inequalities

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|), \qquad |\nabla V(x)| \leq \phi(x), \tag{54}$$

for all $x \in \mathbb{R}^n$, and also satisfy the inequality

$$\nabla V(x)^\top \bar{f}(x) \leq (\kappa_3 - \gamma)|\nabla J(x)|^2 < 0, \tag{55}$$

for all $x \in \mathcal{M}_3$, where the functions α_1 and α_2 are $\mathcal{K}_\infty$ functions, whose existence is guaranteed by the radial unboundedness of V [5, Lemma 4.3]. Since, by assumption $\gamma > \kappa_3$, system (26) satisfies items (a)–(b) in Assumption 2. Moreover, from Assumption 5, we have that $\alpha_J(|x|)|x| \leq |\nabla J(x)|$, where α_J is a class $\mathcal{K}$ function. Hence, system (26) satisfies item (c)–(ii) in Assumption 2 with $\phi(x) := |\nabla J(x)|$. By Theorem 1 we conclude that system (26) is UGUB.

Next, we consider the case when $J^\star \in \mathbb{R}_{>0}$. In this case, there exists $\bar{J} \in \mathbb{R}$ such that $0 < \bar{J} < J^\star$. Since $J^\star$ is the minimum value of the cost, it follows that $J(x) > \bar{J}$ for all $x \in \mathbb{R}^n$, which implies that the previous computations hold with $\delta_1 = 0$. In addition, in this case Assumption 1 holds with $\delta_1 = 0$, and the Lyapunov function candidate V and the positive definite function ϕ in (53) still satisfy the inequalities (54)–(55) for any choice of δ satisfying (7), with $\mathcal{M}_j$, for $j \in \{1,2,3\}$, being the corresponding nested subsets defined in (10). Therefore, by invoking Corollary 3, we conclude that system (26) is UGpAS. Finally, if $J^\star = 0$, we can take $\bar{J} = 0$ and in this case, Assumption 1 will be satisfied for all $\delta_1 \in (0,\infty)$ using $\tilde{m} = \sqrt{2\omega_i\kappa}$ in (50). Therefore, by Corollary 2, we conclude that the closed-loop system is UGpAS. ■

5.7. Proof of Theorem 3

It is easy to see that the right hand side in (26) is C^0 and satisfies item (b) in Assumption 1. In addition, the maps $u_{i,j}$ are C^∞, which implies that the maps $u_{i,j}(J(\cdot),\tau)$ are C^2 for all $x \in \mathbb{R}^n$. Therefore, system (26) satisfies item (c) in Assumption 1. Similar to the proof of Theorem 2, we compute that

$$|\mathrm{D}_x f_1(x,\tau)| \leq |\nabla J(x)|\sum_{i=1}^{r}\sum_{j=1}^{2}|\mathrm{D}_y u_{i,j}(J(x),\tau)||b_{i,j}|.$$
$$|\mathrm{D}_x f_1(x,\tau_1)f_1(x,\tau_2)| \leq \sum_{i,k=1}^{r}\sum_{j,l=1}^{2}|b_{i,j}||b_{k,l}||U_{i,j,k,l}(x,\tau_1,\tau_2)|,$$
$$|\mathrm{D}_x(\mathrm{D}_x f_1(x,\tau_1)f_1(x,\tau_2))| \leq \sum_{i,k=1}^{r}\sum_{j,l=1}^{2}|b_{i,j}||b_{k,l}||\mathrm{D}_x U_{i,j,k,l}(x,\tau_1,\tau_2)|,$$

where the maps $U_{i,j,k,l}$ and $\mathrm{D}_x U_{i,j,k,l}$ are given by

$$U_{i,j,k,l}(x,\tau_1,\tau_2) = \mathrm{D}_y u_{i,j}(J(x),\tau_1)u_{k,l}(J(x),\tau_2)\nabla J(x),$$
$$\begin{aligned}\mathrm{D}_x U_{i,j,k,l}(x,\tau_1,\tau_2) &= \mathrm{D}_y u_{i,j}(J(x),\tau_1)u_{k,l}(J(x),\tau_2)\nabla^2 J(x)\\ &+ \mathrm{D}_y u_{i,j}(J(x),\tau_1)\mathrm{D}_y u_{k,l}(J(x),\tau_2)\nabla J(x)\nabla J(x)^\top\\ &+ \mathrm{D}_y^2 u_{i,j}(J(x),\tau_1)u_{k,l}(J(x),\tau_2)\nabla J(x)\nabla J(x)^\top,\end{aligned}$$

and the maps $\mathrm{D}_y u_{i,j}$ and $\mathrm{D}_y^2 u_{i,j}$ are given by

$$\mathrm{D}_y u_{i,1}(y,\tau) = -\sqrt{2\omega_i}\sin\left(\omega_i\tau + y\right),$$
$$\mathrm{D}_y u_{i,2}(y,\tau) = \sqrt{2\omega_i}\cos\left(\omega_i\tau + y\right),$$
$$\mathrm{D}_y^2 u_{i,1}(y,\tau) = -\sqrt{2\omega_i}\cos\left(\omega_i\tau + y\right),$$
$$\mathrm{D}_y^2 u_{i,2}(y,\tau) = -\sqrt{2\omega_i}\sin\left(\omega_i\tau + y\right),$$

for all $(y,\tau) \in \mathbb{R}_{>0}\times\mathbb{R}_{\geq 0}$. We observe that, for all $x \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$, we have that

$$|u_{i,j}(J(x),\tau)| \leq \sqrt{2\omega_i}, \qquad |\mathrm{D}_y u_{i,j}(J(x),\tau)| \leq \sqrt{2\omega_i},$$
$$|\mathrm{D}_y^2 u_{i,j}(J(x),\tau)| \leq \sqrt{2\omega_i}.$$

Therefore, we obtain that

$$|U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq 2\sqrt{\omega_i\omega_k}|\nabla J(x)|,$$
$$|\mathrm{D}_x U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq 2\sqrt{\omega_i\omega_k}(2|\nabla J(x)|^2 + |\nabla^2 J(x)|).$$

From Assumption 4-(c) and Assumption 5-(b), we have that $|\nabla J(x)| \leq M_J$ and $|\nabla^2 J(x)| \leq L_J$ for all $x \in \mathbb{R}^n$. It follows that

$$|U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq 2M_J\sqrt{\omega_i\omega_k}$$
$$|\mathrm{D}_x U_{i,j,k,l}(x,\tau_1,\tau_2)| \leq 2(2M_J^2 + L_J)\sqrt{\omega_i\omega_k}.$$

for all $x \in \mathbb{R}^n$ and all $\tau \in \mathbb{R}_{\geq 0}$. Combining all of the above, we arrive at the upper bounds

$$|\mathrm{D}_x f_1(x,\tau)| \leq L_1, \qquad |\mathrm{D}_x(\mathrm{D}_x f_1(x,\tau_1)f_1(x,\tau_2))| \leq L_3,$$

for all $(x,\tau)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}$, where the constants L_1 and L_3 are given by

$$L_1=\sqrt{2}M_J\sum_{i=1}^{r}\sum_{j=1}^{2}\sqrt{\omega_i}|b_{i,j}|, \tag{56}$$

$$L_3=2(2M_J^2+L_J)\sum_{i,k=1}^{r}\sum_{j,l=1}^{2}\sqrt{\omega_i\omega_k}|b_{i,j}||b_{k,l}|. \tag{57}$$

Therefore, system (26) satisfies Assumption 1 with $\delta_1=0$. Next, let $\delta_1=\delta_2=0$, and fix an arbitrary choice of $\delta_3\in(0,\infty)$, and let $\mathcal{M}_j$ for $j\in\{1,2,3\}$ be the corresponding nested subsets as defined in (10). Using the formula (9f), the nominal average system (14) corresponding to system (26) on $\mathbb{R}^n$, is given by

$$\dot{x}=\bar{f}(x)=b_0(x)+\sum_{i=1}^{r}\sum_{j=1}^{2}b_{i,j}b_{i,j}^{\top}\nabla J(x). \tag{58}$$

Since the Lyapunov function candidate V and the function ϕ (53) now satisfy the inequalities (54)–(55) for all $x\in\mathcal{M}_3$, for any choice of $\delta_3\in(0,\infty)$, it follows that system (26) satisfies items (a)–(b) in Assumption 2. Moreover, from item (b) in Assumption 5, we have that $\alpha_J(|x|)\leq|\nabla J(x)|$, where α_J is a class $\mathcal{K}$ function. Finally, using item (b) in Assumption 5 and item (d) in Assumption 4, we obtain that $|b_0(x)|\leq|\nabla J(x)|\leq L_J$, for all $x\in\mathbb{R}^n$. Then, we have the following Claim.

Claim 2. *The map g from item (d) in Proposition 1 is uniformly bounded, for all $(x,\tau,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}_{\geq 0}\times[0,\varepsilon_0]$.*

The proof of Claim 2 can be found in Appendix A. From Claim 2, it follows that system (26) satisfies item (c)–(i) in Assumption 2. Therefore, by Corollary 3 we conclude that system (26) is UGpAS. ■

6. Conclusion and future work

We introduced a (second-order) averaging method that allows to study the stability properties of a class of oscillatory systems with periodic flows based on the stability properties of their corresponding averaged systems. In contrast to existing results in the literature, the method is suitable for the study of uniform *global* (practical) stability properties. Such properties are studied under suitable assumptions, which, naturally, are stronger compared to others that only enable local or semi-global practical results. By leveraging the proposed method, we showed that a class of extremum seeking algorithms is able to achieve uniform global practical asymptotic stability for a broad range of cost functions, which include quadratic (with positive definite Hessian), strongly convex, and certain invex functions. Future research will extend these results via singular perturbation theory to study dynamic plants in the loop, as well as systems with hybrid dynamics.

CRediT authorship contribution statement

Mahmoud Abdelgalil: Writing – original draft, Formal analysis, Conceptualization. **Jorge I. Poveda:** Writing – review & editing, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

The authors are grateful to the seven anonymous reviewers for their invaluable feedback, which has significantly improved the quality of the manuscript.

Appendix A. Auxiliary lemmas

Lemma 7. *Let $J:\mathbb{R}^n\to\mathbb{R}$ be a μ-strongly convex C^1 function with L globally Lipschitz gradient. Then, item (d) in Assumption 5 is satisfied.*

Proof. The upper bound follows directly by [42, Thm. 2.1.5]. To obtain the lower bound, note that by μ-strong convexity:

$$\left(\nabla J(x_1)-\nabla J(x_2)\right)^{\top}(x_1-x_2)\geq\mu|x_1-x_2|^2,$$

for all $x_1,x_2\in\mathbb{R}^n$. Using Cauchy-Schwartz inequality, it is easy to see that the following holds $\left|\nabla J(x_1)-\nabla J(x_2)\right|\geq\mu|x_1-x_2|$ for all $x_1,x_2\in\mathbb{R}^n$. It follows that

$$\left|\nabla J(x_1)-\nabla J(x_2)\right|^2\geq\mu^2|x_1-x_2|^2,$$

for all $x_1,x_2\in\mathbb{R}^n$. Let $\alpha_J:[0,\infty)\to[0,\mu)$ be given by $\alpha_J(s):=\mu\tanh(s)$, which is strictly increasing and satisfies $\alpha_J(0)=0$. Therefore, $\alpha_J\in\mathcal{K}$ and, by definition, $\alpha_J(s)<\mu$, for all $s\geq 0$. It follows that

$$\left|\nabla J(x_1)-\nabla J(x_2)\right|^2\geq\alpha_J(|x_1-x_2|)^2|x_1-x_2|^2.$$

for all $x_1,x_2\in\mathbb{R}^n$.

Lemma 8. *Let $J:\mathbb{R}^n\to\mathbb{R}$ be the function defined in Example 3. Then, J satisfies Assumption 3.*

Proof. The cost function J can be written as $J=h\circ H$ where $h(s)=s+3\sin(\sqrt{s})^2$ and $H(x):=|x|^2$ are C^∞ everywhere on their domain. Moreover, $H(x)\geq 0$ for all $x\in\mathbb{R}^n$. Therefore, the function $J=h\circ H$ is C^∞. The derivative of J satisfies

$$\nabla J(x)=\mathrm{D}h(H(x))\nabla H(x)=2\mathrm{D}h(H(x))(x),$$

where

$$\mathrm{D}h(H(x))=\frac{1}{2}\left(2+\frac{3\sin\left(2\sqrt{H(x)}\right)}{\sqrt{H(x)}}\right)\in\mathbb{R}.$$

It follows that

$$|\nabla J(x)|^2=4|\mathrm{D}h(H(x))(x)|^2=4Dh(H(x))^2|x|^2,$$

and it can be verified that $\frac{1}{4}<\mathrm{D}h(H(x))<4$, for all $x\in\mathbb{R}^n$. Therefore, there exists $\mu\in\mathbb{R}_{>0}$ such that, for all $x\in\mathbb{R}^n$, we have that

$$|\nabla J(x)|^2\geq 4\mu^2|x|^2\geq\alpha_J(|x|)^2|x|^2,$$

where $\alpha_J(s):=2\mu\tanh(s)$. Similarly, the second derivative of J satisfies

$$\nabla^2 J(x)=\mathrm{D}^2h(H(x))\nabla H(x)\nabla H(x)^{\top}+\mathrm{D}h(H(x))\nabla^2 H(x),$$

where

$$\mathrm{D}^2h(H(x))=\frac{3\cos\left(2\sqrt{H(x)}\right)}{2H(x)}-\frac{3\sin\left(2\sqrt{H(x)}\right)}{4H(x)^{3/2}},$$

It follows that

$$|\nabla^2 J(x)|\leq|\mathrm{D}^2h(H(x))||\nabla H(x)|^2+|\mathrm{D}h(H(x))||\nabla^2 H(x)|,$$

where $|\mathrm{D}^2h(H(x))|\leq\frac{3}{H(x)}$. Hence, the Hessian satisfies the inequality

$$|\nabla^2 J(x)|\leq\frac{3|\nabla H(x)|^2}{H(x)}+8\leq 20.$$

Appendix B. Proofs of auxiliary claims

B.1. Proof of Claim 1

Proof. The matrix $D_x\Phi(x,\tau)$ is a square matrix, and therefore its singular value decomposition is given by

$$D_x\Phi(x,\tau) = V(x,\tau)\Sigma(x,\tau)U(x,\tau)^\top,$$

where the matrices $V(x,\tau)$ and $U(x,\tau)$ are orthonormal matrices and $\Sigma(x,\tau)$ is a square diagonal matrix with the singular values of $D_x\Phi(x,\tau)$ on the diagonal. Since for all $\varepsilon \in (0,\varepsilon_0)$, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, the eigenvalues of the Jacobian matrix $D_x\Phi(x,\tau)$ are contained in the compact interval $[1-2\tilde{L}_\Psi\varepsilon, 1+2\tilde{L}_\Psi\varepsilon] \subset [1/2,3/2]$, it follows that the singular values of $D_x\Phi(x,\tau)$ coincide with its eigenvalues and therefore are also contained in the compact interval $[1-2\tilde{L}_\Psi\varepsilon, 1+2\tilde{L}_\Psi\varepsilon] \subset [1/2,3/2]$. Moreover, the matrix $D_x\Phi(x,\tau)$ is invertible and its inverse coincides with its pseudo-inverse. From the singular value decomposition of $D_x\Phi(x,\tau)$, we have that its pseudo-inverse $D_x\Phi(x,\tau)^\dagger$ is given by

$$D_x\Phi(x,\tau)^\dagger = U(x,\tau)\Sigma(x,\tau)^\dagger V(x,\tau)^\top,$$

However, $\Sigma(x,\tau)^\dagger$ is simply the inverse of $\Sigma(x,\tau)$ which is well-defined since $\Sigma(x,\tau)$ is a diagonal matrix whose diagonal entries belong to the compact interval $[1-2\tilde{L}_\Psi\varepsilon, 1+2\tilde{L}_\Psi\varepsilon] \subset [1/2,3/2]$. Therefore, we have that

$$D_x\Phi(x,\tau)^{-1} = D_x\Phi(x,\tau)^\dagger = U(x,\tau)\Sigma(x,\tau)^{-1}V(x,\tau)^\top,$$

and, using the properties of the operator norm of matrices, we have that

$$\begin{aligned}
\left|D_x\Phi(x,\tau)\right| &\leq |U(x,\tau)|\,|\Sigma(x,\tau)|\,|V(x,\tau)|,\\
\left|D_x\Phi(x,\tau)^{-1}\right| &\leq |U(x,\tau)|\left|\Sigma(x,\tau)^{-1}\right||V(x,\tau)|.
\end{aligned}$$

Since $U(x,\tau)$ and $V(x,\tau)$ are orthonormal matrices, it follows that $|U(x,\tau)| = |V(x,\tau)| = 1$. In addition, since $\Sigma(x,\tau)$ is a diagonal matrix whose diagonal entries belong to the compact interval $[1-2\tilde{L}_\Psi\varepsilon, 1+2\tilde{L}_\Psi\varepsilon] \subset [1/2,3/2]$, we have that $|\Sigma(x,\tau)| \leq 1+2\tilde{L}_\Psi\varepsilon \leq \frac{3}{2}$, and $\left|\Sigma(x,\tau)^{-1}\right| \leq \frac{1}{1-2\tilde{L}_\Psi\varepsilon} \leq 2$. However, since $0 \leq \varepsilon \leq \frac{1}{4\tilde{L}_\Psi}$, then $\frac{1}{1-2\tilde{L}_\Psi\varepsilon} \leq 1+4\tilde{L}_\Psi\varepsilon$. Therefore, we have that

$$\left|D_x\Phi(x,\tau)\right| \leq 1+2\tilde{L}_\Psi\varepsilon, \qquad \left|D_x\Phi(x,\tau)^{-1}\right| \leq 1+4\tilde{L}_\Psi\varepsilon.$$

It follows that the inverse of the Jacobian matrix $D\Psi(x,\tau)$ is well-defined and is given by (31), [44, p. 146], which concludes the proof of the claim. ■

B.2. Proof of Claim 2

Proof. The map g from item (d) in Proposition 1 has the explicit form

$$g(x,\tau,\varepsilon) = \sum_{k=1}^{5} G_i(x,\tau,\varepsilon)\, g_i(x,\tau,\varepsilon),$$

where the matrix-valued maps G_i are uniformly bounded and the maps g_i are given by

$$\begin{aligned}
g_1(x,\tau) &= v_1\circ\Psi^{-1}(x,\tau), & g_2(x,\tau) &= -f_2\circ\Psi^{-1}(x,\tau),\\
g_3(x,\tau) &= -f_1\circ\Psi^{-1}(x,\tau), & g_4(x,\tau) &= v_2\circ\Psi^{-1}(x,\tau),\\
g_5(x,\tau) &= -f_2\circ\Psi^{-1}(x,\tau).
\end{aligned}$$

In this case, since $\delta_1 = \delta_2 = 0$, we have that $\hat{f}_k(x,\tau) = f_k(x,\tau)$, and therefore we obtain that

$$f_1(x,\tau) = \sum_{i=1}^{r}\sum_{j=1}^{2} b_{i,j}\sqrt{2\omega_i}\xi_i(J(x)+\omega_i\tau), f_2(x,\tau) = b_0(x),$$

where $\xi_1(s) = \cos(s)$ and $\xi_2(s) = \sin(s)$. Clearly, for all $(x,\tau) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, we have that

$$|f_1(x,\tau)| \leq \sum_{i=1}^{r}\sum_{j=1}^{2}\sqrt{2\omega_i}|b_{i,j}|, \qquad |f_2(x,\tau)| \leq |\nabla J(x)| \leq M_J.$$

where used item (b) in Assumption 5. In addition, direct differentiation shows that

$$|D_x f_1(x,\tau)| \leq \sum_{i=1}^{r}\sum_{j=1}^{2}\sqrt{2\omega_i}|b_{i,j}||\nabla J(x)|,$$

which is also uniformly bounded due to item (b) in Assumption 5. Since v_1 is the integral of f_1 with respect to τ and is periodic in τ, it follows that v_1 is also uniformly bounded. Similarly, v_2 is the integral with respect to τ of terms that (smoothly) depend on f_1, $D_x f_1$, and f_2, all of which are uniformly bounded, and is periodic in τ. It follows that v_2 is also uniformly bounded. Finally, since Ψ^{-1} is diffeomorphism, it follows that all the maps g_i for $i \in \{1,\ldots,5\}$ are uniformly bounded. Therefore, the remainder map g is also uniformly bounded. ■

Appendix C. Auxiliary lemmas

Lemma 9. *For any fixed $\delta_1 > 0$, there exists a constant $L_{\delta_1} > 0$ such that: $|f_2(x_1,\tau) - f_2(x_2,\tau)| \leq L_{\delta_1}|x_1 - x_2|$, for all $x_1, x_2 \in \mathbb{R}\backslash(-\delta_1,\delta_1)$ and all $\tau \in \mathbb{R}_{\geq 0}$, where $f_2 : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is the function given by $f_2(x,\tau) = -|x|^{\frac{1}{2}}\mathit{sign}(x)\sin(\tau)^2$.*

Proof. Since $f_2(\cdot,\tau) : \mathbb{R} \to \mathbb{R}$ is absolutely continuous for all $\tau \in \mathbb{R}_{\geq 0}$, and by invoking the fundamental theorem of calculus, we have that

$$|f_2(x_1,\tau) - f_2(x_2,\tau)| \leq \int_0^1 |D_x f_2(x(\lambda),\tau)| d\lambda\, |x_1 - x_2|.$$

where $x(\lambda) := x_2 + \lambda(x_1 - x_2)$, and $D_x f_2(x,\tau) = -\frac{1}{2}|x|^{-\frac{1}{2}}\sin(\tau)^2$. In particular, for all $\tau \in \mathbb{R}_{\geq 0}$ and all $\delta_1 \in (0,\infty)$, the map $D_x f_2(\cdot,\tau) : [-\delta_1,\delta_1] \to \mathbb{R}$ belongs to the function space $\mathcal{L}^1([-\delta_1,\delta_1])$. Fix $\delta_1 \in (0,\infty)$, and let $x_1, x_2 \in \mathbb{R}\backslash(-\delta_1,\delta_1)$ be two arbitrary points. If $x_1 = x_2$, then there is nothing to prove. Therefore, without loss of generality, we may assume that $x_1 > x_2$. Then, only one of the following cases holds:

(C1) $x_1, x_2 \in [+\delta_1, +\infty)$, which implies that $x(\lambda) \in [\delta_1,\infty)$, for all $\lambda \in [0,1]$. Consequently, for all $\tau \in \mathbb{R}_{\geq 0}$:

$$\int_0^1 |D_x f_2(x(\lambda),\tau)|\, d\lambda \leq \frac{1}{2\sqrt{\delta_1}}.$$

(C2) $x_1, x_2 \in (-\infty, -\delta_1]$, which implies that $x(\lambda) \in (-\infty,-\delta_1]$, for all $\lambda \in [0,1]$. Consequently, for all $\tau \in \mathbb{R}_{\geq 0}$:

$$\int_0^1 |D_x f_2(x(\lambda),\tau)|\, d\lambda \leq \frac{1}{2\sqrt{\delta_1}}.$$

(C3) $x_1 \in [+\delta_1,+\infty)$ and $x_2 \in (-\infty,-\delta_1]$, which implies that there exists $\lambda_1, \lambda_2 \in [0,1]$, such that $x(\lambda_1) = \delta_1$, $x(\lambda_2) = -\delta_2$, $\lambda_1 > \lambda_2$, and the following relations hold

$$\begin{aligned}
x(\lambda) &\in (-\infty,-\delta_1], & \forall \lambda &\in [0,\lambda_2],\\
x(\lambda) &\in [-\delta_1,+\delta_1], & \forall \lambda &\in [\lambda_2,\lambda_1],\\
x(\lambda) &\in [+\delta_1,+\infty), & \forall \lambda &\in [\lambda_1,1].
\end{aligned}$$

Using the properties of the integral, we have that

$$\begin{aligned}
\int_0^1 |D_x f_2(x(\lambda),\tau)|\, d\lambda \leq &\int_0^{\lambda_2} |D_x f_2(x(\lambda),\tau)|\, d\lambda\\
&+ \int_{\lambda_2}^{\lambda_1} |D_x f_2(x(\lambda),\tau)|\, d\lambda\\
&+ \int_{\lambda_1}^{1} |D_x f_2(x(\lambda),\tau)|\, d\lambda,
\end{aligned} \tag{C.1}$$

for all $\tau \in \mathbb{R}_{\geq 0}$. Therefore, we have that

$$\int_0^1 |\mathrm{D}_x f_2(x(\lambda), \tau)|\, \mathrm{d}\lambda \leq \frac{1}{\sqrt{\delta_1}} + \int_{-\delta_1}^{\delta_1} |\mathrm{D}_x f_2(y, \tau)|\, \mathrm{d}y. \tag{C.2}$$

Finally, we compute $\int_{-\delta_1}^{\delta_1} |\mathrm{D}_x f_2(y, \tau)|\, \mathrm{d}y = 2\sqrt{\delta_1}$.

By defining L_{δ_1} with $L_{\delta_1} := \frac{1}{\sqrt{\delta_1}} + 2\sqrt{\delta_1}$, the proof of the Lemma is concluded. ■

Lemma 10. *Consider system (4) with $x \in \mathbb{R}$, $f_1(x, \tau) = 0$, and*

$$f_2(x, \tau) = -|x|^{\frac{1}{2}} \mathit{sign}(x) \sin(\tau)^2,$$

and let g be the map generated by Proposition 1. Then, there exists a constant $\bar{L}_g$ such that

$$|g(x, \tau, \varepsilon)| \leq \bar{L}_g(|x|^{\frac{1}{2}} + 1),$$

for all $(x, \tau, \varepsilon) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times (0, \varepsilon_0)$.

Proof. As shown in Lemma 9, for any $\delta_1 \in (0, \infty)$ there exists a constant $L_{\delta_1} > 0$ such that

$$|f_2(x_1, \tau) - f_2(x_2, \tau)| \leq L_{\delta_1} |x_1 - x_2|,$$

for all $x_1, x_2 \in \mathbb{R}$ and all $\tau \in \mathbb{R}_{\geq 0}$. Therefore, and since $f_1 = 0$, system (4) satisfies Assumption 1 for any $\delta_1 \in (0, \infty)$. Consequently, for any (fixed) choice of δ satisfying (7), there exists ε_0 such that all assertions of Proposition 1 are true. In particular, the pushforward of system (4) under the action of the diffeomorphism Ψ generated by Proposition 1 satisfies

$$\dot{x} = \bar{f}(x) + \varepsilon\, g(x, \tau, \varepsilon),$$

where the averaged vector field $\bar{f}$ is given by

$$\bar{f}(x)(x) = -\frac{1}{2}\varphi(x)|x|^{\frac{1}{2}} \operatorname{sign}(x).$$

Using formulas (9), we compute that

$$v_2(x, \tau) = \frac{1}{2}\varphi(x)|x|^{\frac{1}{2}} \sin(2\tau) \implies |v_2(x, \tau)| \leq \frac{1}{2}|x|^{\frac{1}{2}},$$

for all $(x, \tau) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$. From the proof of Proposition 1, we obtain that the map g has the explicit expression

$$\begin{aligned} g(x, \tau, \varepsilon) &= \varepsilon \tilde{F}(x, \tau) v_2 \circ \Psi^{-1}(x, \tau) \\ &\quad - \varepsilon \mathrm{D}_x v_2 \circ \Psi^{-1}(x, \tau) f_2 \circ \Psi^{-1}(x, \tau), \end{aligned}$$

for all $(x, \tau, \varepsilon) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times (0, \varepsilon_0)$, where ε_0 is the constant generated by Proposition 1, and $\tilde{F}$ is the map given by

$$\begin{aligned} \tilde{F}(x, \tau) &= \bar{F}(\pi_1 \circ \Psi^{-1}(x, \tau), x), \\ \bar{F}(x_1, x_2) &= \int_0^1 \mathrm{D}_x \bar{f}(\lambda x_1 + (1 - \lambda) x_2)\, \mathrm{d}\lambda. \end{aligned}$$

In addition, Lemma 3 implies that there exist constants $\bar{L}$ and $\bar{L}_{v,2}$ such that

$$|\mathrm{D}\bar{f}(x)| \leq \bar{L}, \qquad |\mathrm{D}_x v_2(x, \tau)| \leq L_{v,2},$$

$(x, \tau) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$. It follows that

$$|\tilde{F}(x, \tau)| \leq \bar{L}, \qquad |\mathrm{D}_x v_2 \circ \Psi^{-1}(x, \tau)| \leq L_{v,2},$$

$(x, \tau) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$. Hence, the map g satisfies the upper bound

$$\begin{aligned} |g(x, \tau, \varepsilon)| &\leq \varepsilon(\bar{L}|v_2 \circ \Psi^{-1}(x, \tau)| + L_{v,2}|f_2 \circ \Psi^{-1}(x, \tau)|) \\ &\leq \varepsilon(\bar{L} + L_{v,2})|\Psi^{-1}(x, \tau)|^{\frac{1}{2}}, \end{aligned}$$

for all $(x, \tau, \varepsilon) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times (0, \varepsilon_0)$. On the other hand, using Lemma 4, it can be shown that, for all $(x, \tau) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$, the diffeomorphism Ψ^{-1} satisfies the upper bound

$$|\Psi^{-1}(x, \tau)| \leq (1 + L_\Psi \varepsilon)|x| + L_\Psi \varepsilon.$$

Hence, we have that

$$\begin{aligned} |g(x, \tau, \varepsilon)| &\leq \varepsilon(\bar{L} + L_{v,2})((1 + L_\Psi \varepsilon)|x| + L_\Psi \varepsilon)^{\frac{1}{2}} \\ &\leq \varepsilon(\bar{L} + L_{v,2})\left((1 + L_\Psi \varepsilon)^{\frac{1}{2}}|x|^{\frac{1}{2}} + (L_\Psi \varepsilon)^{\frac{1}{2}}\right) \\ &\leq \varepsilon(\bar{L} + L_{v,2})\left(\sqrt{2}|x|^{\frac{1}{2}} + 1\right) \leq \bar{L}_g\left(|x|^{\frac{1}{2}} + 1\right), \end{aligned}$$

for all $(x, \tau, \varepsilon) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times (0, \varepsilon_0)$, where $\bar{L}_g = \varepsilon\sqrt{2}(\bar{L} + L_{v,2})$ is the sought after constant. ■

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a C^0 map and let $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ be a diffeomorphism. Let the map $\Psi_* f : \mathbb{R}^n \to \mathbb{R}^n$ be given by

$$\Psi_* f(x) = \mathrm{D}\Psi \circ \Psi^{-1}(x) f \circ \Psi^{-1}(x). \tag{C.3}$$

Clearly, the map $\Psi_* f$ is C^0. Consider the two ODEs

$$\dot{x} = f(x), \qquad x(0) = x_0 \tag{C.4}$$

$$\dot{\bar{x}} = \Psi_* f(\bar{x}), \qquad \bar{x}(0) = \bar{x}_0. \tag{C.5}$$

Since f and $\Psi_* f$ are both C^0, it follows that both ODEs (C.4) and (C.5) have the local existence of solutions property from any initial condition [48]. Then, we have the following Lemma.

Lemma 11. *Let $x : [0, t_s) \to \mathbb{R}^n$ be any maximal solution to (C.4) where $[0, t_s)$ is the maximal interval of existence of the solution, for some $t_s \in \mathbb{R}_{>0} \cup \{\infty\}$. Then there exists a (unique) maximal solution $\bar{x} : [0, t_s) \to \mathbb{R}^n$ to (C.5) such that $\bar{x}(t) = \Psi(x(t))$, for all $t \in [0, t_s)$. Conversely, if $\bar{x} : [0, t_s) \to \mathbb{R}^n$ is any maximal solution to (C.5), then there exists a (unique) solution $x : [0, t_s) \to \mathbb{R}^n$ to (C.4) such that $x(t) = \Psi^{-1}(\bar{x}(t))$, for all $t \in [0, t_s)$.*

Proof. If $x : [0, t_s) \to \mathbb{R}^n$ is a solution to (C.4), then, by definition, x is C^1 and, for every $t \in [0, t_s)$, we have that

$$\dot{x}(t) = f(x(t)), \qquad x(0) = x_0 \in \mathbb{R}^n. \tag{C.6}$$

Consider the map $\bar{x} = \Psi \circ x : [0, t_s) \to \mathbb{R}^n$. Since Ψ and x are C^1, it follows that $\bar{x}$ is also C^1. Using the chain rule, we have

$$\dot{\bar{x}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\Psi(x(t))) = \mathrm{D}\Psi(x(t))\dot{x}(t) = \mathrm{D}\Psi(x(t)) f(x(t)), \tag{C.7}$$

for all $t \in [0, t_s)$. However, since Ψ is a diffeomorphism, it has a C^1 inverse $\Psi^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, which implies that $x(t) = \Psi^{-1}(\bar{x}(t))$, for all $t \in [0, t_s)$. Therefore, we obtain that the map $\bar{x}$ satisfies

$$\dot{\bar{x}}(t) = \mathrm{D}\Psi(\Psi^{-1}(\bar{x}(t))) f(\Psi^{-1}(\bar{x}(t))) = \Psi_* f(\bar{x}(t)), \qquad \bar{x}(0) = \Psi(x_0),$$

for all $t \in [0, t_s)$. That is, the map $\bar{x} : [0, t_s) \to \mathbb{R}^n$ is a solution of system (C.5). To prove uniqueness, suppose by contradiction that $\tilde{x} : [0, t_s) \to \mathbb{R}^n$ is another maximal solution to (C.5) such that $\tilde{x}(t) = \Psi(x(t))$ for all $t \in [0, t_s)$, and $\exists t_e \in [0, t_s)$ such that $\tilde{x}(t_e) \neq \bar{x}(t_e)$. This implies that

$$\Psi(x(t_e)) = \tilde{x}(t_e) \neq \bar{x}(t_e) = \Psi(x(t_e)),$$

which is a clear contradiction. Conversely, suppose by contradiction that $\hat{x}$ is another maximal solution to (C.4) such that $\bar{x}(t) = \Psi(\hat{x}(t))$ for all $t \in [0, t_s)$, and $\exists t_e \in [0, t_s)$ such that $\hat{x}(t_e) \neq x(t_e)$. Since Ψ is a diffeomorphism, it follows that $\hat{x}(t) = \Psi^{-1}(\bar{x}(t))$, for all $t \in [0, t_s)$. This implies that

$$\Psi^{-1}(\bar{x}(t_e)) = \hat{x}(t_e) \neq x(t_e) = \Psi^{-1}(\bar{x}(t_e)),$$

which is also a clear contradiction. Therefore, if $x : [0, t_s) \to \mathbb{R}^n$ is a maximal solution to (C.4), then $\bar{x} = \Psi \circ x : [0, t_s) \to \mathbb{R}^n$ is the only maximal solution to (C.5) such that $\bar{x}(t) = \Psi(x(t))$, for all $t \in [0, t_s)$. The converse argument is identical if we replace Ψ by Ψ^{-1}. ■

References

[1] M. Leblanc, Sur l'electri''cation des chemins de fer au moyen de courants alternatifs de frequence elevee, Rev. Gen. Electr. (1922).

[2] K.B. Ariyur, M. Krstic, Real-Time Optimization by Extremum-Seeking Control, John Wiley & Sons, 2003.

[3] Y. Tan, W.H. Moase, C. Manzie, D. Nešić, I.M. Mareels, Extremum seeking from 1922 to 2010, in: Proceedings of the 29th Chinese Control Conference, IEEE, 2010, pp. 14–26.

[4] A. Scheinker, 100 Years of extremum seeking: A survey, Automatica 161 (2024) 111481.

[5] H. Khalil, Nonlinear Systems, Prentice Hall, 2002.

[6] J.A. Sanders, F. Verhulst, J. Murdock, Averaging Methods in Nonlinear Dynamical Systems, Vol. 59, Springer, 2007.

[7] M. Krstić, H.-H. Wang, Stability of extremum seeking feedback for general nonlinear dynamic systems, Automatica 36 (4) (2000) 595–601.

[8] Y. Tan, D. Nešić, I. Mareels, On non-local stability properties of extremum seeking control, Automatica 42 (6) (2006) 889–903.

[9] Y. Tan, D. Nešić, I. Mareels, A. Astolfi, On global extremum seeking in the presence of local extrema, Automatica 45 (1) (2009) 245–251.

[10] N. Mimmo, L. Marconi, G. Notarstefano, Uniform quasi-convex optimisation via extremum seeking, in: 2022 IEEE 61st Conference on Decision and Control, CDC, IEEE, 2022, pp. 6306–6311.

[11] D. Nesić, Y. Tan, W.H. Moase, C. Manzie, A unifying approach to extremum seeking: Adaptive schemes based on estimation of derivatives, in: 49th IEEE Conference on Decision and Control, CDC, IEEE, 2010, pp. 4625–4630.

[12] A. Ghaffari, M. Krstić, D. Nešić, Multivariable Newton-based extremum seeking, Automatica 48 (8) (2012) 1759–1767.

[13] T.R. Oliveira, M. Krstić, D. Tsubakino, Extremum seeking for static maps with delays, IEEE Trans. Autom. Control 62 (4) (2016) 1911–1926.

[14] T.R. Oliveira, J. Feiling, S. Koga, M. Krstić, Multivariable extremum seeking for PDE dynamic systems, IEEE Trans. Autom. Control 65 (11) (2020) 4949–4956.

[15] J.I. Poveda, A.R. Teel, A framework for a class of hybrid extremum seeking controllers with dynamic inclusions, Automatica 76 (2017) 113–126.

[16] J.I. Poveda, N. Li, Robust hybrid zero-order optimization algorithms with acceleration via averaging in time, Automatica 123 (2021) 109361.

[17] M. Abdelgalil, D.E. Ochoa, J.I. Poveda, Multi-time scale control and optimization via averaging and singular perturbation theory: From ODEs to hybrid dynamical systems, Annu. Rev. Control 56 (2023) 100926.

[18] H.-B. Dürr, M.S. Stanković, C. Ebenbauer, K.H. Johansson, Lie bracket approximation of extremum seeking systems, Automatica 49 (6) (2013) 1538–1552.

[19] C. Labar, E. Garone, M. Kinnaert, C. Ebenbauer, Newton-based extremum seeking: A second-order Lie bracket approximation approach, Automatica 105 (2019) 356–367.

[20] M. Abdelgalil, H. Taha, Recursive averaging with application to bio-inspired 3-D source seeking, IEEE Control Syst. Lett. 6 (2022) 2816–2821.

[21] H.-B. Dürr, M.S. Stanković, K.H. Johansson, C. Ebenbauer, Examples of distance-based synchronization: An extremum seeking approach, in: 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2013, pp. 366–373.

[22] H.-B. Dürr, M.S. Stanković, K.H. Johansson, C. Ebenbauer, Extremum seeking on submanifolds in the Euclidian space, Automatica 50 (10) (2014) 2591–2596.

[23] M. Abdelgalil, J.I. Poveda, On Lie-bracket averaging for a class of hybrid dynamical systems with applications to model-free control and optimization, 2023, arXiv preprint arXiv:2308.15732.

[24] R. Suttner, Extremum-seeking control for a class of mechanical systems, IEEE Trans. Autom. Control 68 (2) (2022) 1200–1207.

[25] R. Suttner, M. Krstić, Extremum seeking control for fully actuated mechanical systems on Lie groups in the absence of dissipation, Automatica 152 (2023) 110945.

[26] R. Suttner, M. Krstić, Attitude optimization by extremum seeking for rigid bodies actuated by internal rotors only, Internat. J. Robust Nonlinear Control (2023).

[27] R. Suttner, M. Krstić, Nonholonomic source seeking in three dimensions using pitch and yaw torque inputs, Systems Control Lett. 178 (2023) 105584.

[28] A. Scheinker, M. Krstić, Extremum seeking with bounded update rates, Systems Control Lett. 63 (2014) 25–31.

[29] A. Scheinker, M. Krstić, Non-C2 Lie bracket averaging for nonsmooth extremum seekers, J. Dyn. Syst. Meas. Control 136 (1) (2014).

[30] M. Abdelgalil, H. Taha, Lie bracket approximation-based extremum seeking with vanishing input oscillations, Automatica 133 (2021) 109735.

[31] V. Grushkovskaya, A. Zuyev, C. Ebenbauer, On a class of generating vector fields for the extremum seeking problem: Lie bracket approximation and stability properties, Automatica 94 (2018) 151–160.

[32] W. Jongeneel, E. Moulay, Topological Obstructions to Stability and Stabilization: History, Recent Advances and Open Problems, Springer Nature, 2023.

[33] C.G. Mayhew, A.R. Teel, On the topological structure of attraction basins for differential inclusions, Systems Control Lett. 60 (12) (2011) 1045–1050.

[34] N. Mimmo, L. Marconi, G. Notarstefano, An extremum seeking scheme with global convergence properties, IFAC-PapersOnLine 54 (14) (2021) 454–459.

[35] C.K. Lauand, S. Meyn, Quasi-stochastic approximation: Design principles with applications to extremum seeking control, IEEE Control Syst. Mag. 43 (5) (2023) 111–136.

[36] S. Sastry, M. Bodson, J.F. Bartram, Adaptive control: stability, convergence, and robustness, 1990.

[37] A.R. Teel, D. Nesic, Averaging for a class of hybrid systems, Dyn. Contin. Discrete Impuls. Syst. 17 (6) (2010) 829–851.

[38] A. Scheinker, X. Pang, L. Rybarcyk, Model-independent particle accelerator tuning, Phys. Rev. Spec. Top. Accel. Beams 16 (10) (2013) 102803.

[39] R. Suttner, S. Dashkovskiy, Exponential stability for extremum seeking control systems, IFAC-PapersOnLine 50 (1) (2017) 15464–15470.

[40] J.M. Lee, Smooth Manifolds, Springer, 2012.

[41] A. Scheinker, M. Krstić, Minimum-seeking for CLFs: Universal semiglobally stabilizing feedback under unknown control directions, IEEE Trans. Autom. Control 58 (5) (2012) 1107–1122.

[42] Y. Nesterov, Lectures on Convex Optimization, Springer, 2018.

[43] H. Karimi, J. Nutini, M. Schmidt, Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition, in: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2016, Riva Del Garda, Italy, September 19-23, 2016, Proceedings, Part I 16, Springer, 2016, pp. 795–811.

[44] D.S. Bernstein, Matrix mathematics: Theory, Facts, and Formulas, Princeton University Press, 2009.

[45] J. Nestruev, A. Bocharov, S. Duzhin, Smooth Manifolds and Observables, Springer, 2003.

[46] C. Cai, A.R. Teel, Characterizations of input-to-state stability for hybrid systems, Systems Control Lett. 58 (1) (2009) 47–53.

[47] B. Polyak, Introduction to Optimization, Optimization Software Inc., NY, 2010.

[48] A.F. Filippov, Differential Equations with Discontinuous Righthand Sides: Control Systems, Vol. 18, Springer Science & Business Media, 1985.