Extremum seeking control of nonlinear dynamic systems using Lie bracket approximations

Victoria Grushkovskaya1,2 | Christian Ebenbauer3

1Institute of Mathematics, Alpen-Adria University of Klagenfurt, Klagenfurt, Austria
2Institute of Applied Mathematics and Mechanics, National Academy of Sciences of Ukraine, Sloviansk, Ukraine
3Institute for Systems Theory and Automatic Control, University of Stuttgart, Stuttgart, Germany

Correspondence
Victoria Grushkovskaya, Institute of Mathematics, Alpen-Adria University of Klagenfurt, Klagenfurt, Austria. Email: viktoriia.grushkovska@aau.at

Summary
In this article, we consider extremum seeking problems for a general class of nonlinear dynamic control systems. The main result of the article is a broad family of control laws which optimize the steady-state performance of the system. We prove practical asymptotic stability of the optimal steady-state and, moreover, propose sufficient conditions for the asymptotic stability in the sense of Lyapunov. The results generalize and extend existing results which are based on Lie bracket approximations. In particular, our approach does not rely on singular perturbation theory, as commonly used in extremum seeking of nonlinear dynamic systems.

KEYWORDS
Chen-Fliess series, extremum seeking, Lie bracket approximations, nonlinear dynamic systems, stability

1 | INTRODUCTION

Many problems in applications require the stabilization of a control system at some optimal operating point. An optimal operating point is often characterized as a state of the system which is an equilibrium point and where a given state-dependent cost (performance) function takes its minimal or maximal value. For complex systems such an optimal state cannot be determined easily because a model is often inaccurate or is not available at all. In addition, in applications the optimal state is often time-varying and state measurements are not available. Thus, the development of model-free real-time optimizing control laws, known as extremum seeking control laws, that stabilize an a priori unknown optimal operating point is both challenging in theory and highly relevant in practical applications.

Extremum seeking has been extensively studied in the control literature and can be traced back to the early 1920s. However, a solid theoretical foundation of extremum seeking control has been developed only over the last two decades (see, eg, Reference 1 for the literature overview). Today there exist a number of results, both on theoretical studies and practical applications (see, eg, References 2-20). Classical extremum seeking control laws exploit time-periodic oscillating input perturbations in order to find and stabilize the optimal steady-state of the system without the knowledge of a system model. One such perturbation-based framework for the analysis and design of extremum seeking systems was proposed in Reference 9 and relies on Lie bracket approximations. The main idea therein is that trajectories of the extremum seeking system approximate trajectories of a so-called Lie bracket system which corresponds to a gradient-like dynamics which...
optimizes the cost function. Based on the Lie bracket system and its corresponding extremum seeking system, a whole analysis and design framework has been established, see, for example, References 16,17,21-32. In particular, extremum seeking for dynamic nonlinear systems using Lie bracket approximations has been addressed in Reference 25. In that article, a combination of Lie bracket approximations and singular perturbations techniques (time-scale separation) has been proposed. In general, for dynamic extremum seeking systems, a singular perturbation approach is quite common, see, for example, the articles 8,24,11,17,33,34 In this article, however, we propose an alternative approach without singular perturbation theory using the techniques developed in Reference 16. In article 16, we extended the results of Reference 9 and gave a rather general description for a whole family of extremum seeking control laws. Using the Chen-Fliess series, it has been shown in Reference 16 that the proposed control laws achieve practical stability and, under additional assumptions, asymptotic stability in the sense of Lyapunov. However, the results in Reference 16 do not directly apply to nonlinear control systems with dynamic maps.

Therefore, the goal of this article is to extend the results of paper 16 to dynamic systems utilizing Chen-Fliess series techniques and to extend the results in Reference 25 in terms of a broader class of control laws and more general and stronger stability results. In particular, the main contributions of this article are as follows: First, we introduce a family of extremum seeking control laws for rather general nonlinear dynamic systems with a well-defined steady-state map. Second, we analyze the asymptotic stability properties of the extremum seeking system with the proposed control laws. Specifically, we prove practical asymptotic stability of the optimal steady-state, and in addition, we provide conditions to achieve asymptotic stability in the sense of Lyapunov. In contrast to References 2,4,11,17,25,33, our prove method solely relies on Chen-Fliess series techniques and does not depend on classical singular perturbation arguments. Hence our approach provides a general and monolithic way to the analysis and design of extremum seeking systems based on the Lie bracket approximation framework.

The remainder of this article is organized as follows. The problem statement is given in Section 2. Section 3 contains the main results of the article. The extremum seeking control laws are described in Subsection 3.1, while sufficient conditions for practical asymptotic stability and asymptotic stability in the sense of Lyapunov are stated in Subsection 3.2. In Section 4, we illustrate the obtained results by an academic example. The appendix contains some auxiliary statements and the proofs of the main results.

**Notations and definitions.**

Throughout the article, we use the following notations:

- \( k = 1, \ldots, n \)—the integer number \( k \) varies from \( 1 \) to \( n \in \mathbb{N} \);
- \( \delta \)—Kronecker delta: \( \delta_{ij} = 1 \) and \( \delta_{ij} = 0 \) whenever \( i \neq j \);
- \( \mathbb{R}^+ \)—the set of all nonnegative real numbers;
- \( B_\delta(x^*) \)—\( \delta \)-neighborhood of an \( x^* \in \mathbb{R}^n \) with \( \delta > 0 \);
- \( \partial M, \overline{M} \)—the boundary and the closure of a set \( M \subset \mathbb{R}^n \), respectively; \( \overline{M} = M \cup \partial M \);
- \( \mathcal{K} \)—the class of continuous strictly increasing function \( \phi : \mathbb{R}^+ \to \mathbb{R}^+ \) such that \( \phi(0) = 0 \);
- \( [f,g](x) \)—the Lie bracket of \( f, g : \mathbb{R}^n \to \mathbb{R}^n \) at the point \( x \in \mathbb{R}^n \), \( [f,g](x) = L_f g(x) - L_g f(x) \).
- \( L_{jf}(x) = \lim_{s \to 0} \frac{f(x + s g(x)) - f(x)}{s} \). Notice that the limit may exist at \( x \in \mathbb{R}^n \) even if \( f,g \) are not differentiable at \( x \).

**Definition 1.** A point \( x^* \in \mathbb{R}^n \) is said to be **singly practically uniformly asymptotically stable** for the system

\[
\dot{x} = f^* \eta(t,x), \quad x \in \mathbb{R}^n, \quad t \in \mathbb{R}^+, \quad \epsilon, \eta > 0, \quad f^* : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n,
\]

with respect to a set \( S \subseteq \mathbb{R}^n \), if \( x^* \) is:

- **singly practically uniformly stable** for the above system, that is, for any \( \rho \) there exist \( \delta > 0, \bar{\epsilon} > 0 \), and \( \overline{\eta} : (0, +\infty) \to (0, +\infty) \) such that, for any \( t_0 \in \mathbb{R}^+ \), \( \epsilon \in (0, \bar{\epsilon}] \), \( \eta \in [\overline{\eta}(\epsilon), +\infty) \), if \( x(t_0) \in B_\delta(x^*) \) then \( x(t) \in B_\rho(x^*) \) for all \( t \in [t_0, +\infty) \);
- **singly practically uniformly attractive with respect to a set** \( S \subseteq \mathbb{R}^n \), that is, for any \( \delta > 0 \) such that \( B_\delta(x^*) \subset S \) and any \( \rho > 0 \) there exist \( \tau \geq 0, \bar{\epsilon} > 0 \), and \( \overline{\eta} : (0, +\infty) \to (0, +\infty) \) such that, for any \( t_0 \in \mathbb{R}^+ \), \( \epsilon \in (0, \bar{\epsilon}] \), \( \eta \in [\overline{\eta}(\epsilon), +\infty) \), if \( x(t_0) \in S \) then \( x(t) \in B_\rho(x^*) \) for all \( t \in [t_0 + \tau, +\infty) \).

If the above conditions hold for any \( \epsilon \in (0, \bar{\epsilon}] \), \( \eta \in [\overline{\eta}(\epsilon), +\infty) \) with some fixed \( \bar{\epsilon} > 0 \) and \( \overline{\eta}(\epsilon) > 0 \), then a point \( x^* \) is said to be **uniformly asymptotically stable** for the given system with respect to a set \( S \subseteq \mathbb{R}^n \).
Let us emphasize that the parameter \( \eta \) depends on parameter \( \varepsilon \) in Definition 1, which distinguishes it from the definition of practical asymptotic stability. Furthermore, the case \( S = \mathbb{R}^n \) means global attractivity (and asymptotic stability), while \( S \subseteq \mathbb{R}^n \) means that the above properties hold locally (in a set \( S \)).

## 2 PROBLEM STATEMENT

Consider the system

\[
\begin{align*}
\dot{x} &= f(x, u), \\
y &= h(x, u),
\end{align*}
\]

where \( x \in D_x \subseteq \mathbb{R}^{n_x} \) is the state vector, \( x(t_0) = x^0 \in D_x \) and \( t_0 \in \mathbb{R}^+ \) are the initial data, \( u \in D_u \subseteq \mathbb{R}^{n_u} \) is the control input, \( y \in D_y \subseteq \mathbb{R}^{n_y} \) is the system output, \( f \in C^1(D_x \times D_u; \mathbb{R}^{n_x}), h \in C^0(D_x \times D_u; \mathbb{R}^{n_y}) \), and \( D_x, D_u, D_y \) are domains. If the output of system (1) depends only on \( u \), that is, \( y(t) = h(u(t)) \), then we say (1) is a static system (or static map), otherwise, we say (1) is a dynamic system (or dynamic map).

**Definition 2.** Consider system (1) and assume that there exists a map \( \ell : D_u \to D_x \) such that for each \( \bar{u} \in D_u \), \( x^0 \in D_x \), \( t_0 \in \mathbb{R}^+ \), the solution \( x(t) \) of system (1) with \( u(t) \equiv \bar{u} \) and \( x(t_0) = x^0 \) satisfy

\[
h(x(t), \bar{u}) \to h(\ell(\bar{u}), \bar{u}) \text{ as } t \to +\infty.
\]

Then the map \( u \to h(\ell(u), u) \) is called the steady-state map of system (1).

In this article, we consider the following problem setup.

**Problem 1.** Consider system (1) and assume its steady-state map exists. Let \( J : D_y \to \mathbb{R} \) be a cost (performance) function which depends on the output of system (1) and assume that there exists a unique minimizer of the function

\[
\hat{J}(u) := J(h(\ell(u), u))
\]

in \( D_u \). We aim to construct a control law \( u(J(y(t)), t) \) which optimizes the steady-state performance in the following sense: for any \( x^0 \in D_x \), \( t_0 \in \mathbb{R} \), the solutions \( x(t) \) of system (1) with input \( u(J(y(t)), t) \) and \( x(t_0) = x^0 \) possess the property

\[
J(y(t)) \to J^* = \min_{u \in D_u} \hat{J}(u) \text{ as } t \to +\infty.
\]

Hence system (1) converges to the so-called optimal steady-state \( x = \ell(u^*) \) if the optimal steady-state input

\[
u^* = \arg \min_{u \in D_u} \hat{J}(u)
\]

is applied to the system. Notice that we assume that the control laws only depend on performance output measurements. We do not assume that \( f, h, \ell, J \) are known.

As it is shown in Reference 25, under suitable conditions, the dynamic control law

\[
u = \frac{1}{\eta \sqrt{\varepsilon}} \sum_{j=1}^{n_e} \sqrt{j} \left( J(y) \cos \left( \frac{jt}{\eta \varepsilon} \right) + \sin \left( \frac{jt}{\eta \varepsilon} \right) \right) e_j
\]

solves Problem 1 in the sense that it guarantees singular practical uniform asymptotic stability of the point \( q^* = (\ell(u^*), u^*) \) for the composed system given by system (1) and the dynamic control law given above. Here, \( j \in \mathbb{N}, e_j \) denotes the \( j \)th unit vector in \( \mathbb{R}^{n_u} \), and \( \varepsilon, \eta \) are some positive parameters. The main idea behind the approach of the article Reference 25 is to deduce the singular practical asymptotic stability of (1) with the above control law from the asymptotic stability properties of the so-called boundary layer model.
\[
\dot{x} = f(\dot{x} + \varepsilon(\tilde{u}), \tilde{u})
\]

and the practical asymptotic stability properties of the so-called reduced system

\[
\dot{\tilde{u}} = \frac{1}{\eta \sqrt{\varepsilon}} \sum_{j=1}^{n_u} \sqrt{j} \left( J(\tilde{u}) \cos \left( \frac{j \eta t}{\eta \varepsilon} \right) + \sin \left( \frac{j \eta t}{\eta \varepsilon} \right) \right) e_j
\]

with \( \tilde{x}(t_0) \in D_x, \tilde{u}(t_0) \in D_u \). The asymptotic practical stability for the reduced system is established by using the Lie bracket approximation approach,\(^{21}\) while the stability properties of the actual closed-loop system are inferred by the stability properties of the boundary layer model and the reduced system using singular perturbation results.\(^{35,36}\) In particular, the parameter \( \varepsilon > 0 \) guarantees a time-scale separation when chosen small enough, while \( \eta > 0 \) guarantees a gradient-like dynamics for \( \tilde{J}(u) \) when chosen large enough (relatively to \( \varepsilon \)). In other words, the role of parameter \( \varepsilon > 0 \) is to ensure that the trajectories of the reduced system tend asymptotically to a neighborhood of \( \tilde{u}^* \), while \( \eta \) ensures that the system (1) is always in a quasi-steady-state (and thus is quasi static system with respect to the dynamics of dynamic control law).

### 3 MAIN RESULTS

#### 3.1 Extremum seeking control laws

In this section, we will introduce a whole family of control laws for the solution of Problem 1 provided that system (1) satisfies certain assumptions.

Consider system (1) and let \( u^* \in D_u \) be the optimal steady-state input, that is, the unique minimizer of the function \( J(\varphi'(u), u) \). For stabilizing the optimal steady-state \( x = \varphi(u^*) \), we propose the following family of dynamic output control laws:

\[
\dot{u} = \sum_{j=1}^{2n_u} g_j(J(y))v_j^{\eta}(t)e_j, \quad u(t_0) = u^0 \in D_u,
\]

where \( e_j \) denotes the unit vector in \( \mathbb{R}^{n_u} \) with nonzero \( j \)-th entry if \( j \leq n_u \) or with nonzero \((j - n_u)\)-th entry if \( n_u + 1 \leq j \leq 2n_u \). The time-varying dithers \( v_j^{\eta}(t) \) are given by

\[
v_j^{\eta}(t) = \begin{cases} 
\frac{2}{\eta \sqrt{\varepsilon}} \cos \left( \frac{2\kappa_j t}{\eta \varepsilon} \right), & \text{for } j = 1, n_u, \\
\frac{2}{\eta \sqrt{\varepsilon}} \sin \left( \frac{2\kappa_j t}{\eta \varepsilon} \right), & \text{for } j = n_u + 1, 2n_u,
\end{cases}
\]

with \( \varepsilon > 0, k_j \in \mathbb{N}, k_{j_1} \neq k_{j_2} \) for all \( j_1 \neq j_2 \), and the so-called control functions \( g_j, g_{j+n_u} : \mathbb{R} \rightarrow \mathbb{R} \) are such that for each \( \xi \in \mathbb{R} \),

\[
[g_j, g_{j+n_u}](\xi) = -\gamma, \quad \gamma > 0, \ j = 1, n_u.
\]

In Reference 16, we have proposed the formula

\[
g_{j+n_u}(\xi) = -\gamma g_j(\xi) \int \frac{d\xi}{g_j^2(\xi)}
\]

which fulfills the last condition, that is, that the Lie bracket is constant. In this article, we use a more convenient formula, that is, we parametrize the functions \( g_j, g_{j+n_u} \) by

\[
g_j(\xi) = r_j(\xi) \sin \phi_j(\xi), \quad g_{j+n_u}(\xi) = r_j(\xi) \cos \phi_j(\xi),
\]

where \( r_j, \phi_j \) are such that \( r_j^2(\xi) \frac{d\phi_j(\xi)}{d\xi} = \gamma \).
with \( r_j \in C(\mathbb{R}; \mathbb{R}) \), \( \phi_j \in C^1(\mathbb{R}; \mathbb{R}) \). Such a representation of the control functions \( g_j, g_{j+n} \) has also been used in Reference 37. The proposed choice of \( g_j, g_{j+n} \) ensures that the trajectories of the system

\[
\dot{u} = \sum_{j=1}^{2n} g_j(J(\bar{u}))v_j^\nu(t)e_j
\]

approximate trajectories of the so-called Lie bracket system, which in this case takes the form of a gradient descent dynamics

\[
\ddot{u} = -\gamma \nabla J(\bar{u}), \quad \bar{u} \in D_u, \quad \bar{u}(t_0) = u(t_0),
\]

see References 16, 21 for more details. If the functions \( g_j \circ J \circ h^1, g_{j+n} \circ J \circ h \in C^2(D_x \times D_u; \mathbb{R}) \), then this property can be exploited to prove the singular practical asymptotic stability of \( q^* = (\mathcal{C}(u^*), u^*) \) for system (1) and (2) using the approach of Reference 25. However, many functions satisfying (3) fail to satisfy the \( C^2 \) assumption at \( q^* \) (see Section 4), so that the results of Reference 25 are no longer applicable.

### 3.2 Stability results

In this subsection, we establish stability properties of the point \( q^* = (\mathcal{C}(u^*), u^*) \) for the closed-loop system (1) and (2).

We make the following assumptions:

A1.1) \( f \in C^1(D_x \times D_u; \mathbb{R}^{n_x}) \) and there exists a function \( \mathcal{C} \in C^2(D_u; \mathbb{R}) \) such that, for each fixed \( u \in D_u, f(x, u) = 0 \) if and only if \( x = \mathcal{C}(u) \). Moreover, the image \( \mathcal{C}[D_u] = \{\mathcal{C}(u) | u \in D_u\} \subseteq D_x \).

A2.1) The function \( J \circ h \in C^2(D_x \times D_u; \mathbb{R}), \tilde{J}(u) = J(h(\mathcal{C}(u), u)) \), satisfies the following properties for all \( u \in D_u \):

\[
\begin{align*}
\alpha_{11}(\|u - u^*\|) &\leq \tilde{J}(u) - \tilde{J}(u^*) \leq \alpha_{12}(\|u - u^*\|), \\
\alpha_{21}(\|u - u^*\|) &\leq \|\nabla \tilde{J}(u)\|^2 \leq \alpha_{22}(\|u - u^*\|), \\
\left\| \frac{\partial^2 \tilde{J}(u)}{\partial u^2} \right\| &\leq \alpha_3(\|u - u^*\|),
\end{align*}
\]

with some functions \( \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathcal{K} \), and a nondecreasing function \( \alpha_3 : \mathbb{R}^+ \to \mathbb{R}^+ \).

A3.1) For all \( j_1, j_2, j_3 \in \{1, \ldots, 2n_u\} \), the functions \( g_{j1} \circ J \circ h \in C(D_x \times D_u) \), \( g_{j2} \circ \tilde{J} \in C^2(D_u \setminus \{u^*\}; \mathbb{R}) \), and \( L_{e_{j1}}(g_{j2} \circ J)(g_{j2} \circ \tilde{J}), L_{e_{j2}}(g_{j1} \circ J) - L_{e_{j1}}(g_{j2}) \circ (g_{j1} \circ J) \in C(D_u; \mathbb{R}) \). Furthermore, for any compact subsets \( D'_u \subseteq D_u \) and \( D' \subseteq D_x \), the functions \( g_{j1} \circ J \) are Lipschitz continuous on \( D'_u \times D'_x \) with respect to its first argument, and the functions \( g_{j1} \circ J \) are Lipschitz continuous on \( D'_u \).

\(^*\) To simplify the presentation, in the rest of the article we put \( D_u := B_{\Delta_u}(u^*), D_x := B_{\Delta_x}(\mathcal{C}(u^*)) \), with some \( \Delta_x, \Delta_u \in (0, +\infty] \).

### 3.2.1 Results on singular practical asymptotic stability

Our first result concerns sufficient conditions for singular practical asymptotic stability of the point \( q^* = (\mathcal{C}(u^*), u^*) \) for system (1) and (2).

**Theorem 1.** Consider the system (1)-(2) with a cost function \( J \). Suppose the Assumptions A1.1), A2.1), and A3.1) for \( \gamma > 0 \) hold. Moreover, assume that

A4.1) the trivial solution \( x = 0 \) of the system \( \dot{x} = f(x + \mathcal{C}(u), u) \) is uniformly asymptotically stable with respect to \( D_x \),

uniformly in the parameter \( u \);
A5.1) for any $\delta_{\alpha u} > 0$ there exist $\nu_\varepsilon \geq 0$, $\eta_0 : (0, +\infty) \to (0, +\infty)$ such that, for any $\varepsilon > 0$, $\eta \in [\eta_0(\varepsilon), +\infty)$, the solutions of system (1)-(2) with initial conditions $x(t_0) = x^0 \in D_x$, $u(t_0) = u^0 \in D_u$ satisfy the properties $\|x(t) - \ell(u(t))\| \leq \nu_\varepsilon$ and $x(t) \in D_x$ for $t \in [t_0, t_0 + \varepsilon \eta]$ whenever $u(t) \in D_u$ and $\|x^0 - \ell(u^0)\| \leq \delta_{\alpha u}$.

Then $q^* = (\ell(u^*), u^*)$ is singularly practically uniformly asymptotically stable for system (1)-(2) with respect to $D_x \times B_{\eta_1(\Delta u)}(u^*)$.

The proof is in Appendix B1. It has to be emphasized that Theorem 1 ensures global asymptotic stability properties in case of radially unbounded $\alpha_{12}$ and $\Delta_x = +\infty$, $\Delta_u = +\infty$.

Note that Assumption A4.1) is common in extremum seeking literature, see, for example, References 17,25,38. Together with Assumption A1.1), this implies the existence of a Lyapunov function for system (1).39 However, we expect that a similar result can also be obtained for systems which do not admit a Lyapunov function. One of such cases is considered in Reference 40.

Assumption A5.1) ensures that the distance between the trajectory $x(t)$ and the quasi-steady-state $\ell(u(t))$ cannot be infinitely large over a time interval of length $\eta \varepsilon$, which means that subsystem (2) has to vary slow enough. In particular, for a static steady-state $\ell(u) \equiv \text{const}$ it means that $x(t)$ cannot escape to infinity in finite time $t = \eta \varepsilon$. In many cases, Assumption A5.1) directly follows from A4.1). One particular case is considered below.

Lemma 1. Consider the system (1)-(2) with a cost function $J$ and let Assumptions A1.1) to A3.1) be satisfied. Moreover, assume that

A4.2) there exist a function $V \in C^1(D_u \times D_x; \mathbb{R})$ and positive constants $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}$ such that, for all $x \in D_x$, $u \in D_u$,

$$\sigma_{11}\|x - l(u)\|^2 \leq V(x, u) \leq \sigma_{12}\|x - l(u)\|^2,$$

$$\frac{\partial V(x, u)}{\partial x} f(x, u) \leq -\sigma_{21} V(x, u),$$

$$\left\| \frac{\partial V(x, u)}{\partial u} \right\| \leq \sigma_{22}\|x - l(u)\|.$$

Then Assumptions A4.1) to A5.1) in Theorem 1 hold.

Note that Assumption A4.2) is widely used in singular perturbation theory (see, eg., References 25,41). The proof of Lemma 1 is in Appendix C1.

The proofs of the obtained results represent constructive procedure for choosing small enough $\varepsilon$ and large enough $\overline{\eta}(\varepsilon)$. In particular, it can be seen that, under the conditions of Lemma 1, it suffices to take $\overline{\eta} = \frac{\sigma}{\varepsilon \sqrt{\varepsilon}}$ with some $\sigma > 0$. This refines the estimate $\overline{\eta} = \frac{1}{\varepsilon^k}$ with $k > 2$ proposed in Reference 25. Unlike the approaches of papers,25,36 we do not apply singular perturbation theory to prove the obtained results. Instead, our proofs are similar to the techniques introduced in our article16 and exploit the Chen-Fliss series expansion of the $u$-component of the solutions of closed-loop system (1)-(2) and a thorough analysis of the behavior of the functions $J$ and $V$ along the trajectories of (1)-(2). In the particular case of the static system with $y = u$, the Problem 1 and its solution coincide with Reference 16. Consequently, the asymptotic stability properties of $\ell(u^*)$ for subsystem (1) can be established by the known results for systems with slowly varying parameters38 or stability conditions for a family of sets.27

### 3.2.2 Results on asymptotic stability in the sense of Lyapunov

Following the ideas of Reference 16, it is also possible to ensure that the solutions of system (1)-(2) with some fixed $\varepsilon$ and $\eta$ tend asymptotically to $q^* = (\ell(u^*), u^*)$. For this purpose, we require the following additional assumptions:

A2.2) the function $\hat{J}(u) = J(h(\ell(u), u))$ satisfies the following properties for all $u \in D_u$ with some $\bar{a}_{11}, \bar{a}_{12}, \bar{a}_{21}, \bar{a}_{22}, \bar{a}_3 > 0$:

$$\bar{a}_{11}\|u - u^*\|^2 \leq \hat{J}(u) - \hat{J}(u^*) \leq \bar{a}_{12}\|u - u^*\|^2,$$

$$\bar{a}_{21}(\hat{J}(u) - \hat{J}(u^*)) \leq \|\nabla \hat{J}(u)\|^2 \leq \bar{a}_{22}(\hat{J}(u) - \hat{J}(u^*)),$$

$$\left\| \frac{\partial^2 \hat{J}(u)}{\partial u^2} \right\| \leq \bar{a}_3.$$
A3.2) For all \( j_1, j_2, j_3 \in \{1, \ldots, 2n_u\} \), the functions \( g_{i} \) satisfy Assumption 1.3 and, moreover, there exist \( M_g > 0 \), \( M_{3g} \geq 0 \) such that

\[
\|g_i(\hat{J}(u))\| \leq M_g \sqrt{\hat{J}(u) - \hat{J}(u^*)},
\]

\[
\|L_{\varepsilon_{i_1} \varepsilon_{i_2} \varepsilon_{i_3}} g_i(\hat{J}(u))\| \leq M_{3g} \sqrt{\hat{J}(u) - \hat{J}(u^*)}, \quad \text{for all } x \in \mathcal{D}_x, u \in \mathcal{D}_u.
\]

**Theorem 2.** Consider the system (1)-(2) with a cost function \( J \). Suppose the Assumptions A1.1), A2.2), and A3.2) for \( \gamma > 0 \) hold. Moreover, assume that there exists a function \( V \) satisfying A4.2). Then the point \( q^* = (u^*, \ell(u^*)) \) is asymptotically stable with respect to \( \Delta_x = \Delta_u = +\infty \) provided that \( J(u^*) \leq J(u) \) along the trajectories of system (1)-(2), in general, cannot be guaranteed.

Assumption A4.2) implies that, for each fixed value of \( u \), the corresponding equilibrium point \( x = \ell(u) \) is exponentially stable for system (1), while A2.2) ensures an exponential decay of the function \( J \) along the trajectories of system

\[
\tilde{u} = \sum_{j=1}^{2n_u} g_j(\hat{J}(\tilde{u}))\nu^{*j}(\tilde{u}), \quad \tilde{u} \in \mathcal{D}_u.
\]

As it will be demonstrated in Section 4, the property of asymptotic stability in the sense of Lyapunov may not be satisfied if subsystem (1) exhibits only asymptotic (but not exponential) stability properties. Nevertheless, we expect that it is possible to prove Lemma 1 and Theorem 2 under relaxed assumptions on \( J \) and \( V \). We leave this issue for future studies.

Let us underline that in A3.2) the optimal cost value \( J(u^*) \) (but not the optimal state) is assumed to be known in order to guarantee that the vector fields of system (1)-(2) tend to 0 as the cost function approaches its optimum. Such situations naturally arise in many control problems where only the optimal cost value is defined a priori (eg, in regression, synchronization or target tracking tasks), or when the deviations between the cost function and its optimal value can be measured. Another application arises from vibrational stabilization problems, as discussed, for example, in Reference 16.

### 4 | Example

In this section, the proposed extremum control laws are illustrated by the following academic example. Consider a spring-mass system oscillating in horizontal plane with friction and control, that is, described by the equation

\[
m\ddot{l} + kl + \phi(l) + u = 0,
\]

where \( l > 0 \) is length of the spring, \( k > 0 \) is spring constant, \( m > 0 \) is mass, \( \phi(l) \geq 0 \) is damping, and \( u \) is control. In the sequel, we assume \( m = k = 1 \). We aim to stabilize system (4) at the state \( l = l^* \), \( \dot{l} = 0 \), assuming that only the measurements of \( y = l - l^* \) are available for control design. Obviously, this problem can be solved by using time-invariant controls. However, we apply the dynamic control laws of type (2) and use the cost function \( J(y) = y^2 \) to illustrate the main results of the article.

#### 4.1 | Case 1: Linear damping

Assume first that damping is linear, that is, \( \phi(l) = \mu l \) with some \( \mu > 0 \), and rewrite equation (4) in the variables \( x_1 = l \), \( x_2 = \dot{l} \):

\[
\dot{x}_1 = x_2,
\]

\[
\dot{x}_2 = -kx_1 - \mu x_2 - u,
\]

\[
y = x_1 - l^*.
\]
For each fixed value of $u$ system (5) possesses the exponentially stable equilibrium $\left( -\frac{u}{k}, 0 \right)$, so that A4.2 holds. Thus, the quasi-steady-state is given by $\mathcal{E}(u) = \left( -\frac{u}{k}, 0 \right)$. Furthermore, as $J(y) = y^2 = (x_1 - l^*)^2$, the function $\dot{J}(u) = \frac{1}{k^2}(u + kl^*)^2$ satisfies A2.2). Following the approach of Section 3, one may put

$$
\dot{u} = \frac{2\sqrt{\pi}}{\eta \sqrt{\epsilon}} \left( g_1(J(y)) \cos \frac{2\pi t}{\eta \epsilon} + g_2(J(y)) \sin \frac{2\pi t}{\eta \epsilon} \right),
$$

(6)

where $g_2(\xi) = -\gamma g_1(\xi) \int \frac{d\xi}{g_1(\xi)}, \gamma > 0$. We consider four pairs of functions $g_1, g_2$ with different properties. Namely, functions exploited in Reference 25,

$$
g_1(J(y)) = \sqrt{\gamma} J(y), \quad g_2(J(y)) = \sqrt{\gamma},
$$

(7)

uniformly bounded functions (see Reference 22)

$$
g_1(J(y)) = \sqrt{\gamma} \sin(J(y)), \quad g_2(J(y)) = \sqrt{\gamma} \cos(J(y)),
$$

(8)

and functions vanishing at the optimal points (see Reference 29)

$$
g_1(J(y)) = \sqrt{\gamma J(y)} \sin(\ln(J(y))), \quad g_2(J(y)) = \sqrt{\gamma J(y)} \cos(\ln(J(y))).
$$

(9)

In addition, we propose the function combining the properties of boundedness and vanishing at the optimal point (another example is given in Reference 16):

$$
g_1(J(y)) = \sqrt{\gamma \tanh \left( \frac{J(y)}{2} \right) \sin(2 \ln(e^{J(y)}) - 1) - J(y)},
$$

$$
g_2(J(y)) = \sqrt{\gamma \tanh \left( \frac{J(y)}{2} \right) \cos(2 \ln(e^{J(y)}) - 1) - J(y)}.
$$

(10)

For the numerical simulations we put $k = 10, \mu = 5, l^* = 1, \epsilon = \frac{1}{4}, \eta = 25, \gamma = 10$. The results of numerical simulations illustrate that the controls with the functions $g_1, g_2$ given by (7) and (8) lead to nonvanishing oscillations in a neighborhood of the optimal point, that is, only singular practical asymptotic stability is ensured (see Figure 1).

However, the functions $g_1, g_2$ chosen according to (9) or (10) satisfy the additional property A3.2), so that the asymptotic stability in the sense of Lyapunov holds (see Figure 2)

![FIGURE 1](https://wileyonlinelibrary.com) Time plots of $x_1(t), x_2(t), u(t)$, and $\|\dot{u}(t)\|$ for system (5)-(6) with functions $g_1, g_2$ given by (7) (left), (8) (right). Initial condition is $x_1(0) = 2, x_2(0) = -1, u(0) = 0$ [Colour figure can be viewed at wileyonlinelibrary.com]
4.2 Case 2: Nonlinear damping

Consider now the case of nonlinear damping $\phi(l) = \mu l^3$, so that system (5) takes the form

$$
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= -kx_1 - \mu x_3^3 - u, \\
y &= x_1 - l^*.
\end{align*}
$$

System (11) has the same quasi-steady-state $\mathcal{E}(u) = \left(-\frac{u}{k}, 0\right)$ which, however, is not exponentially stable for constant $u$. Indeed, the matrix of linear approximation for (11) in a neighborhood of $\mathcal{E}(u)$ always has purely imaginary eigenvalues $\pm i\sqrt{k}$.

Using Barbashin-Krasovskii theorem (or LaSalle invariance principle), one can show that, for each fixed $u$, $\mathcal{E}(u)$ is asymptotically stable for (11) and the conditions of Theorem 1 are satisfied. This can be performed, for example, with the Lyapunov function

$$
V(x) = k \left( x_1 + \frac{u}{k} \right)^2 + x_2^2.
$$

Note that although system (11) does not admit a Lyapunov function satisfying A4.2), it is possible to construct a higher order polynomial Lyapunov function. The results of numerical simulations for system (11) with controls (7) and (10) are shown on Figure 3. The values of system’ and control parameters are taken the same as before. It is interesting to note that in this case even controls with vanishing amplitude do not yield asymptotic stability in the sense of Lyapunov.
5 | CONCLUSIONS AND FUTURE WORK

In this article, we have addressed extremum seeking problems for a general class of nonlinear dynamic systems with a well-defined steady-state map. We proposed a broad family of dynamic control laws for extremum seeking problems based on Lie bracket approximation ideas, which generalizes the results in References 16,25 and which guarantees singular practical asymptotic stability of the optimal state-steady. Unlike the existing results in the literature, the control functions in the proposed control law are not required to be twice continuously differentiable at the optimal steady-state, which provides more flexibility for designing control laws. In particular, the relaxation of regularity requirement is crucial for the control functions $g_j, g_{j+n}$ that vanish at the optimal operating point and for establishing asymptotic stability results in the sense of Lyapunov. The proof techniques proposed in this article extend the approach for analyzing static systems as introduced in Reference 16 and do not rely on classical results from singular perturbations theory, but solely on Chen-Fliess series techniques.

The main results of the article are proved under the assumption that the system admits a well-defined steady-state map and a family of Lyapunov functions for all steady-states. Although this assumption is rather common in the extremum seeking literature, it imposes certain restrictions on the class of systems. In our future work, we aim to show that the proposed approach can be also applied to systems, for which the existence of a steady-state map and a family of Lyapunov functions is not known. Other possible research directions include, for example, extremum seeking problems with a time-varying optimal operating point or problems with constraints, as well as stabilization of underactuated control systems with fast and slow dynamics.

ACKNOWLEDGEMENTS

This work was supported in part by the German Research Foundation (projects GR 5293/1-1, EB 425/4-1) and NAS of Ukraine (budget program KPKBK 6541230).

ORCID

Victoria Grushkovskaya  https://orcid.org/0000-0003-0439-6834

REFERENCES

1. Tan Y, Moase W, Manzie C, Nešić D, Mareels IMY. Extremum seeking from 1922 to 2010. Paper presented at: Proceedings of the 29th Chinese Control Conference; 2010:14-26.
2. Krstić M, Wang HH. Stability of extremum seeking feedback for general nonlinear dynamic systems. Automatica. 2000;36(4):595-601.
3. Krstić M, Arijur KB. Real-Time optimization by Extremum Seeking Control. New York: Wiley-Interscience; 2003.
4. Tan Y, Nešić D, Mareels I. On non-local stability properties of extremum seeking control. Automatica. 2006;42(6):889-903.
5. Nešić D. Extremum Seeking Control: Convergence Analysis. European Journal of Control. 2009;15(3-4):331–347.
6. Nešić D, Tan Y, Moase WH, Manzie C. A unifying approach to extremum seeking: adaptive schemes based on estimation of derivatives. Paper presented at: Proceedings of the 49th IEEE Conference on Decision and Control; 2010:4625-4630.
7. Fu L, Özgüner Ü. Extremum seeking with sliding mode gradient estimation and asymptotic regulation for a class of nonlinear systems. Automatica. 2011;47(12):2595-2603.
8. Liu SJ, Krstić M. Stochastic Averaging and Stochastic Extremum Seeking. London: Springer Science & Business Media; 2012.
9. Dürr HB, Zeng C, Ebenbauer C. Saddle Point Seeking for Convex Optimization Problems. IFAC Proceedings Volumes. 2013;46(23):540–545.
10. Haring M, Van De Wouw N, Nešić D. Extremum-seeking control for nonlinear systems with periodic steady-state outputs. Automatica. 2013;49(6):1883-1891.
11. Guay M, Dochain D. A time-varying extremum-seeking control approach. Automatica. 2015;51:356-363.
12. KhongSZ, Tan Y, Manzie C, Nešić D. Extremum seeking of dynamical systems via gradient descent and stochastic approximation methods. Automatica. 2015;56:44-52.
13. Poveda JI, Quijano N. Shahshahani gradient-like extremum seeking. Automatica. 2015;58:51-59.
14. Benosman M. Learning-Based Adaptive Control: An Extremum Seeking Approach—Theory and Applications. Oxford: Butterworth-Heinemann; 2016.
15. Oliveira TR, Krstić M, Tsubakino D. Extremum seeking for static maps with delays. IEEE Trans Autom Control. 2017;62(4):1911-1926.
16. Grushkovskaya V, Zuyev A, Ebenbauer C. On a class of generating vector fields for the extremum seeking problem: Lie bracket approximation and stability properties. Automatica. 2018;94:151-160.
17. Guay M, Atta KT. Dual mode extremum-seeking control via Lie-bracket averaging approximations. Paper presented at: Proceedings of the 2018 Annual American Control Conference; 2018:2972-2977.
18. Rušiti D, Oliveira TR, Mills G, Krstić M. Deterministic and stochastic Newton-based extremum seeking for higher derivatives of unknown maps with delays. Eur J Control. 2018;41:72-83.
19. Scheinker A, Scheinker D. Constrained extremum seeking stabilization of systems not affine in control. *Int J Robust Nonlinear Control*. 2018;28(2):568-581.

20. Wildhagen S, Michalowsky S, Feiling J, Ebenbauer C. Characterizing the learning dynamics in extremum seeking: the role of gradient averaging and non-convexity. Paper presented at: Proceedings of the 2018 IEEE Conference on Decision and Control; 2018:21-26.

21. Dürr HB, Stanković MS, Ebenbauer C, Johansson K. Lie bracket approximation of extremum seeking systems. *Automatica*. 2013;49:1538-1552.

22. Scheinker A, Krstić M. Extremum seeking with bounded update rates. *Syst Control Lett*. 2014;63:25-31.

23. Scheinker A, Krstić M. Non-C2 Lie bracket averaging for nonsmooth extremum seekers. *J Dyn Syst Meas Control*. 2014;136(1):011010-1-011010-10.

24. Grushkovskaya V, Ebenbauer C. Multi-agent coordination with Lagrangian measurements. *IFAC-PapersOnLine*. 2016;49(22):115-120.

25. Dürr HB, Krstić M, Scheinker A, Ebenbauer C. Extremum seeking for dynamic maps using Lie brackets and singular perturbations. *Automatica*. 2013;49:83-99.

26. Ebenbauer C, Michalowsky S, Grushkovskaya V, Gharesifard B. Distributed optimization over directed graphs with the help of Lie brackets. *IFAC-PapersOnLine*. 2017;50(1):15343-15348.

27. Grushkovskaya V, Dürr HB, Ebenbauer C, Zuyev A. Extremum seeking for time-varying functions using Lie bracket approximations. *IFAC-PapersOnLine*. 2017;50(1):5522-5528.

28. Scheinker A, Krstić M. *Model-free Stabilization by Extremum Seeking*. New York, NY: Springer; 2017.

29. Suttner R, Dashkovskiy S. Exponential stability for extremum seeking control systems. *IFAC-PapersOnLine*. 2017;50(1):15464-15470.

30. Grushkovskaya V, Michalowsky S, Zuyev A, May M, Ebenbauer C. A family of extremum seeking laws for a unicycle model with a moving target: theoretical and experimental studies. Paper presented at: Proceedings of the 2018 European Control Conference; 2018:912-917.

31. Labar C, Garone E, Kinnaert M, Ebenbauer C. Newton-based extremum seeking: a second-order Lie bracket approximation approach. *Automatica*. 2019;105:356-367.

32. Mandic F, Miskovic N, Loncar I. Underwater Acoustic Source Seeking Using Time-Difference-of-Arrival Measurements. *IEEE Journal of Oceanic Engineering*. 2019;1-13.

33. Ghaessifard B, Grushkovskaya V, Ebenbauer C. Multi-agent coordination with Lagrangian measurements. *IFAC-PapersOnLine*. 2016;49(22):115-120.

34. Ghaessifard B, Grushkovskaya V, Ebenbauer C. Distributed optimization over directed graphs with the help of Lie brackets. *IFAC-PapersOnLine*. 2017;50(1):15343-15348.

35. Labar C, Garone E, Kinnaert M, Ebenbauer C. Newton-based extremum seeking: a second-order Lie bracket approximation approach. *Automatica*. 2019;105:356-367.

36. Grushkovskaya V, Zuyev A. Optimal stabilization problem with minimax cost in a critical case of q pairs of purely imaginary eigenvalues. *Nonlinear Anal Theory Methods Appl*. 2013;80:156-178.

37. Grushkovskaya V, Zuyev A. Extremum seeking approach for nonholonomic systems with multiple time scale dynamics. Paper presented at: Proceedings of the 21st IFAC World Congress; 2020; arXiv preprint: arXiv:2005.11370.

38. Khalil HK, Kokotovic PV. *Nonlinear Systems*. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 1996.

39. Khalil HK, Kokotovic PV. On stability properties of nonlinear systems with slowly varying inputs. *IEEE Trans Autom Control*. 1991;36(2):229.
APPENDIX A. AUXILIARY RESULTS

Lemma 2. Let $D \subseteq \mathbb{R}^n$, $\xi(t) \in D$, $0 \leq t \leq r$, be a solution of system $\dot{\xi} = \sum_{i=1}^{l} h_i(\xi)w_i(t)$, and let the functions $h_i$ be Lipschitz continuous in $D$ with Lipschitz constant $L$. Then

$$||\dot{\xi}(t) - \xi(0)|| \leq t\max_{i \geq 1} ||h_i(\xi(0))||e^{Lt}, \quad t \in [0, r].$$  \hspace{1cm} (A1)

with $v = \max_{i \in [0, r]} \sum_{i=1}^{l} |w_i(t)|$.

The proof is based on the Grönwall-Bellman inequality. Another version of this lemma can be found, for example, in References 16,44,45.

Lemma 3. Let the vector fields $h_i$ be Lipschitz continuous in a domain $D \subseteq \mathbb{R}^n$, and $h_i \in C^2(D \setminus \Xi; \mathbb{R})$, where $\Xi = \{\xi \in D : h_i(\xi) = 0 \text{ for all } 1 \leq i \leq l\}$. Assume, moreover, that $L_{h_i}h_i, L_{h_i}L_{h_j}h_i \in C(D; \mathbb{R}^n)$, for all $i,j,l = 1, l$. If $\xi(t) \in D$, $t \in [0, r]$, is a solution of system $\dot{\xi} = \sum_{i=1}^{l} h_i(\xi)w_i(t)$ with $w \in C([0, r]; \mathbb{R}^m)$ and $x(0) = x_0 \in D$, then $\xi(t)$ can be represented by the Chen-Fliess series:

$$\dot{\xi}(t) = \dot{\xi}^0 + \sum_{i=1}^{l} h_i(\xi^0) \int_{0}^{t} w_i(v)dv + \sum_{i,j=1}^{l} L_{h_j}h_i(\xi^0) \int_{0}^{t} \int_{0}^{v} w_i(v)w_j(s)dsdv + R(t), \quad t \in [0, r].$$  \hspace{1cm} (A2)

where

$$R(t) = \sum_{i,j=1}^{l} \int_{0}^{t} \int_{0}^{s} L_{h_j}h_i(\xi(p))w_i(v)w_j(s)w_i(p)dpdsdv$$

is the remainder of the Chen-Fliess series expansion.

Lemma 4. Let $\Gamma \subseteq \mathbb{R}^m$ be compact, and the function $f \in C^1(\mathbb{R}^n \times \Gamma; \mathbb{R}^n)$ be such that, for every $u \in \Gamma$, the equation $f(x, u) = 0$ has a twice continuously differentiable isolated root $x = \ell(u)$, that is, $f(\ell(u), u) = 0$. Assume also that the equilibrium point $\xi = 0$ of the system $\dot{z} = f(z + \ell(u), u)$ is uniformly asymptotically stable in $B_r(0)$ $(r > 0)$, uniformly in the parameter $u$. Then there exists a Lyapunov function $W(z, u)$ such that

$$\alpha_1(||z||) \leq W(z, u) \leq \alpha_2(||z||),$$

$$\frac{\partial W(z, u)}{\partial z} f(z + \ell(u), u) \leq -\alpha_3(||z||),$$

$$\left|\left|\frac{\partial W(z, u)}{\partial u}\right|\right| \leq c_1, \quad \left|\left|\frac{\partial W(z, u)}{\partial z}\right|\right| \leq c_2,$$

for all $z \in B_r(0)$ and $u \in \Gamma$, where $\alpha_1, \alpha_2, \alpha_3, \in \mathbb{K}$, $c_1, c_2 \geq 0$.

APPENDIX B. PROOF OF THEOREM 1

Step 1. This step is aimed at ensuring a well-definiteness of solutions of system (1)-(2).

More precisely, for given $\delta_u \in (0, a_{12}^{-1}(a_{11}(\Delta_u)))$, $\delta_\chi \in (0, \Delta_\chi)$, and $\eta_0$ satisfying A5.1, we will choose a small enough $\varepsilon$ such that, for any $\eta > \eta_0(\varepsilon)$ and any $u^0 \in B_{\delta_\chi}(\ell(u^0))$, $x^0 \in B_{\delta_\chi}(\ell(u^0))$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial condition $x(0) = x^0, u(0) = u^0$ belong to a some compact subset of $D = D_x \times D_u$ as $t \in [0, \eta_1]$.

Recall that system (1)-(2) evolves in the space $D_x \times D_u$, where $D_u = B_{\Delta_\chi}(u^*)$, $D_x = B_{\Delta_\chi}(\ell(u^*))$ with $u^*$ being the unique minimizer of the function $J(u) = J(h(\ell(u), u))$ in $D_u$, and $\Delta_\chi, \Delta_u \in (0, +\infty]$. Let us underline that the assumption $D_u = B_{\Delta_\chi}(u^*)$, $D_x = B_{\Delta_\chi}(\ell(u^*))$ is made only to simplify the technical details in the proofs, and our main results can be proved without that. Without loss of generality, assume $t_0 = 0, \tilde{J}(u^*) = 0$.

We fix a $\delta_u \in (0, a_{12}^{-1}(a_{11}(\Delta_u)))$, $\delta_\chi \in (0, \Delta_\chi)$, and take $\nu, \eta_0$ from Assumption A5.1.
Next, we take \( \delta' \in (a_{11}^{-1}(a_{12}(\delta_u)), \Delta_u) \), \( \delta_u \in (\delta', \Delta_u) \) and denote \( \hat{D}_u = B_{\delta_u}(u^*) \) (in case \( \Delta_u < \infty \) we take \( \delta_u = \Delta_u \) and \( \hat{D}_u = D_u \)). Obviously,

\[
B_{\delta_u}(u^*) \subset B_{\delta}(u^*) \subset \hat{D}_u \subset D_u.
\]

To introduce a compact set \( \hat{D}_x \subset D_x \), denote \( \delta_{ux} = \delta_x + L_{\epsilon} \delta_u \) with \( L_{\epsilon} \) such that \( \| \ell(u) - \ell(v) \| \leq L_{\epsilon} \| u - v \| \), and observe that

\[
x^0 \in B_{\delta}(\ell(u^*)) \Rightarrow \| x^0 - \ell(u^*) \| \leq \delta_{ux}.
\]

By Assumption A5.1, as soon as \( u(t) \in \hat{D}_u \) and \( \| x^0 - \ell(u^*) \| \leq \delta_{ux} \), we have \( x(t) \in \hat{D}_x \) and, moreover,

\[
\| x(t) - \ell(u^*) \| \leq \| x(t) - \ell(u(t)) \| + \| \ell(u(t)) - \ell(u^*) \| \leq \nu_{\epsilon} := \nu_{\epsilon} + L_{\epsilon} \delta_u.
\]

where \( \nu_{\epsilon} \) is defined from A5.1). Thus, in this case \( x(t) \in \hat{D}_x = B_{\delta_x}(\ell(u^*)) \), where \( \delta_x = \min \{ \nu_{\epsilon}, \Delta_x \} \). Obviously, \( \hat{D}_x \) is compact.

To ensure that the solutions \( (x(t), u(t)) \) of system (1)-(2) with initial condition \( x(0) = x^0, u(0) = u^0 \) belong to the compact set \( \hat{D} = \hat{D}_x \times \hat{D}_u \subset D \) as \( t \in [0, \eta \epsilon] \), we begin with constructing an a priori bound on \( \| u(t) - u^0 \| \). For this purpose, we exploit the integral representation of the \( u \)-component of the solutions of system (1)-(2) with the initial condition \( u(0) = u^0 \in B_{\delta_x}(x^0) \):

\[
\| u(t) - u^0 \| = \left\| \sum_{j=1}^{2n_u} \int_0^t g_j(\hat{J}(x(s), u(s)))u^t_s x(s) e_j ds \right\|
\leq \left( t \| g(\hat{J}(u^0)) \| + \int_0^t \| g(\hat{J}(x(s), u(s))) - g(\hat{J}(u^0)) \| ds + \int_0^t \| g(\hat{J}(u(s))) - g(\hat{J}(u^0)) \| ds \right) \max_{0 \leq \epsilon \leq \eta} \sum_{j=1}^{2n_u} |v^t_s\eta(t)|.
\]

Here and in the sequel, \( \hat{J}(x, u) = J(h(x, u)) \). It is easy to see that, for any \( \tau > 0 \),

\[
\max_{0 \leq \epsilon \leq \eta} \sum_{j=1}^{2n_u} |v^t_s\eta(t)| \leq \frac{c_w}{\eta \sqrt{\epsilon}}
\]

with \( c_w = 2 \sum_{j=1}^{n_u} \sqrt{2 \pi \kappa_j} \). Assumption A3.1) implies that there exist constants \( M_g, L_g, L_j > 0 \) such that, for all \( u, v \in \hat{D}_u \),

\[
\| g(\hat{J}(u)) \| \leq M_g,
\| g(\hat{J}(x, u)) - g(\hat{J}(u)) \| \leq L_g(\| x - \ell(u) \|),
\| g(\hat{J}(x)) - g(\hat{J}(v)) \| \leq L_j(\| u - v \|).
\]

Then

\[
\| u(t) - u^0 \| \leq \frac{c_w}{\eta \sqrt{\epsilon}} \left( M_g t + L_g \int_0^t \| x(s) - \ell(u(s)) \| ds + L_j \int_0^t \| u(s) - u^0 \| ds \right).
\]

Using Grönwall-Bellman inequality together with A5.1), we obtain the estimate

\[
\| u(t) - u^0 \| \leq c_w \sqrt{\epsilon}(M_g + L_g \nu_{\epsilon}) e^{L_j \epsilon \sqrt{\epsilon}} \quad \text{for all} \quad t \in [0, \eta \epsilon].
\]

Let \( \epsilon_0 \) be such that

\[
c_w \sqrt{\epsilon_0}(M_g + L_g \nu_{\epsilon}) e^{L_j \epsilon_0} = d_u.
\]
Then for any $\varepsilon \in (0, \varepsilon_0]$, $\eta \in [\eta_0(\varepsilon), +\infty)$, and $u^0 \in B_{\delta_u}(u^\ast)$, $x^0 \in B_{\delta_x}(c'(u^\ast))$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial condition $x(0) = x^0$, $u(0) = u^0$ satisfy the estimate

$$\|u(t) - u^0\| \leq d_u \text{ for all } t \in [0, \varepsilon],$$  \hfill (B2)

and

$$\|u(t) - u^\ast\| \leq \|u^0 - u^\ast\| + d_u \text{ for all } t \in [0, \varepsilon].$$  \hfill (B3)

Next, we define the set

$$L_c = \{ u \in D_u : \hat{J}(u) \leq c \},$$

and take $c_J = a_{11}(\delta_u')$. It is easy to show that the following set inclusions hold:

$$B_{\delta_x}(c'(u^\ast)) \subseteq L_{c_J} \subseteq B_{\delta_x}(u^\ast).$$

Let us put $d_u = \min\{d_1, d_2, d_3\}$, where $d_1, d_2 > 0$ will be define in the next paragraph, and $d_3$ will be defined in Step 5.

Taking $d_1 = a_{12}^{-1}(a_{11}(\delta_u')) - \delta_u > 0$ by the definition of $\delta_u'$ and $d_2 = \delta_u - \delta_u' > 0$, we get the following two properties:

- If $u(0) \in B_{\delta_x}(u^\ast)$, then $u(t) \in L_{c_J}$ for all $t \in [0, \varepsilon]$.

Indeed, estimate (B3) implies

$$\hat{J}(u(t)) \leq a_{11}(\|u^0 - u^\ast\| + d_u) \leq a_{11}(\delta_u + d_1) = a_{12}(a_{12}^{-1}(a_{11}(\delta_u'))) = a_{11}(\delta_u') = c_J \text{ for all } t \in [0, \varepsilon].$$

P2) If $u(0) \in L_{c_J}$, then $u(t) \in D_u$ for all $t \in [0, \varepsilon]$.

In this case, estimate (B3) yields

$$\|u(t) - u^\ast\| \leq \|u^0 - u^\ast\| + d_u \leq a_{11}^{-1}(\hat{J}(u^0)) + d_2 \leq a_{11}^{-1}(c_J) + \delta_u - \delta_u' = \delta_u \text{ for all } t \in [0, \varepsilon].$$

Together with (B1), this implies that for any $\varepsilon \in (0, \varepsilon_0]$, $\eta \in [\eta_0(\varepsilon), +\infty)$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial conditions $u^0 \in L_{c_J} \supset B_{\delta_x}(c'(u^\ast))$, $x^0 \in B_{\delta_x}(c'(u^\ast))$ are well-defined in $D_x \times D_u$ for all $t \in [0, \varepsilon]$.

**Step 2.** The goal of this step is to analyze the behavior of a Lyapunov-like function along the trajectories of system (1)-(2), to ensure that after a time $\eta \varepsilon$, the maximal deviation between the trajectories $x(t)$ and $c'(u(t))$ becomes less than $\varepsilon$.

Namely, we will choose a small enough $\varepsilon_1 > 0$ and a large enough $\eta_1(\varepsilon)$ in such a way that, for any $\varepsilon \in (0, \varepsilon_1)$, $\eta > \eta_1(\varepsilon)$, and any $u^0 \in L_{c_J}$, $x^0 \in B_{\delta_x}(c'(u^\ast))$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial condition $x(0) = x^0$, $u(0) = u^0$ satisfy the estimate $\|x(t) - c'(u(t))\| \leq \varepsilon$ for all $t \in [\eta \varepsilon, 2\eta \varepsilon]$.

From Lemma 4 and Assumptions A1.1) and A4.1), there exists a function $V \in C^1(D_x \times D_u; \mathbb{R}^+)$, functions $\delta_{11}, \delta_{12}, \delta_{21} \in \mathcal{K}$ and a positive constant $\delta_{22}$ such that, for all $x \in D_x$, $u \in D_u$,

$$\delta_{11}(\|x - l(u)\|) \leq V(x, u) \leq \delta_{12}(\|x - l(u)\|),$$

$$\frac{\partial V(x, u)}{\partial x} f(x, u) \leq -\delta_{21}(\|x - l(u)\|),$$

$$\left\| \frac{\partial V(x, u)}{\partial u} \right\| \leq \delta_{22}.$$  \hfill (2.1)

Since $(x(t), u(t)) \in D_x \times D_u$ for all $t \in [0, \eta \varepsilon]$, we have that

$$\|\dot{u}(t)\| \leq \frac{c_0 M_8}{\eta \sqrt{\varepsilon}} \text{ for all } t \in [0, \eta \varepsilon].$$
Then estimating the time derivative of the function $V$ along the trajectories of system (1)-(2), we get

$$V(x(t), u(t)) \leq -\tilde{\sigma}_{21}||x(t) - \ell'(u(t))|| + \frac{c_u M_g}{\eta \sqrt{\varepsilon}} \tilde{\sigma}_{22} \quad \text{for} \quad t \in [0, \eta \varepsilon]. \quad (B4)$$

Observe that, for any $\rho_0 > 0$ and for all $\eta \geq \frac{c_u M_g}{\varepsilon \tilde{\sigma}_{21}(\rho_0)}$,

$$||x(t) - \ell'(u(t))|| > \rho_0 \Rightarrow V(x(t), u(t)) < -\tilde{\sigma}_{21}(\rho_0) + \frac{c_u M_g}{\eta \sqrt{\varepsilon}} \tilde{\sigma}_{22} < 0.$$  

With such a choice of $\eta$,

$$\{ (x, u) \in \tilde{D}_x \times \tilde{D}_u : V(x, u) \geq 0 \} \subseteq \{ (x, u) \in \tilde{D}_x \times \tilde{D}_u : ||x - \ell'(u)|| \leq \rho_0 \}$$

Indeed, if at some point $(x, u) \in \tilde{D}_x \times \tilde{D}_u$ the function $V(x, u)$ is nonnegative, then estimate (B4) yields

$$0 \leq -\tilde{\sigma}_{21}||x - \ell'(u)|| + \frac{c_u M_g}{\eta \sqrt{\varepsilon}} \tilde{\sigma}_{22} \leq -\tilde{\sigma}_{21}||x - \ell'(u)|| + \tilde{\sigma}_{21}(\rho_0),$$

that is $||x - \ell'(u)|| \leq \rho_0$.

Furthermore, if $||x^0 - \ell'(u^0)|| \leq \rho_0$ and $V(x(t), u(t)) < 0$ for $t \in [0, T_1]$ ($T_1 \leq \eta \varepsilon$), then the function $V(x(t), u(t))$ decreases along the trajectories of system (1)-(2) on $[0, T_1]$ and

$$||x(t) - \ell'(u(t))|| \leq \tilde{\sigma}^{-1}_{11}(V(x(t), u(t))) < \tilde{\sigma}^{-1}_{11}(V(x^0, u^0)) \leq \tilde{\sigma}^{-1}_{11}(|x^0 - \ell'(u^0)||) \leq \tilde{\sigma}^{-1}_{11}(\tilde{\sigma}_{12}(\rho_0)).$$

These two observations imply that

$$||x^0 - \ell'(u^0)|| \leq \rho_0 \Rightarrow ||x(t) - \ell'(u(t))|| \leq \min \{ \rho_0, \tilde{\sigma}^{-1}_{11}(\tilde{\sigma}_{12}(\rho_0)) \} = \tilde{\sigma}^{-1}_{11}(\tilde{\sigma}_{12}(\rho_0)) \quad \text{for all} \quad t \in [0, \eta \varepsilon]. \quad (B5)$$

Consider now the case $||x^0 - \ell'(u^0)|| > \rho_0$ and assume that $||x(t) - \ell'(u(t))|| > \rho_0$ for all $t \in [0, \eta \varepsilon]$. Let us take

$$\eta_1(\varepsilon) > \frac{c_u M_g}{\varepsilon \tilde{\sigma}_{21}(\rho_0)} + \frac{\tilde{\sigma}_{12}(\varepsilon)}{\tilde{\sigma}_{21}(\rho_0)}.$$

Then integrating (B4) we get

$$V(x(t), u(t)) < V(x^0, u^0) - t \left( \tilde{\sigma}_{21}(\rho_0) - \frac{c_u M_g}{\eta \sqrt{\varepsilon}} \tilde{\sigma}_{22}(\varepsilon) \right) \quad \text{for each} \quad t \in [0, \eta \varepsilon].$$

Under the above choice of $\eta$, the right-hand side of the above inequality is negative at $t = \eta \varepsilon$ while the left-hand side is nonnegative for all $t \geq 0$. The obtained contradiction proves that there exists a $T_2 \in (0, \eta \varepsilon)$ such that

$$||x(t) - \ell'(u(t))|| > \rho_0 \quad \text{for} \quad t \in [0, T_2),$$

$$||x(T_2) - \ell'(u(T_2))|| \leq \rho_0.$$  

Together with (B5), this yields $||x(t) - \ell'(u(t))|| \leq \tilde{\sigma}^{-1}_{11}(\tilde{\sigma}_{12}(\rho_0))$ for all $t \in [T_2, \eta \varepsilon]$.

Putting

$$\varepsilon_1 = \min \{ \varepsilon_0, \delta_{xu} \},$$
and taking for any $\varepsilon \in (0, \epsilon_1]$ the values $\rho_0 = \delta_{12}^{-1}(\delta_{11}(\varepsilon))$ and $\eta \in [\overline{\eta} (\varepsilon) = \max\{\eta_0 (\varepsilon), \eta_1 (\varepsilon)\}, +\infty)$, we conclude that

$$||x(t) - \ell(u(t))|| \leq \varepsilon \quad \text{for all} \quad t \in [T_2, \etae].$$

Recall that, by the property P1) obtained in Step 1, $u(\etae) \in \mathcal{L}_{\mathcal{C}_1}$. Then by P2) and the obtained estimate $||x(\etae) - \ell(u(\etae))|| \leq \varepsilon \leq \delta_{au}$, we may conclude that the solutions $(x(t), u(t))$ of system (1)-(2) with initial conditions $u^0 \in \mathcal{B}_{\mathcal{U}}(x^\varepsilon)$, $x^0 \in \mathcal{B}_{\mathcal{U}}(\ell(u^\varepsilon))$ are well-defined in $\tilde{D}_{\mathcal{C}} \times \tilde{D}_{\mathcal{U}}$ for all $t \in [\etae, 2\etae]$.

Then we repeat the previous argumentation of Step 2 for $u(\etae) \in \mathcal{L}_{\mathcal{C}_1}$, $x(\etae) \in \mathcal{B}_{\mathcal{U}}(\ell(u^\etae))$ ($\delta_{au} = \min\{\delta_{au}, \Delta_{1}\}$) and the same choice of $\varepsilon, \eta(\varepsilon)$, and conclude that

$$||x(t) - \ell(u(t))|| \leq \varepsilon \quad \text{for all} \quad t \in [\etae, 2\etae].$$

Note that the latter estimate holds and $x(t) \in \tilde{D}_{\mathcal{C}}$ for any time-interval $[\etae, T]$ ($T > \etae$) provided that $u(t) \in \tilde{D}_{\mathcal{U}}$ for $t \in [\etae, T]$.

To sum up Steps 1 and 2, we have ensured that the solutions of system (1)-(2) with initial conditions $u^0 \in \mathcal{B}_{\mathcal{U}}(x^\varepsilon)$, $x^0 \in \mathcal{B}_{\mathcal{U}}(\ell(u^\varepsilon))$ are well-defined in $\tilde{D}_{\mathcal{C}} \times \tilde{D}_{\mathcal{U}}$ for all $t \in [0, 2\etae]$. Furthermore, after a time interval of length $\etae$, the trajectories $x(t)$ and $\ell(u(t))$ became $\varepsilon$-close. In the next steps of the proof we switch to the analysis of the behavior of solutions on the time interval $[\etae, 2\etae]$.

**Step 3. In this step we consider an auxiliary system corresponding to (2) in the case $h(x,u) = h(\ell(u), u)$, and estimate possible deviations of the trajectories of such system from the $u$-component of the trajectories of system (1)-(2) on the time interval $t \in [\etae, 2\etae]$.**

Namely, we will show that, if the initial condition of such auxiliary system at $t_0 = \etae$ is equal to $u(\etae) \in \mathcal{L}_{\mathcal{C}_1}$, then for any $\varepsilon \in (0, \epsilon_1)$, $\eta > \eta_1(\varepsilon)$, and any $u(\etae) \in \mathcal{L}_{\mathcal{C}_1}$, $x(\etae) \in \mathcal{B}_{\mathcal{U}}(\ell(u^\etae))$, norm of the difference between $u(t)$ and the corresponding trajectory of the auxiliary system is of order $\varepsilon \sqrt{\varepsilon}$ for $t \in [\etae, 2\etae]$.

Consider the system

$$\sum_{j=1}^{2n\varepsilon} \mathcal{G}_j(J(\overline{u}))v^\varepsilon_j(t)\xi_j, \quad t \in [\etae, 2\etae],$$

with the initial condition $\overline{u}(\etae) = u(\etae) \in \mathcal{L}_{\mathcal{C}_1}$. Note that as soon as $\overline{u}(\etae) \in \mathcal{L}_{\mathcal{C}_1}$, the solutions of system (B6) are well-defined in $\overline{D}_{\mathcal{U}}$ for all $t \in [\etae, 2\etae]$, for any $\varepsilon \in (0, \epsilon_0)$ and $\eta > \eta_0(\varepsilon)$, where $\epsilon_0$ is defined in Step 1 and $\eta_0(\varepsilon)$ is defined by Assumption A5.1).

Using the integral representations for the $u$-component of the solutions of system (1)-(2) and for the solutions of system (B6) with initial conditions $\overline{u}(\etae) = u(\etae)$, we may write the estimate

$$||u(t) - \overline{u}(t)|| \leq \frac{c_w}{\eta \sqrt{\varepsilon}} \int_{\etae}^{t} (L_{2}||u(s) - \overline{u}(s)|| + L_{3}||x(s) - \ell(u(s))||)ds, \quad \text{for} \quad t \in [\etae, 2\etae].$$

The Grönwall-Bellman inequality implies

$$||u(t) - \overline{u}(t)|| \leq \frac{c_wL_{2}}{\eta \sqrt{\varepsilon}} e^{c_wL_{2} \frac{\sqrt{\varepsilon}}{\eta \sqrt{\varepsilon}}} \int_{\etae}^{t} ||x(s) - \ell(u(s))|| ds \quad \text{for} \quad t \in [\etae, 2\etae].$$

As proved in Step 2, $||x(t) - \ell(u(t))|| \leq \varepsilon$ for $t \in [\etae, 2\etae]$. Hence,

$$||u(t) - \overline{u}(t)|| \leq \frac{c_wL_{2}}{\eta \sqrt{\varepsilon}} e^{c_wL_{2} \frac{\sqrt{\varepsilon}}{\eta \sqrt{\varepsilon}}} \sqrt{\varepsilon} = \theta \sqrt{\varepsilon} \quad \text{for} \quad t \in [\etae, 2\etae],$$

where $\theta = c_wL_{2} e^{c_wL_{2} \sqrt{\varepsilon}}$.

Let us also observe that, as it follows from Lemma 3 and Assumption A3.1), the solutions of system (B6) can be represented as

$$\overline{u}(2\etae) = u(\etae) - \varepsilon \gamma \nabla J(u(\etae)) + R(\varepsilon, \eta),$$

(B8)
where

\[ R(\epsilon, \eta) = \sum_{j_1,j_2,d_1=1}^{2n_b} \int_{\eta \epsilon}^{2\eta \epsilon} \int_{\eta \epsilon}^{2\eta \epsilon} \int_{\eta \epsilon}^{2\eta \epsilon} L_{\epsilon_1,g_1} L_{\epsilon_2,g_2} g_1 \circ \hat{J}(\hat{u}(s_3)) e_{j_1} v_{j_2}^\eta(s_3) v_{j_3}^\eta(s_2) v_{j_4}^\eta(s_1) ds_3 ds_2 ds_1. \]

In (B8), we exploited the fact that \( \hat{u}(\eta \epsilon) = u(\eta \epsilon) \). Since the functions \( L_{\epsilon_1,g_1} L_{\epsilon_2,g_2} g_1 \circ \hat{J} \) are continuous in \( D_u \) From Assumption A3.1, the remainder \( R(\epsilon, \eta) \) of representation (B8) can be estimated as

\[ \|R(\epsilon, \eta)\| \leq \left( \frac{\eta \epsilon}{6} \right)^3 \max_{1 \leq j_1,j_2,j_3 \leq 2n_b} \sup_{u \in D_u} \|L_{\epsilon_1,g_1} L_{\epsilon_2,g_2} g_1(\hat{J}(u))\| \max_{0 \leq t \leq \eta \epsilon} \left( \sum_{j=1}^{2n_b} \|v_{j}^\eta(t)\| \right)^3 \leq \zeta \epsilon \sqrt{\epsilon}, \quad (B9) \]

where \( \zeta = \frac{\epsilon^3}{6} \max_{1 \leq j_1,j_2,j_3 \leq 2n_b} \sup_{u \in D_u} \|L_{\epsilon_1,g_1} L_{\epsilon_2,g_2} g_1(\hat{J}(u))\| \). Note that the obtained estimate does not depend on \( \eta \) and \( u(\eta \epsilon) \).

**Step 4.** The goal of this step is to ensure the decay of the function \( \hat{J}(u(t)) \) outside a neighborhood of \( u^* \) after the time intervals of length \( \eta \epsilon \). Namely, for any \( \rho_1 \in (0, \delta_u) \), we will find a small enough \( \epsilon_2 > 0 \) and a large enough \( \bar{\eta}(\epsilon) \) in such a way that, for any \( \epsilon \in (0, \epsilon_2) \), \( \eta > \bar{\eta}(\epsilon) \), and for any \( u(\eta \epsilon) \in L_{\epsilon_1} \setminus B_{\rho_1}(u^*) \), \( x(\eta \epsilon) \in B_{\delta_u}(\epsilon(u^*)) \), the \( u(t) \)-component of the solutions of system (1)-(2) satisfies the property \( \hat{J}(u(2\eta \epsilon)) < \hat{J}(u(\eta \epsilon)) \).

Note that the vector fields of subsystem (2) depend on the function \( \hat{J}(x, u) \), so that we cannot obtain the above-mentioned properties from the analysis of the behavior of \( \hat{J}(u) \) along the trajectories of (2), as it was done in Reference 16. Instead, we consider the behavior of the function \( \hat{J}(u) \) along the trajectories of auxiliary system (B6) and exploit representation (B8) and estimates (B7), (B9).

With Taylor's formula, the function \( \hat{J}(u(t)) \) can be represented as

\[ \hat{J}(u(2\epsilon \eta)) = \hat{J}(u(\epsilon \eta)) + (\nabla \hat{J}(u(\epsilon \eta)), u(2\epsilon \eta) - u(\epsilon \eta)) + \frac{1}{2}(u(2\epsilon \eta) - u(\epsilon \eta)) + \frac{\partial J(\xi)}{\partial u^2}(u(2\epsilon \eta) - u(\epsilon \eta)), \]

where \( \frac{\partial J(\xi)}{\partial u^2} \) denotes the Hessian matrix of \( \hat{J} \) calculated at \( \xi = u(\epsilon \eta) + \theta(u(2\epsilon \eta) - u(\epsilon \eta)) \), for some \( \theta \in [0, 1] \). Adding and subtracting \( \bar{u}(2\epsilon \eta) \) into multipliers \( u(2\epsilon \eta) - u(\epsilon \eta) \), we estimate \( \hat{J}(u(2\epsilon \eta)) \) as

\[ \hat{J}(u(2\epsilon \eta)) \leq \hat{J}(u(\epsilon \eta)) + (\nabla \hat{J}(u(\epsilon \eta)), \bar{u}(2\epsilon \eta) - u(\epsilon \eta)) + \frac{1}{2} \left\| \frac{\partial^2 \hat{J}(\xi)}{\partial u^2} \right\| \left\| u(2\epsilon \eta) - u(\epsilon \eta) \right\|^2 \]

\[ + \|\nabla \hat{J}(u(\epsilon \eta))\| \left\| u(2\epsilon \eta) - \bar{u}(2\epsilon \eta) \right\| + \frac{1}{2} \left\| \frac{\partial^2 \hat{J}(\xi)}{\partial u^2} \right\| \left\| u(2\epsilon \eta) - \bar{u}(2\epsilon \eta) \right\|^2. \quad (B10) \]

Substituting (B7), (B8), and (B9) into (B10) and taking into account that \( u(t) \in \bar{D}_u \) for \( t \in [0, 2\epsilon \eta] \), we may estimate \( \hat{J}(u(2\epsilon \eta)) \) as

\[ \hat{J}(u(2\epsilon \eta)) \leq \hat{J}(u(\epsilon \eta)) - \epsilon \gamma \left\| \nabla \hat{J}(u(\epsilon \eta)) \right\|^2 \left( 1 - \frac{\gamma^2 a_3(\delta_u)}{2} \epsilon \right) \]

\[ + \epsilon \sqrt{\epsilon} \left\| \nabla \hat{J}(u(\epsilon \eta)) \right\|(1 + \epsilon \gamma a_3(\delta_u))(\xi + \theta) + \frac{\epsilon^3 a_3(\delta_u)}{2 \gamma^2}(\xi + \theta)^2. \]

Let us take an arbitrary small \( \rho_1 \in (0, \delta_u) \), and assume that \( \left\| u(\epsilon \eta) - u^* \right\| > \rho_1 \) (the case \( u(\epsilon \eta) \in \bar{B}_{\rho_1}(u^*) \) will be covered in Step 5). Then \( \left\| \nabla \hat{J}(u(\epsilon \eta)) \right\| \geq a_{21}(\rho_1) \), and

\[ \hat{J}(u(2\epsilon \eta)) \leq \hat{J}(u(\epsilon \eta)) - \epsilon \left\| \nabla \hat{J}(u(\epsilon \eta)) \right\|^2 \left( \gamma - \epsilon \left( \frac{\gamma^2 a_3(\delta_u)}{2} + \frac{\sqrt{\epsilon}}{a_{21}(\rho_1)}(1 + \epsilon \gamma a_3(\delta_u))(\xi + \theta) + \frac{\epsilon^2 a_3(\delta_u)}{2 a_{21}^2(\rho_1)}(\xi + \theta)^2 \right) \right). \]

For any \( \lambda \in (0, \gamma) \), let \( \epsilon_2 \) be the positive root of the equation

\[ \epsilon_2 \left( \frac{\gamma^2 a_3(\delta_u)}{2} + \frac{\sqrt{\epsilon}}{a_{21}(\rho_1)}(1 + \epsilon \gamma a_3(\delta_u))(\xi + \theta) + \frac{\epsilon^2 a_3(\delta_u)}{2 a_{21}^2(\rho_1)}(\xi + \theta)^2 \right) = \gamma - \lambda. \]
Then, for all $\varepsilon \in (0, \min \{\varepsilon_0, \varepsilon_1, \varepsilon_2\}]$, $\eta \in [\overline{\eta}(\varepsilon), +\infty)$,

$$J(u(2\eta)) \leq J(u(\eta)) - \varepsilon \lambda \|\nabla J(\eta)\|^2 < J(u(\eta)). \quad (B11)$$

Let us emphasize that the constant $\lambda$ does not depend on $u(\eta)$.

**Step 5.** On this final step, we show that the solutions of system (1)-(2) enter a prescribed neighborhood of $(\ell(u^*), u^*)$ after a finite time, and then remain there.

Namely, given a $\rho > 0$ we will find an $\overline{\varepsilon} > 0$ such that, for any $\varepsilon \in (0, \overline{\varepsilon})$ and $\eta > \overline{\eta}(\varepsilon)$, there exists an $\mathcal{N} \in \mathbb{N}$ such that, for any $u^0 \in \mathcal{C}_1, x^0 \in B_{\delta_{u^0}}(\ell(u^*))$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial condition $x(0) = x^0$, $u(0) = u^0$ satisfy the property $\|x(t) - \ell(u^*)\| \leq \rho$ and $\|u(t) - u^*\| \leq \frac{\rho}{\delta_{u^0}}$ for all $t \geq \mathcal{N} \eta$.

Since $u(\eta) \in \mathcal{C}_1$, estimate (B11) ensures $u(2\eta) \in \mathcal{C}_1$. Similarly to Steps 1 and 2, we may show that $(x(t), u(t)) \in \hat{D}_x \times \hat{D}_u$ for $t \in [0, 2\eta)$, and $\|x(t) - \ell(u(t))\| \leq \varepsilon$ for all $t \in [\eta, 2\eta]$. Furthermore, we may consider system (B6) for $t \in [2\eta, 3\eta]$ with the initial condition $\overline{u}(2\eta) = u(2\eta) \in \mathcal{C}_1$, and show that $J(u(3\eta)) < J(u(2\eta))$ whenever $\|u(2\eta) - u^*\| > \rho_1$. Iterating this procedure for $u(j\eta) \in \mathcal{C}_1 \setminus B_{\delta_{u^0}}(u^*)$, we conclude that there exists an $\mathcal{N} \in \mathbb{N}$ such that

$$\|u(j\eta) - u^*\| > \rho_1 \text{ for } j = 0, 1, \ldots, \mathcal{N} - 1,$$

$$\|u(\mathcal{N}\eta) - u^*\| \leq \rho_1.$$

Furthermore,

$$J(u(\mathcal{N}\eta)) < \ldots < J(u(2\eta)) < J(u(\eta)),$$

that is, $u(\mathcal{N}\eta) \in \mathcal{C}_1$, and $(x(t), u(t)) \in \hat{D}_x \times \hat{D}_u$ for $t \in [0, \mathcal{N}(\mathcal{N} + 1)\eta]$, $\|x(t) - \ell(u(t))\| \leq \varepsilon$ for all $t \in [\eta, \mathcal{N}(\mathcal{N} + 1)\eta]$. For an arbitrary $t \in [\eta, \mathcal{N}(\mathcal{N} + 1)\varepsilon]$, we denote the integer part of $\frac{t}{\eta}$ as $\left[ \frac{t}{\eta} \right]$ and observe that $0 < t - \left[ \frac{t}{\eta} \right] \eta < \eta$. Then we apply the triangle inequality together with (B2):

$$\|u(t) - u^*\| \leq \left\| u \left( \left[ \frac{t}{\eta} \right] \eta \right) - u^* \right\| + \left\| u(t) - u \left( \left[ \frac{t}{\eta} \right] \eta \right) \right\|$$

$$\leq \alpha_1^{-1} \left( J \left( \left[ \frac{t}{\eta} \right] \eta \right) \right) + d_u \leq \alpha_1^{-1}(a_{12}(\delta_u)) + d_3 \text{ for } t \in [\eta, \mathcal{N}(\mathcal{N} + 1)\eta]. \quad (B12)$$

Recall that $d_3$ in (B12) has been introduced in Step 1 and is assumed to be defined on the current step. For $t \in [\mathcal{N}\eta, \mathcal{N}(\mathcal{N} + 1)\eta]$, we denote that $u(\mathcal{N}\eta) \in B_{\delta_{u^0}}(u^*)$, and

$$\|u(t) - u^*\| \leq \|u(\mathcal{N}\eta) - u^*\| + d_u \leq \rho_1 + d_3 \text{ for } t \in [\mathcal{N}\eta, \mathcal{N}(\mathcal{N} + 1)\eta], \quad (B13)$$

which follows from (B3). Consequently,

$$J(u((\mathcal{N} + 1)\eta)) \leq a_{12}(\rho_1 + d_3).$$

Consider two possibilities:

i1) If $\|u((\mathcal{N} + 1)\eta) - u^*\| > \rho_1$, we repeat the step and conclude that

$$J(u((\mathcal{N} + 2)\eta)) < J(u((\mathcal{N} + 1)\eta)) \leq a_{12}(\rho_1 + d_3).$$

$(x(t), u(t)) \in \hat{D}_x \times \hat{D}_u$ for $t \in [0, (\mathcal{N} + 2)\eta]$, and $\|x(t) - \ell(u(t))\| \leq \varepsilon$ for all $t \in [\eta, (\mathcal{N} + 2)\eta]$. Arguing similarly to (B12), we also get that, for $t \in [(\mathcal{N} + 1)\eta, (\mathcal{N} + 2)\eta]$,

$$\|u(t) - u^*\| \leq \alpha_1^{-1} \left( J \left( \left[ \frac{t}{\eta} \right] \eta \right) \right) + d_3 = \alpha_1^{-1}(J((\mathcal{N} + 1)\eta)) + d_3 \leq \alpha_1^{-1}(a_{12}(\rho_1 + d_3)) + d_3.$$

i2) If $\|u((\mathcal{N} + 1)\eta) - u^*\| \leq \rho_1$, the property (B13) yields $u(t) \in B_{\rho_1 + d_3}(u^*)$ for all $t \in [(\mathcal{N} + 1)\eta, (\mathcal{N} + 2)\eta]$. The other assertions of Step 5 are shown in a similar way.
Repeating i1, i2, we prove that $(x(t), u(t)) \in D_x \times D_u$ and
\[ \|x(t) - \ell'(u(t))\| \leq \varepsilon \text{ for all } t \in [0, \infty). \] (B14)

Furthermore,
\[ \|u(t) - u^*\| \leq \max\{\rho_1, \rho_1 + d_3, \alpha_{11}^{-1}(\alpha_{12}(\rho_1 + d_3)) + d_3\} = \alpha_{11}^{-1}(\alpha_{12}(\rho_1 + d_3)) + d_3 \text{ for all } t \in [\tilde{N} \eta \varepsilon, \infty). \]

Consequently,
\[ \|x(t) - \ell(u(t))\| \leq \|x(t) - \ell'(u(t))\| + \|\ell'(u(t)) - \ell'(u^*)\| \leq \varepsilon + L_{\ell'}(\alpha_{11}^{-1}(\alpha_{12}(\rho_1 + d_3)) + d_3) \text{ for all } t \in [\tilde{N} \eta \varepsilon, \infty). \]

For any $\rho > 0$, let us take, for example, $\rho_1$ and $d_3$ small enough to ensure that $d_3 \leq \frac{\rho}{4L_{\ell'}}$, $\rho_1 + d_3 \leq \alpha_{12}^{-1}\left(\alpha_{11}\left(\frac{\rho}{4L_{\ell'}}\right)\right)$, and
\[ \tilde{\varepsilon} = \min\left\{\varepsilon_0, \varepsilon_1, \varepsilon_2, \frac{\rho}{2}\right\}. \]

Then
\[ \|u(t) - u^*\| \leq \frac{\rho}{2L_{\ell'}} \text{ and } \|x(t) - \ell(u(t))\| \leq \rho \text{ for all } t \in [\tilde{N} \eta \varepsilon, \infty). \] (B15)

**Conclusions.** To sum up, we have proved that under Assumptions A1.1 to A5.1, for any given $\delta_u \in (0, \alpha_{12}^{-1}(\alpha_{11}(\Delta_u)))$ and $\delta_\varepsilon \in (0, \Delta_\varepsilon)$, there exist an $\tilde{\varepsilon} > 0$ and $\nu : (0, +\infty) \rightarrow (0, +\infty)$ such that, for any $\varepsilon \in (0, \tilde{\varepsilon})$, $\nu > \eta(\varepsilon)$, and any $u^0 \in B_{\delta_u}(u^*)$, $x^0 \in B_{\delta_\varepsilon}(\ell(u^*))$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial condition $x(0) = x^0$, $u(0) = u^0$ belong to a some compact subset of $D_x \times D_u \subseteq \mathbb{R}^2$ as $t \in [0, +\infty)$.

In addition, the estimates (B3), (B12), and (B15) imply that
\[ \|u(t) - u^*\| \leq \alpha_{11}^{-1}(\alpha_{12}(\delta_u)) + d_u \text{ for all } t \in [0, +\infty). \] (B16)

Besides, with the use of estimates (B5) and (B14) we obtain
\[ \|x(t) - \ell(u^*)\| \leq \|x(t) - \ell'(u(t))\| + \|\ell'(u(t)) - \ell'(u^*)\| \leq \delta_\varepsilon + L_{\ell'}(\delta_u + \alpha_{11}^{-1}(\alpha_{12}(\delta_u)) + du) \text{ for all } t \in [0, +\infty). \] (B17)

Here, we also exploited the fact that $\varepsilon < \delta_{\varepsilon u} = \delta_\varepsilon + L_{\ell'}\delta_u$ by the construction. Recall that $d_u$ can be made arbitrary small with the appropriate choice of $\varepsilon$, as it was shown in Step 1. Then estimate (B16) and (B17) guarantee that for any $\rho > 0$ there exist $\delta_u > 0$, $\delta_\varepsilon > 0$ and $\tilde{\varepsilon} > 0$ such that, for any $\varepsilon \in (0, \tilde{\varepsilon})$, $\nu > \eta(\varepsilon)$, and any $u^0 \in B_{\delta_u}(u^*)$, $x^0 \in B_{\delta_\varepsilon}(\ell(u^*))$, the solutions $(x(t), u(t))$ of system (1)-(2) with initial condition $x(0) = x^0$, $u(0) = u^0$ satisfy the estimates $\|u(t) - u^*\| \leq \frac{\rho}{2L_{\ell'}}$ and $\|x(t) - \ell(u^*)\| \leq \rho$ for all $t \in [0, +\infty)$. This gives the singular practical uniform stability of $(\ell(u^*), u^*)$ for system (1)-(2). Consequently, its singular practical uniform attractivity directly follows from (B15). All in all, we have proved the singular practical uniform asymptotic stability of $(\ell(u^*), u^*)$ for system (1)-(2).

**APPENDIX C. PROOF OF LEMMA 1**

We keep the notations from the proof of Theorem 1, define the sets $L_{\varepsilon, c}$ and $\tilde{D}_u \subseteq D_u$ in the same way, and put $\delta_{\varepsilon u} \in (0, \sqrt{\sigma_{11}/\sigma_{12} \Delta_\varepsilon}), \tilde{D}_x = B_{\delta_{\varepsilon u}}(\ell(u^*))$.

Assumption A4.2) obviously implies Assumption A4.1) of Theorem 1. Let us show that Assumption A5.1) is satisfied as well.

Using the triangular inequality and Assumption A3.1), we get that
\[ \|\dot{u}(t)\| \leq \frac{c_w}{\eta \sqrt{\varepsilon}} (M_g + L_g \|x(t) - \ell(u(t))\| + L_j \|u(t) - u(0)\|), \]
Thus, there exist an $\epsilon_0 > 0$ such that $\|x(t) - \ell'(u^*)\| < \sqrt{\frac{\sigma_{12}}{\sigma_{11}}} \Delta_x$ for $t \in [0, \eta \epsilon]$ provided that $\epsilon < \epsilon_0$. This observation ensures Assumption A5.1 of Theorem 1 and completes the proof.
APPENDIX D. PROOF OF THEOREM 2

From Theorem 1 and Lemma 1, there exist $\tilde{\varepsilon}_0 > 0$, $\tilde{\eta}_0(\varepsilon) > 0$ such that for any $\varepsilon > 0$, $\eta > \tilde{\eta}_0(\varepsilon)$, and $x^0 \in D_x$, $u^0 \in B_{\tilde{\varepsilon}_0}(u^*)$ such that $\|x^0 - \ell'(u^0)\| \leq \delta_{u^0}$, the $u$-component of the solutions of system (1)-(2) with initial conditions $u(0) = u^0$, $x(0) = x^0$ are well-defined in $\tilde{D}_u$ for all $t \in [0, +\infty)$. Below we will prove that the trajectories of system (1)-(2) converge asymptotically to $(u^*, \ell'(u^*))$ with some fixed $\varepsilon, \eta$.

Taking into account Assumptions A2.2) to A3.2), we may construct more tiny a priori estimates in the following way. Using the triangular inequality and Assumption A3.2), we get that

$$||\dot{u}(t)|| \leq \frac{c_w}{\sqrt{\varepsilon}} (M_g ||u(0) - u^*|| + L_g ||x(t) - \ell'(u(t))|| + L_f ||u(t) - u(0)||),$$

where $M_g = M_g \sqrt{\alpha_{ij}}$. This implies that estimates (C1) and (C2) can be written as

$$||u(t) - u(0)|| \leq c_w e^{L_f \sqrt{\varepsilon}} \left( \sqrt{\varepsilon M_g} ||u(0) - u^*|| + L_g \int_0^t ||x(s) - \ell'(u)(s)|| ds \right) \quad \text{for} \quad t \in [0, \eta \varepsilon]$$

and

$$||\dot{u}(t)|| \leq \frac{c_w}{\eta \sqrt{\varepsilon}} \left( \zeta_1 ||u(0) - u^*|| + \zeta_2 \sqrt{V(u(t), x(t))} + \frac{c_w \zeta_3}{\eta \sqrt{\varepsilon}} \int_0^t \sqrt{V(u(s), x(s))} ds \right),$$

where $\zeta_1 = M_g (1 + L_f c_w \sqrt{\varepsilon} e^{L_f \sqrt{\varepsilon}})$. Note that differently from Reference 16, the solutions of system (1)-(2) do not admit the estimate of form $||u(t) - u(0)|| \sim O(\|u(0) - u^*\|)$ as $u(0) \to u^*$.

If $\eta_1(\varepsilon)$ is defined as in (C3), then for any $\varepsilon \in (0, \tilde{\varepsilon}_0], \eta \in [\max \{\tilde{\eta}_0(\varepsilon), \eta_1(\varepsilon)\}, +\infty)$ estimate (C4) takes the form

$$\sqrt{V(x(t), u(t))} \leq \left( \sqrt{V(x(0), u(0))} e^{-\frac{\eta_1}{2} t} + \varepsilon_{\tilde{\varepsilon}} ||u(0) - u^*|| \right) e^{\varepsilon t} \quad \text{for} \quad t \in [0, \eta \varepsilon],$$

with $\varepsilon_{\tilde{\varepsilon}} = c_w \bar{\zeta}_1 \sigma_{22}$ and $\zeta_5$ defined in the proof of Lemma 1.

As in the proof of Theorem 1, we expand the solutions of system (B6) with $\bar{u}(0) = u(0)$ into the Chen-Fliess series. However, we do not assume that $||u(t) - u^*|| > 0$ for $t \in [0, \eta \varepsilon]$ while exploiting instead Lemma 3 together with Assumptions A2.2) to A4.2).

Again, we may write the representation

$$\bar{u}(\eta \varepsilon) = u(0) - \varepsilon \gamma \sqrt{\hat{J}(u(0))} + R(\varepsilon, \eta).$$

With Lemma 2, the remainder of the Chen-Fliess series expansion can be estimated as

$$\|R(\varepsilon, \eta)\| \leq \tilde{M}_{3g} c_w \int_0^{\eta \varepsilon} \int_0^{\tilde{x}_1} \int_0^{\tilde{x}_2} ||u(s) - u^*|| ds_2 ds_1 ds \leq \tilde{\xi} \varepsilon \sqrt{\varepsilon \hat{J}(u(0))},$$

where $\tilde{M}_{3g} = M_{3g} \sqrt{\alpha_{ij}}$, $\tilde{\xi} = \tilde{M}_{3g} c_w \sqrt{\varepsilon} e^{L_f \sqrt{\varepsilon}}$. Furthermore, using (B7) and (D3), we conclude that for any $\varepsilon \in (0, \tilde{\varepsilon}_0], \eta \in [\max \{\tilde{\eta}_0(\varepsilon), \eta_1(\varepsilon)\}, +\infty)$,

$$||\bar{u}(\eta \varepsilon) - u(\eta \varepsilon)|| \leq \varepsilon \theta_1 \sqrt{V(x(0), u(0))} + \varepsilon \sqrt{\varepsilon \hat{J}(u(0))},$$

where $\theta_1 = 2c_w L_f \sqrt{\sigma_{22}} e^{(c_w L_f + \bar{\zeta}_1) \sqrt{\varepsilon}}$, $\theta_2 = c_w \sigma_{22} \sqrt{\tau_{12}} e^{(c_w L_f + \bar{\zeta}_1) \sqrt{\varepsilon}}$. Substituting the above estimates into (B10), we obtain
\[
\hat{J}(u(\eta e)) \leq \hat{J}(u(0)) - \varepsilon \hat{J}(u(0))(\gamma \bar{a}_{21} - \sqrt{\varepsilon} \sqrt{\bar{a}_{22}(\zeta + \theta_2) - \alpha_3 \varepsilon (\gamma^2 \bar{a}_{22} + \varepsilon (\zeta^2 + \theta_2^2))} \\
+ \varepsilon \theta_1 \sqrt{\bar{a}_{22}} \hat{J}(u(0))V(x(0), u(0)) + \varepsilon^2 \bar{a}_3 \theta_1^2 V(x(0), u(0)).
\]

For any \( \lambda \in (0, \gamma \bar{a}_{21}) \), let \( \bar{e}_1 \) be the smallest positive solution of the equation
\[
\sqrt{\varepsilon} \sqrt{\bar{a}_{22}(\zeta + \theta_2) + \alpha_3 \varepsilon (\gamma^2 \bar{a}_{22} + \varepsilon (\zeta^2 + \theta_2^2))} + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} = \gamma \bar{a}_{21} - \lambda.
\]

Then for any \( \varepsilon \in (0, \bar{e}_1] \)
\[
\hat{J}(u(\eta e)) \leq \hat{J}(u(0)) - \varepsilon \left( \lambda + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} \right) \hat{J}(u(0)) + \varepsilon \theta_1 \sqrt{\bar{a}_{22}} \hat{J}(u(0))V(x(0), u(0)) + \varepsilon^2 \bar{a}_3 \theta_1^2 V(x(0), u(0)).
\]

Unlike the results of paper\(^{16}\) the obtained estimate does not guarantee that the function \( \hat{J}(u) \) is decaying along the trajectories of system (1)-(2) on infinite time interval in the sense \( \hat{J}(u(0)) \geq \hat{J}(u(\eta e)) \geq \hat{J}(u(\eta e)) \geq \ldots \). Instead, we will ensure such property for the function
\[
W(u, x) = \hat{J}(u) + V(u, x).
\]

From (D3) and (D4), we conclude that
\[
W(u(\eta e), x(\eta e)) \leq \hat{J}(u(0)) - \varepsilon \lambda \hat{J}(u(0)) + 2 \varepsilon \sqrt{\hat{J}(u(0))V(x(0), u(0))} \left( \frac{\theta_1 \sqrt{\bar{a}_{22}}}{2} + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} \right) \\
+ V(x(0), u(0)) (\gamma^2 \bar{a}_{22} + \varepsilon (\zeta^2 + \theta_2^2)) \\
= \hat{J}(u(0)) - \varepsilon \lambda \left( \sqrt{\hat{J}(u(0))} - \sqrt{V(x(0), u(0))} \left( \frac{\theta_1 \sqrt{\bar{a}_{22}}}{2} + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} \right) \right)^2 \\
+ V(x(0), u(0)) \left( \frac{\varepsilon \theta_1 \sqrt{\bar{a}_{22}}}{2} + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} \right) + \varepsilon^2 \bar{a}_3 \theta_1^2 + e^{2 \varepsilon \sqrt{\varepsilon}} \\
\leq \hat{J}(u(0)) + V(x(0), u(0)) \left( \frac{\varepsilon \theta_1 \sqrt{\bar{a}_{22}}}{2} + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} \right)^2 + \varepsilon^2 \bar{a}_3 \theta_1^2 + e^{2 \varepsilon \sqrt{\varepsilon}}.
\]

Take \( \eta_2(\varepsilon) \) such that \( e^{2 \varepsilon \sqrt{\varepsilon} - \frac{\alpha_3 \varepsilon}{2} \lambda} = \frac{1}{2} \), that is,
\[
\eta_2(\varepsilon) = \frac{2}{\alpha_2 \varepsilon} (\ln 2 + 2 \varepsilon \sqrt{\varepsilon}),
\]

and let \( \bar{e}_2 \) be the smallest positive root of the equation
\[
\frac{\varepsilon}{4 \lambda} \left( \frac{\theta_1 \sqrt{\bar{a}_{22}}}{2} + \frac{\varepsilon \theta_1^2}{\bar{a}_1} e^{2 \varepsilon \sqrt{\varepsilon}} \right)^2 + \varepsilon^2 \bar{a}_3 \theta_1^2 = \frac{1}{4}.
\]

Then for any \( \varepsilon \in (0, \bar{e} = \min \{ \bar{e}_0, \bar{e}_1, \bar{e}_2 \}] \), \( \eta \in [\bar{\eta} = \max \{ \hat{\eta}_0(\varepsilon), \eta_1(\varepsilon), \eta_2(\varepsilon) \}, +\infty) \),
\[
W(u(\eta \varepsilon), x(\eta e)) \leq \hat{J}(u(0)) + \frac{3}{4} V(x(0), u(0)) \leq \hat{J}(u(0)) + V(x(0), u(0)) = W(x(0), u(0)),
W(u(\eta \varepsilon), x(\eta e)) = W(x(0), u(0)) \Leftrightarrow u(0) = u^* \text{ and } x(0) = \varepsilon(u(0)) = \varepsilon(u^*).
\]

Iterating the above procedure for \( u(\varepsilon) \in L_{C^0}, \ x(\varepsilon) \in D_{k} \) we conclude that \( W(u((j+1)\varepsilon), x((j+1)\varepsilon)) \leq W(u(j\varepsilon), x(j\varepsilon)), j = 0, 1, \ldots \).
Consider the discrete-time dynamical system
\[ q^j = \phi(q^{j-1}), \quad j = 1, 2, \ldots, \tag{D6} \]
where \( q = (x, u) \in D_x \times D_u, \phi : D_x \times D_u \to D_x \times D_u \) maps any \( q^0 = (\xi_1, \xi_2) \in D_x \times D_u \) to the solution of system (1)-(2) with the initial condition \( x|_{t=0} = \xi_1, u|_{t=0} = \xi_2 \) and controls (2) and (3) evaluated at \( t = \eta \epsilon \), and \( \phi(q^0) = q^0 \) if \( \xi_1 = \ell^\epsilon(u^*), \xi_2 = u^* \).

One can see that \( q^j = (x(j \eta \epsilon), u(j \eta \epsilon)), j = 0, 1, 2, \ldots, \) where \((x(t), u(t))\) is the solution of system (1)-(2) with the initial condition \( x|_{t=0} = \xi_1, u|_{t=0} = \xi_2 \) and controls (2) and (3). As it has been already proved, \((x(t), u(t))\) \( \in D_x \times D_u \) for all \( t \geq 0 \). Then the invariance principle\(^{47}\) for (D6) implies that
\[ q^j \to (\ell^\epsilon(u^*), u^*) \quad \text{as} \quad j \to +\infty. \tag{D7} \]

Furthermore, with (D3) and (C3) estimate (D1) takes the form
\[ \|u(t) - u(0)\| \leq \tilde{\theta}_1 \sqrt{\epsilon} \|u(0) - u^*\| + \tilde{\theta}_2 \epsilon \sqrt{\epsilon} \sqrt{V(u(0), z(0))} \quad \text{for} \quad t \in [0, \eta \epsilon], \tag{D8} \]
where \( \tilde{\theta}_1 = c_w e^{\sqrt{\epsilon} (c_u L + \xi_5)} (\tilde{M} + \xi_4 \sqrt{\epsilon}), \quad \tilde{\theta}_2 = \frac{-2 L_c c_w e^{\sqrt{\epsilon} (c_u L + \xi_5)}}{\sqrt{\xi_3}}. \) The latter estimate completes the proof.