Extremum Seeking Approach for Nonholonomic Systems with Multiple Time Scale Dynamics

Victoria Grushkovskaya, Alexander Zuyev

Abstract: In this paper, a class of nonlinear driftless control-affine systems satisfying the bracket generating condition is considered. A gradient-free optimization algorithm is developed for the minimization of a cost function along the trajectories of the controlled system. The algorithm comprises an approximation scheme with fast oscillating controls for the nonholonomic dynamics and a model-free extremum seeking component with respect to the output measurements. Exponential convergence of the trajectories to an arbitrary neighborhood of the optimal point is established under suitable assumptions on time scale parameters of the extended system. The proposed algorithm is tested numerically with the Brockett integrator for different choices of generating functions.

Keywords: nonholonomic systems, extremum seeking, stability of nonlinear systems, output feedback control, Lyapunov methods.

1. INTRODUCTION

Extremum seeking theory aims at designing universal control algorithms which steer the trajectories of dynamical systems with uncertainties to the minimum (or maximum) of a cost function whose analytical representation may be partially or completely unknown. The first results in this direction date back to the twenties of the last century, while the first thorough analysis of the stability properties of extremum seeking systems has been carried out only in the early 2000s, cf. Krstić and Wang (2000). Since then, many new extremum seeking algorithms and their applications have been developed (see, e.g., Krstić and Åström (2000); Tan et al. (2006); Nesić et al. (2010); Fu and Özkutlu (2011); Liu and Krstić (2012); Dühr et al. (2013b); Haring et al. (2013); Guay and Dochain (2015); Benosman (2016); Grushkovskaya and Ebenbauer (2016); Ebenbauer et al. (2017); Poveda and Teel (2017); Scheinker and Krstić (2017); Suttner and Dashkovskiy (2017); Grushkovskaya et al. (2018); Guay and Atta (2018); Labar et al. (2019); Mandić et al. (2019)). A special place in these extremum seeking studies is given to nonlinear systems with dynamic input-output maps of the form

\[ \dot{x} = f(x, \xi), \quad x \in \mathbb{R}^n, \xi \in \mathbb{R}^m, \quad f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, \]

\[ y = h(x, \xi), \quad y \in \mathbb{R}^p, \quad h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p. \]

(1)

The classical extremum seeking problem statement for system (1) is to define the input \( \xi \) in such a way that the output of system (1) is optimized in the sense of minimization (or maximization) of an output-dependent cost function \( J : \mathbb{R}^p \to \mathbb{R} \). In this direction one can mention, e.g., the papers by Krstić and Wang (2000); Tan et al. (2006); Ghaffari et al. (2012); Guay and Dochain (2015); Haring and Johansen (2017); Dühr et al. (2017); Guay and Atta (2018). Typically, extremum seeking approaches for (1) are based on the construction of a dynamic extension \( \dot{\xi} = g(J(y), t) \), where \( g : \mathbb{R} \times [0, \infty) \to \mathbb{R}^m \) is chosen to ensure the desired vicinity of the trajectories of (1) to an optimal point. The analysis of the resulting system relies on singular perturbation theory and requires that system (1) admits a steady-state \( x = \ell(\xi) \), which is asymptotically stable for each fixed value of \( \xi \). Furthermore, a crucial assumption in such studies is the existence of certain Lyapunov function for system (1). However, there are many important classes of systems which do not admit a control Lyapunov function with desired properties.

In this paper, we consider a class of nonholonomic systems governed by driftless control-affine systems, in which the number of inputs can be significantly smaller than the number of state variables. In general, the linearization of these systems is not controllable. Moreover, as it was proved in the famous work by Brockett (1983), such nonholonomic systems cannot be stabilized by a continuous feedback law. To stabilize such systems one can use, e.g., discontinuous (e.g., Astolfi (1994); Clarke et al. (1997)) or time-varying feedback laws (e.g., Zuyev (2016); Grushkovskaya and Zuyev (2018)). Consequently, the resulting closed-loop system becomes discontinuous or non-autonomous and, in general, does not admit a regular Lyapunov function of the form \( V(x) \).
The goal of our paper is to construct extremum seeking controls for a class of nonholonomic systems with time-varying inputs adapted from Grushkovskaya and Zuyev (2018). We propose a novel solution of the extremum seeking problem for nonholonomic systems based on combination of stabilizing strategies for nonholonomic systems and gradient-free extremum seeking controllers. Although the main idea of our control design approach is inspired by singular perturbation techniques, we do not apply them directly in the proof. Instead, we propose a novel approach for dynamic stabilization of nonholonomic systems and generalize the techniques introduced in Grushkovskaya et al. (2018) to systems with multiple time scales.

The rest of this paper is organized as follows. In Section 2, we introduce basic notations, formulate the problem statement, and describe the main idea of our control design approach. Section 3 provides the main results of the paper, which are illustrated with an example in Section 4. Section 5 contains concluding remarks. Some auxiliary statements are given in Appendix A, and the proof of the main result is contained in Appendix B.

2. PRELIMINARIES

2.1 Notations and Definitions: $\delta_{ij}$ is the Kronecker delta; dist$(x, S)$ is the Euclidean distance between an $x \in \mathbb{R}^n$ and an $S \subseteq \mathbb{R}^n$: $B_r(x^*)$ is a $\delta$-neighborhood of an $x^* \in \mathbb{R}^n$; $\partial M$, $\overline{M}$ is the boundary and the closure of a set $M \subseteq \mathbb{R}^n$, respectively; $M = M \cup \partial M$; $|S|$ is the cardinality of a set $S$; $\mathcal{K}$ is the class of continuous strictly increasing functions $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\varphi(0) = 0$; $[f, g](x)$ is the Lie bracket of vector fields $f, g : \mathbb{R}^n \to \mathbb{R}^n$ at a point $x \in \mathbb{R}^n$, $[f, g](x) = Lg f(x) - Lf g(x)$, where $Lg f(x) = \lim_{s \to 0} \frac{f(x + sg(x) - f(x))}{s}$.

Similarly to Clarke et al. (1997); Zuyev (2016), we exploit the sampling approach for the stabilization of nonholonomic systems. Given an $\varepsilon > 0$, we define the partition $\pi_{\varepsilon}$ of $[0, +\infty)$ into the intervals $I_j = [t_j, t_{j+1})$, $t_j = j \varepsilon$, $j \in \mathbb{N} \cup \{0\}$.

Definition 1. Assume given a feedback $u = \varphi(x, \xi, t)$, $\varphi : D \times D \times [0, +\infty) \to \mathbb{R}^n$, $\varepsilon > 0$, and $x^0, \xi^0 \in D \subseteq \mathbb{R}^n$. A $\pi_{\varepsilon}$-solution of the system

$$\dot{x} = f(x, u), \ \xi = g(x, \xi, t), \ x, \xi \in D \subseteq \mathbb{R}^n, u \in \mathbb{R}^m,$$

(2) corresponding to $(x^0, \xi^0, \varphi)$, is an absolutely continuous function $(x^t(\cdot), \xi^t(\cdot))$ in $D \times D$, defined for $t \in [0, +\infty)$, which satisfies the initial conditions $x(0) = x^0$, $\xi(0) = \xi^0$ and the differential equations

$$\dot{x}(t) = f(x(t), \varphi(x(t), \xi^t(t), t)), \quad \dot{\xi}(t) = g(x(t), \xi(t), t) \quad \text{for each } j = 0, 1, 2, \ldots.$$

The above definition will be applied for the stabilization of nonholonomic systems using the approach of Zuyev (2016); Grushkovskaya and Zuyev (2018). However, the extremum seeking scheme proposed in this paper can also be used for output stabilization of systems with well-defined classical solutions.

2.2 Problem statement & Main idea. Consider a class of nonholonomic systems governed by driftless control-affine equations with single output:

$$\dot{x} = \sum_{i=1}^{m} u_i f_i(x), \quad y = J(x),$$

(3)

where $x = (x_1, \ldots, x_n)$ is the state, $x(0) = x^0 \in D$, $u = (u_1, \ldots, u_m)$ is the control, $m < n$, $y \in \mathbb{R}$ is the output of the system, $J : D \to \mathbb{R}$ is the cost function, and the vector fields $f_i : D \to \mathbb{R}^n$ are linearly independent. Let the following rank condition be satisfied in $D$:

$$\text{span}\{f_i(x), [f_i, f_j](x) | i \in S_1, j \in S_2\} = \mathbb{R}^n, \quad (4)$$

where $S_1 \subseteq \{1, 2, \ldots, m\}$ and $S_2 \subseteq \{1, 2, \ldots, m\}$ are some sets of indices, $|S_1| + |S_2| = n$. We study the following extremum seeking problem:

**Problem 1.** Let $J \in C^2(D; \mathbb{R})$ be a strongly convex function, and let $x^* \in D$ be such that $J(x) > J(x^*)$ for all $x \in D \setminus \{x^*\}$. The goal is to construct a control law $u = u(t, x, J(x))$ such that the trajectories $x(t)$ of system (3) with the initial conditions from $D$ tend asymptotically to an arbitrary small neighborhood of $x^*$.

The main idea of the control algorithm proposed in this paper can be described in two stages:

1. Model-based stabilizing component. For each value $\xi \in D$, we construct time-periodic fast oscillating control laws with state-dependent coefficients to ensure that the corresponding steady-state $x = \xi$ of (3) is asymptotically (and even exponentially) stable. Further we assume that $\xi(t)$ evolves according to certain differential equations, so the result of Zuyev (2016); Grushkovskaya and Zuyev (2018) cannot be directly applied for establishing stability properties of the extended system (2). Note that, in general, (3) does not admit a control Lyapunov function. Instead, we will prove that with the proposed choice of the control $u$ the trajectory $x(t)$ remains in a sufficiently small neighborhood of $\xi(t)$ for $t \in [0, \infty)$. These controls are model-based, i.e. the dynamics (control vector fields) and the coordinates of the system are assumed to be known, but not the analytical expression of $J$ and the optimal point $x^*$. We will apply sampling controllers, that is the solutions of (3) will be defined in the sense of Definition 1.

2. Model-free extremum seeking component. To optimize the state $x = \xi$ with respect to minimizing the cost function $J(x)$ along the trajectories of (3), we construct a dynamic extension $\xi = g(y, t)$, where $g(y, t)$ is taken in the form of fast oscillating time-periodic functions with output-dependent coefficients from (Grushkovskaya et al. (2018)). Thus, this part of the controller is model-free.

Remark 1. In Problem 1, we assume that the cost function $J$ depends only on the state variable $x$, but not on the control input $u$. This assumption is not crucial and is made in order to simplify the proof. Besides, if $J$ depends only on $u$, the stability properties directly follow from (Grushkovskaya et al. (2018)) and (Grushkovskaya and Zuyev (2018)) with the same proof techniques.

3. MAIN RESULTS

3.1 Control design. In this section, we formalize the control algorithm announced in Subsection 2.2. Namely, the overall system has the following form:

$$\dot{x} = \sum_{i=1}^{m} u_i f_i(x), \quad u_i = \varphi_i^r(x, \xi, t), \quad y = J(x), \quad x(0) = x^0, \quad (5a)$$

$$\xi = g(y, t), \quad g(y, t) = \sum_{j=1}^{2n} g_j(y) u_j^r(t) e_j, \quad \xi(0) = x^0. \quad (5b)$$

In (5a), the stabilizing component $u_i = \varphi_i^r(x, \xi, t)$ is
\[ \varphi_t^i(x, \xi) = \sum_{i_1 \in S_1} a_{i_1}(x, \xi) \delta_{i_1t} + \sqrt{\epsilon} \sum_{(i_1, i_2) \in S_2} \sqrt{\kappa_{ii_2}} [a_{i_1i_2}(x, \xi)] \times \delta_{ii_2} \text{sign}(a_{i_1i_2}(x, \xi)) \ \frac{2N}{\epsilon} \sigma(t) + \delta_{ii_2} \frac{2N}{\epsilon} \sigma(t). \] (6)

Here \( \kappa_{ii_2} \in \mathbb{N}, \kappa_{ii_2} \neq \kappa_{i_1i_2} \) for all \((i_1, i_2) \neq (i_2, i_1)\), and \(a(x, \xi) = \left( \left( a_{i_1i_2}(x, \xi) \right)_{i_1 \in S_1} \left( a_{i_1i_2i_3}(x, \xi) \right)_{(i_1, i_2) \in S_2} \right) \in \mathbb{R}^n\) is defined as

\[ a(x, \xi) = -\gamma_1 F^{-1}(x)(x - \xi) \] (7)

with \( F^{-1}(x) \) being the \( n \times n \) matrix inverse to

\[ F(x) = \left( \left( f_{i,j}(x) \right)_{j \in S_1} \left( f_{i,j}(x) \right)_{(j, j_2) \in S_2} \right), \]

and the control gain \( \gamma_2 > 0 \) to be defined later in the proof of the main result. Such a choice of \( \gamma_2 \) is aimed to ensure that the trajectories \( x(t) \) are close enough to \( \xi(t) \) for all \( t \geq 0 \) and all initial conditions \( x(0) \). Note that the rank condition (4) implies nonsingularity of \( F(x) \) for any \( x \in \mathcal{D} \).

In (5b), \( g(y, t) \) is the extremum seeking component. Here \( e_j \) denotes the unit vector in \( \mathbb{R}^n \) if \( j \leq n \), and non-zero \( (j - n) \)-th entry if \( n + 1 \leq j \leq 2n \), the functions \( g_j, g_{j+n} \) have the satisfaction

\[ [g_j(z), g_{j+n}(z)] = -\gamma_2, \gamma_2 > 0, j = \frac{n}{2}. \]

For example, the choice \( g_j, g_{j+n} = -\gamma_2 g_j, g_{j+n} \) for all \( j \neq \frac{n}{2} \) was proposed in (Grushkovskaya et al. (2018)). In this paper, we propose to parameterize the functions \( g_j, g_{j+n} \) as

\[ g_j(z) = r_j(z) \sin \phi_j(z), g_{j+n}(z) = r_j(z) \cos \phi_j(z), \]

with \( r_j, \phi_j \) such that \( r_j^2(z) \phi_j^2(z) = \gamma_2. \) (8)

The discrete-time version of the above parameterization has also been used by Feiilig et al. (2019).

Next, the inputs \( v_j^{(k)}(t) \) are given by

\[ v_j^{(k)}(t) = \begin{cases} \sqrt{\frac{4k_j}{\mu}} \cos \frac{2k_jt}{\mu} & \text{if } j = \frac{n}{2}, \\ \sqrt{\frac{4k_j}{\mu} \sin \frac{2k_jt}{\mu}} & \text{if } j = n + \frac{1}{2}, \end{cases} \]

where \( \mu > 0, k_j \in \mathbb{N}, k_j \neq k_{j+1}, \) for all \( j \neq \frac{n}{2} \).

Remark 2. Although the choice of \( g_j, g_{j+n} \) (8) may look similar, there are many extremum seeking systems whose control vectors field satisfies this relation. For example, the functions \( g_j(z) = z, g_{j+n}(z) = 1 \) have been exploited by D"urr et al. (2013a); D"urr et al. (2017); \( g_j(z) = \sin z, g_{j+n}(z) = \cos z \) by Schenker and Krstić (2017); \( g_j(z) = \sqrt{2} \sin \left( \text{ln} |z| \right), g_{j+n}(z) = \sqrt{2} \cos \left( \text{ln} |z| \right) \) by Sattner and Duchesne (2017); \( g_j(z) = \sqrt{\frac{1+e^{-t}}{1+e^{-t}}} \cos(e^t + 2 \ln(e^t - 1)), g_{j+n}(z) = \sqrt{\frac{1+e^{-t}}{1+e^{-t}}} \cos(e^t + 2 \ln(e^t - 1)) \) by Grushkovskaya et al. (2018). One more example will be given in Section 4.

3.2 Stability conditions. Assume that the cost function \( J \in C^2(\mathcal{D}; \mathbb{R}) \) satisfies the following properties in \( \mathcal{D} \):

\[ \sigma_{11} ||x - x^*||^2 - J(x) - J^* \leq \sigma_{12} ||x - x^*||^2, \]

\[ \sigma_{21} (J(x) - J^*) \leq ||\nabla J(x)||^2 \leq \sigma_{22} (J(x) - J^*), \]

\[ \frac{\partial^2 J(x)}{\partial x^2} \leq \sigma_3, \]

with \( x^* \in \mathcal{D} \) and some positive constants \( \sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}, \sigma_3 \). The main result of this paper is as follows. Theorem 1. Given system (3) and a function \( J \in C^2(\mathcal{D}; \mathbb{R}) \) satisfying (9), assume that:

- the vector fields \( f \in C^2(\mathcal{D}; \mathbb{R}^n) \) in (3) satisfy (4) in \( \mathcal{D} \), and there is an \( \alpha > 0 \) such that \( ||F^{-1}(x)|| \leq \alpha \) for all \( x \in \mathcal{D} \);
\[ \xi_j = \frac{4\pi^2}{\mu} g_1(y_1) \cos(2\pi k_j t/\mu) + g_2(y_2) \sin(2\pi k_j t/\mu) \epsilon_j, \quad (13) \]

\( j = 1, 2, 3 \).

In this example, we take \( y = J(x) = \|x\|^2 \), \( \gamma_1 = 20 \), and \( \gamma_2 = 1, k_1 = 4, k_2 = 2, k_3 = 3 \), and consider two types of functions \( g_1, g_2 \). The results of numerical simulations with the functions from Dürr et al. (2017),

\[ g_1(z) = z, \quad g_2(z) = 1, \quad (14) \]

are depicted on Fig. 1 (left). Here \( \epsilon = 0.1 \) and \( \mu = 0.5 \).

To improve the qualitative behavior of (11)–(13), we can apply another pair of the generating functions satisfying (8), which vanish when \( J \) takes its minimal value, e.g.,

\[ g_1(z) = \sqrt{\tanh z/2} \sin \left( 2 \ln(e^z - 1) - z \right), \quad (15) \]

\[ g_2(z) = \sqrt{\tanh z/2} \cos \left( 2 \ln(e^z - 1) - z \right) \] if \( z > 0, g_1(0) = g_2(0) = 0 \).

In this case, we took \( \epsilon = 0.25 \), \( \mu = 1 \). Note that, unlike the results of Grushkovskaya et al. (2018), the trajectories of (11)–(13) exhibit non-vanishing oscillations in a neighborhood of the extremum point (which are, however, considerably smaller than with the functions (14)) (see Fig. 1, right). Thus, an interesting question is whether it is possible to achieve asymptotic stability in the sense of Lyapunov with the proposed control algorithm.

In both case, we take the initial conditions \( x(0) = (1, -1, 1) \), \( \xi(0) = (-1, 1, 1) \) to illustrate that the proposed approach can be applied also for \( x^0 \neq \xi^0 \).

5. CONCLUSIONS & FUTURE WORK

To simplify the presentation, we consider only the class of nonholonomic systems (3) satisfying one-step bracket generating condition in this paper, i.e., we assume that the vector fields together with their Lie brackets span the whole \( n \)-dimensional space at each state \( x \in D \subseteq \mathbb{R}^n \). Another hypothesis is put in (9), so that the cost \( J \) possesses properties of a quadratic function. This hypothesis is introduced in order not to overcomplicate the proof of the main results. It should be emphasized that information about the analytical expression of \( J \) and its minimizer \( x^* \) is not required for the control design. Furthermore, all the constants in (9) may also be unknown. In future work, we expect to address broader classes of cost functions possessing polynomial convergence properties, similarly to the results of Grushkovskaya et al. (2018). We also plan to extend the proposed control design approach to nonholonomic systems under higher order controllability conditions with iterated Lie brackets.

REFERENCES


### Appendix A. AUXILIARY RESULTS

This section contains several technical results which be used for the proof of Theorem 1.

**Lemma 1.** Let $D \subseteq \mathbb{R}^n$, $\xi(t) \in D$, $t \in [0, \tau]$, be a solution of the system $\dot{\xi} = \sum_{i=1}^{m} h_i(\xi) \omega_i(t)$, and let the vector fields $h_i$ be Lipschitz continuous in $D$ with the Lipschitz constant $L$. Then $|\xi(t) - \xi(0)| \leq \nu t \max_{t \in [0, \tau]} \|h_i(\xi(0))\|e^{L \nu t}, t \in [0, \tau]$, with $\nu = \max_{t \in [0, \tau]} \sum_{i=1}^{m} \|w_i(t)\|$.

Lemma 1 follows from the Grönwall–Bellman inequality.

**Lemma 2.** (Zuyev and Grushkovskaya (2017)). Let vector fields $h_i$ be Lipschitz continuous in a domain $D \subseteq \mathbb{R}^n$, and $h_i \in C^2(D \times \mathbb{R}^n)$, where $\xi = \{\xi \in D : h_i(\xi) = 0 \text{ for } 1 \leq i \leq l\}$, and $L_{h_i} h_i, L_{h_i} h_i, L_{h_i} h_i \in C(D \times \mathbb{R}^n)$ for all $i,j,l \in \mathbb{N}$. If $t \xi(t) \in D$, $t \in [0, \tau]$, is a solution of $\dot{\xi} = \sum_{i=1}^{m} h_i(\xi) \omega_i(t)$ with $w_i(0) = 0$, $i \in D$, then $\xi(t)$ can be represented by the Chen–Fliess series:

$$\xi(t) = \sum_{i=1}^{\infty} h_i(\xi(t)) \int_{0}^{t} \omega_i \left( \frac{t}{s} \right) ds + \sum_{i,j,l=1}^{m} L_{h_i} h_j(\xi(t)) \int_{0}^{t} \int_{0}^{s} \omega_i \left( \frac{t}{s} \right) \omega_j \left( \frac{s}{r} \right) ds dr + \cdots$$

(App. A.1)

**Remark 1.** Let $D \subseteq \mathbb{R}^n$ be a bounded convex domain, $W \in C^2(D, \mathbb{R})$, $x \in D$, and let the following inequalities hold:

$$\sigma_{11} \|x - x^*\|^2 \leq W(x) \leq \gamma_{12} \|x - x^*\|^2,$$

$$\sigma_{21} W(x)^2 \leq \|\nabla W(x)\|^2 \leq \sigma_{22} W(x)^2 \leq \frac{1}{\sigma_{11}},$$

$$\frac{\sigma_{12}^2 \|x\|}{\sigma_{11}^2} \leq \sigma_{22} W(x)^2 \leq \frac{1}{\sigma_{11}},$$

where $m \geq 1$ and $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}, \sigma_{33}$ are positive constants. Then, for any $x^0 \in \mathbb{R}^n$ and any function $x : [0, \epsilon] \rightarrow D$ satisfying the conditions $x(0) = x^0, x(\epsilon) = x^0 - \gamma \nabla W(x^0) + r_0, \gamma > 0, r_0 \in \mathbb{R}^n$, the function $W$ satisfies the estimate:

$$W(x(\epsilon)) \leq W(x^0) \left(1 - \frac{\sigma_{11}}{m} \frac{1}{\sigma_{22}} \frac{\|x^0\|^2}{2m^2} \right),$$

where $x_1 = \gamma \sigma_{21} - \sqrt{2\sigma_{12}^2} \|r_0\| W \frac{1}{m} \|x^0\|^{-1}(x^0)/\epsilon, x_2 = ((m - 1)\sigma_{22} + m\sigma_{12} \left(\gamma \sqrt{\sigma_{12}^2 + \|r_0\| W \frac{1}{m} \|x^0\|^{-1}(x^0)/\epsilon \right)^2.$

**Appendix B. PROOF OF THEOREM 1**

For the sake of clarity, we divide the proof into several steps resulting in intermediate statements.

**Step 0. Notations and preliminary constructions.** To practically stabilize system (5) at $(0, x^*)$, we will focus on three parameters: $\gamma_1, \epsilon, \mu$, assuming that $\epsilon < \mu$. In the proof, we will determine big enough $\gamma_1 = \gamma_1(\mu)$, small enough $\epsilon = \epsilon(\gamma_1, \mu)$, and small enough $\mu$. It can be seen from the proof that such a choice is always possible. We use the following notations in the proof: for all $\tau \in [0, \epsilon]$,
Now let us put\\n\[ U(x(\cdot),\xi(\cdot),\tau) = \max_{x(\cdot)\in\mathbb{S}_t} \sum_{i=1}^m \|\varphi^*_i(x(\cdot),\xi(\cdot),\tau)\|, \]
\[ W(\tau) = \max_{0\leq\tau<\infty} \sum_{j=1}^n \|u_j^*(\tau)\| \leq \frac{\sqrt{2\pi}c_w}{\sqrt{\epsilon}}, \quad c_w = 2 \sum_{j=1}^n \sqrt{2\pi k_j}. \]

Recall that the state-dependent control coefficients are defined by (7), which implies that, for any \( x(0) = x_0 \in D, \xi(0) = \xi_0 \in D, [a(x_0,\xi_0)] \leq \gamma_0 \|x_0 - \xi_0\|. \) The Hölder inequality implies that, for any \( \epsilon > 0 \) and all \( \tau \in [0,\epsilon], \)
\[ U(x_0,\xi_0,\tau) = \epsilon \leq \epsilon \sum_{i\in S} a_i(x_0,\epsilon_0) + 2\sqrt{2\pi}\epsilon \]
\[ \times \sum_{(i,j)\in S} \kappa_{i,j}(x_0,\xi_0) \leq c_u \sqrt{\gamma_1\epsilon\|x_0 - \xi_0\|^2}. \quad (B.1) \]

and the \( \pi_\epsilon \)-solution \( x(\cdot) \) of system (5) can be represented my means of the Chen–Fliss series: \( x(\epsilon) = x_0 - \epsilon \gamma_1(x_0 - \xi_0) + R_1(\epsilon) + R_2(x_0,\xi_0,\epsilon). \) (B.8)

This proves the following intermediate statement.

**Statement 1.** For any \( \mu \in (0,\mu_0), \gamma_1 \in (\gamma_1,\infty), \epsilon \in (0,\epsilon(\gamma_1,\mu)], \)
\( x_0 \in D_\epsilon, \) the \( \pi_\epsilon \)-solutions of system (5) with the initial conditions \( x(0) = x_0, \xi(0) = \xi_0 \) satisfy the following property:
\[ \|x^0 - \xi^0\| \leq \frac{\rho_1 \mu^3 \sqrt{\epsilon}}{3} \Rightarrow \|x(t) - \xi(t)\| \leq \rho_1 \mu^3 \sqrt{\epsilon} \quad \text{for} \ t \in [0,\epsilon]. \]

Furthermore, if \( \xi_0 \in D_\epsilon \subseteq D' \) then \( x(t) \) is well-defined in \( D' \) for \( t \in [0,\epsilon]. \)

**Step 2.** Our next goal is to ensure that the \( x \)-component of the \( \pi_\epsilon \) solution of system (5) is in a sufficiently small neighborhood of the \( \xi \)-component. For this, we apply Lemma A.1. Namely, assume that \( x(t) \in D_\epsilon \) for \( t \in [0,\epsilon], \xi(0) = \xi_0 \in B_{\nu_1}(x_0). \)

Denote \( R(x_0,\xi_0,\epsilon) = R_1(\epsilon) + R_2(x_0,\xi_0,\epsilon). \) Using (B.1) and notations from (B.2), we get \( \|R_1(\epsilon)\| \leq M_3 f_{\rho_0}(\|x_0^0 - \xi_0^0\|/3) \quad \text{for all} \ t \in [0,\epsilon], \)
\[ \|R(x_0,\xi_0,\epsilon)\| \leq \tilde{\gamma}(\|x_0^0 - \xi_0^0\|) \]
\[ \left(\frac{2\mu^3}{3}\|x_0^0 - \xi_0^0\|/3 \right)^{3/4}\right) \quad \text{for} \ t \in [0,\epsilon]. \]

Combining (B.9), (B.7), and (B.8), we come to the following estimate:
\[ \|x(t) - \xi(t)\| \leq (1 - \epsilon \gamma_1)\|x(0) - \xi(0)\|^2 + \gamma_1(\|x(0) - \xi(0)\|^2)^{3/4} + \frac{\nu\epsilon}{\sqrt{\epsilon}} \]
For any \( \gamma_1 > \gamma_1, \lambda_1 \in (\gamma_1,\gamma_1) \) and define
\[ \epsilon_1(\gamma_1,\mu) = \min \left\{ \epsilon(\gamma_1,\mu), \frac{\gamma_1 - \lambda_1}{\lambda_1}\right\}. \]

Recall that \( \epsilon(\gamma_1,\mu) < \epsilon_1(\gamma_1,\mu) < 1. \) Then, for any \( \epsilon \in (0,\epsilon_1(\gamma_1,\mu)) \),
\[ \|x(t) - \xi(t)\| < (1 - \epsilon_1(\gamma_1,\mu))\|x(0) - \xi(0)\|^2 + \frac{\nu\epsilon}{\sqrt{\epsilon}} \]
Recall that \( \gamma_1 \) is given by (B.5), which implies \( \frac{\nu\epsilon}{\sqrt{\epsilon}} = \frac{\gamma_1\mu^3 \sqrt{\epsilon}}{3}. \) This together with Statement 1 gives us the next intermediate result.

**Statement 2.** Assume that \( x(t) \in D' \) for all \( t \in [0,\epsilon_1], x(0) \in D_\epsilon. \)
Then, for any \( \mu \in (0,\mu_0), \gamma_1 \in (\gamma_1,\infty), \epsilon \in (0,\epsilon(\gamma_1,\mu)], \mu \)
the following properties hold:
\[ \|x(t) - \xi(t)\| \leq \frac{\rho_1 \mu^3 \sqrt{\epsilon}}{3} \quad \text{then} \|x(t) - \xi(t)\| \leq \frac{\rho_1 \mu^3 \sqrt{\epsilon}}{3} \]
\[ \|x(t) - \xi(t)\| \leq \frac{\rho_1 \mu^3 \sqrt{\epsilon}}{3} \quad \text{for all} \ t \in [0,\epsilon]. \]

**Step 2.** Now let us put \( x(0) = x_0 = \xi_0 = \xi(0), x_0 \in D_\epsilon. \)
Then \( x(t) \equiv x(0) = x_0 \in D_\epsilon \) for all \( t \in [0,\epsilon], \)
\[ \|x(t) - \xi(t)\| \leq \|x(0) - \xi(0)\| \leq \frac{\nu\epsilon}{\sqrt{\epsilon}} + \delta_\epsilon \]
5471
The goal of this step is to ensure the decay of the \( \xi(t) \in D_{\xi} \subset D' \) for \( t \in [0, \varepsilon] \). Besides, Statement 2 implies \( \|x(\varepsilon) - \xi(\varepsilon)\| < \rho_1 \mu^3 \). From Statements 1 and 2, the \( x \)-component of the \( \pi_c \)-solution of system (5) is also well-defined in \( D' \) for \( t \in [0, \varepsilon] \). Again, it is easy to see that \( \|\xi(t) - \xi_0\| \leq \sqrt{\rho} + \delta_2 \) for \( t \in [0, \varepsilon] \), i.e. \( \xi(\varepsilon) = D_{\xi} \) and \( \|x(\varepsilon) - \xi(\varepsilon)\| < \rho_1 \mu^3 \). Without loss of generality, we may assume that \( \varepsilon = N_1 \), with some \( N_1 \in \mathbb{N} \). Repeating Steps 1–2 \( \varepsilon = N_1 \), we come to the following statement.

**Statement 3.** For any \( \mu \in (0, \mu_0) \), \( \gamma_1 \in (\gamma_1, \infty) \), \( \varepsilon \in (0, \varepsilon(\gamma_1, \mu)) \), the \( \pi_c \)-solutions \( (x(t), \xi(t)) \) of system (5) with the initial conditions \( x(0) = \xi(0) = D_{\xi} \) are well-defined in \( D' \times D' \) for all \( t \in [0, (N_1 + 1)\varepsilon] \), \( \|x(t) - \xi(t)\| \leq \rho_1 \mu^3 \sqrt{\rho} \) for all \( t \in [0, \mu] \), \( \|x(\mu) - \xi(\mu)\| < \rho_1 \mu^3 \sqrt{\rho} \). Thus, for any \( \mu \in (0, \mu_0) \), we can take \( \gamma_1(\mu, \mu) \), such that \( x(t), \xi(t) \in D' \) for \( t \in [0, \mu] \). In the next steps, we will find sufficiently small \( \mu \) independently on \( \varepsilon \) and \( \gamma_1 \).

**Step 4.** The goal of this step is to ensure the decay of the cost function \( J(x) \) along the trajectories of system (5) by choosing sufficiently small \( \mu \).

For this purpose, we apply again Lemma A.1. Since \( x(t), \xi(t) \in D' \) for \( t \in [0, \mu] \), we may consider the Chen–Fliss series expansion of the \( \xi \)-component of solution of system (5) on the interval \( [0, \mu]\):

\[
\xi(\mu) = \xi^0 + \mu \gamma_2 \nabla J(\xi^0) + R_3(\mu),
\]

where

\[
R_3(\mu) = \sum_{j=1}^{2n} \left[ \int_0^\mu \left( g_j \circ J(x(s)) - g_j \circ J(\xi(s)) \right) c_j v_j(s) ds \right]
+ \sum_{j,j',k=1}^{2n} \left[ \int_0^\mu \int_0^{s_1} L_{j,k,j',k} g_j \circ J(x(s_2)) - g_j \circ J(\xi(s_2)) c_j v_j(s_1) c_{j'} v_{j'}(s_2) ds_1 ds_2 \right]

+ \sum_{j,j',k=1}^{2n} \left[ \int_0^\varepsilon \int_0^{s_1} \int_0^{s_2} g_j \circ J(x(s_3)) c_j v_j(s_1) c_{j'} v_{j'}(s_2) ds_1 ds_2 ds_3 \right]

\]

Under the assumptions of Theorem 1, we conclude that \( \|R_3(\mu)\| \leq c_3 \mu^{1+\xi} \), where \( \xi = \min(\xi_1, 1/2) \), \( c_3 = c_4 \mu^2 \max(2, \varepsilon^{-1}/2) \rho_1 (L_G + \sqrt{\mu} L_2 \sigma_2 \mu_0) \max(\varepsilon^{-1/2}, \mu_0^{-1}) \sqrt{2} \).

Thus, applying Statement 3 we get \( \|R_3(\mu)\| \leq c_3 \mu^{1+\xi} \), where \( \xi = \min(\xi_1, 1/2) \), \( c_3 = c_4 \mu^2 \max(2, \varepsilon^{-1}/2) \rho_1 (L_G + \sqrt{\mu} L_2 \sigma_2 \mu_0) \max(\varepsilon^{-1/2}, \mu_0^{-1}) \sqrt{2} \).

Using Taylor's formula for the function \( J(\xi) \),

\[
J(\xi(t)) = J(\xi^0) + \nabla J(\xi^0)(\xi(t) - \xi^0)
\]

and exploiting (9), we obtain

\[
J(\xi(t)) \leq J(\xi^0) - \mu \gamma_2 \sigma_2 J(\xi^0) + \mu^{1+\xi} \sigma_2 \sqrt{2} \mu^{1/2} J(\xi^0)
+ \sigma_1 \mu^2 \sigma_2^2 J(\xi^0) + \xi \mu^2 \sigma_2 J(\xi^0) + \mu^{1+\xi} J(\xi^0)
\]

\[
= J(\xi^0) \left( 1 - \mu \gamma_2 \sigma_2 + \mu \gamma_2 \sigma_2 \sigma_2 + \mu \gamma_2 \sigma_2 \sigma_2 \right)
\]

\[
+ \mu^{1+\xi} \sigma_2 \sqrt{2} \mu^{1/2} J(\xi^0) + \sigma_1 \mu^2 \sigma_2 J(\xi^0) + \mu^{1+\xi} \sigma_2 \mu^{1+\xi} J(\xi^0).
\]

Let \( L_c = \{ x \in D : J(x) < c \} \), \( c \in [\sigma_1, \delta_2^2] \). Then

\[
\frac{\partial^2 J(x)}{\partial \sigma_i \partial \sigma_j} \leq c_j \sigma_i \delta_2^2.
\]

For any \( \rho \in (0, \delta_j] \), \( \rho \in (0, \sigma_2 \delta_2 \rho_1) \), we define

\[
\rho_1 = \min(\mu_0, 1/2, \mu_1).
\]