Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Extremum Seeking Approach for Nonholonomic Systems with Multiple Time Scale Dynamics ⋆

Victoria Grushkovskaya * ,**** **Alexander Zuyev** ** ,*** ,****

* *Institute of Mathematics, Alpen-Adria University of Klagenfurt, Austria (e-mail: viktoriia.grushkovskay@aau.at).*
** *Otto von Guericke University Magdeburg, Germany*
*** *Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany (e-mail: zuyev@mpi-magdeburg.mpg.de)*
**** *Institute of Applied Mathematics & Mechanics, NAS of Ukraine*

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 Ariyur (2003); Tan et al. (2006); Nešić et al. (2010); Fu and Özgüner (2011); Liu and Krstić (2012); Dürr 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

$$\begin{aligned} &\dot{x} = f(x,\xi), \quad x \in \mathbb{R}^n, \xi \in \mathbb{R}^m,\ f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, \\ &y = h(x,\xi), \quad y \in \mathbb{R}^p,\ h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p. \end{aligned} \tag{1}$$

The classical extremum seeking problem statement for system (1) is to define the input ξ 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); Dürr et al. (2017); Guay and Atta (2018). Typically, extremum seeking approaches for (1) are based on the construction of a dynamic extension $\dot{\xi} = g\big(J(y), t\big)$, 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 ξ. 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 $V(x)$.

⋆ This work was supported in part by the German Research Foundation (projects GR 5293/1-1 and ZU 359/2-1).

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: δ_{ij} is the Kronecker delta; $\operatorname{dist}(x,S)$ is the Euclidian distance between an $x\in\mathbb{R}^n$ and an $S\subset\mathbb{R}^n$; $B_\delta(x^*)$ is a δ-neighborhood of an $x^*\in\mathbb{R}^n$; ∂M, $\overline{M}$ is the boundary and the closure of a set $M\subset\mathbb{R}^n$, respectively; $\overline{M}=M\cup\partial M$; $|S|$ 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 $x\in\mathbb{R}^n$, $[f,g](x)=L_fg(x)-L_gf(x)$, where $L_gf(x)=\lim_{s\to0}\frac{f(x+sg(x))-f(x)}{s}$.

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 π_ε of $[0,+\infty)$ into the intervals $I_j=[t_j,t_{j+1})$, $t_j=\varepsilon j$, $j\in\mathbb{N}\cup\{0\}$.

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

$$\dot x=f(x,u),\ \dot\xi=g(x,\xi,t),\ x,\xi\in D\subseteq\mathbb{R}^n,u\in\mathbb{R}^m,\qquad(2)$$

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

$$\dot x(t)=f\big(x(t),\varphi(x(t_j),\xi(t_j),t)\big),\quad t\in I_j=[t_j,t_{j+1}),$$
$$\dot\xi(t)=g(x(t),\xi(t),t)\text{ for each }j=0,1,2,\dots\ .$$

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:

$$\dot x=\sum_{i=1}^m u_if_i(x),\ y=J(x),\qquad(3)$$

where $x=(x_1,\dots,x_n)^\top\in D\subseteq\mathbb{R}^n$ is the state, $x(0)=x^0\in D$, $u=(u_1,\dots,u_m)^\top\in\mathbb{R}^m$ 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:

$$\operatorname{span}\big\{f_i(x),[f_{j_1},f_{j_2}](x)\,|\,i\in S_1,(j_1,j_2)\in S_2\big\}=\mathbb{R}^n,\qquad(4)$$

where $S_1\subseteq\{1,2,...,m\}$ and $S_2\subseteq\{1,2,...,m\}^2$ 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 $\dot\xi=g(y,t)$, 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:

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

$$\dot\xi=g(y,t),\ g(y,t)=\sum_{j=1}^{2n}g_j(y)v_j^\mu(t)e_j,\qquad\xi(0)=x^0.\qquad(5b)$$

In (5a), the stabilizing component $u_i=\varphi_i^\varepsilon(x,\xi,t)$ is

$$\varphi_i^\varepsilon(x,\xi,t)=\sum_{i_1\in S_1} a_{i_1}(x,\xi)\delta_{ii_1}+\sqrt{\frac{4\pi}{\varepsilon}}\sum_{(i_1,i_2)\in S_2}\sqrt{\kappa_{i_1i_2}|a_{i_1i_2}(x,\xi)|}$$
$$\times\Big(\delta_{ii_1}\operatorname{sign}(a_{i_1,i_2}(x,\xi))\cos\frac{2\pi\kappa_{i_1i_2}}{\varepsilon}t+\delta_{ii_2}\sin\frac{2\pi\kappa_{i_1i_2}}{\varepsilon}t\Big). \quad (6)$$

Here $\kappa_{i_1i_2}\in\mathbb{N}$, $\kappa_{i_1i_2}\neq\kappa_{i_3i_4}$ for all $(i_1,i_2)\neq(i_3,i_4)$, and $a(x,\xi)=\Big((a_{i_1}(x,\xi))_{i_1\in S_1}\ (a_{i_1i_2}(x,\xi))_{(i_1,i_2)\in S_2}\Big)^\top\in\mathbb{R}^n$ is defined as

$$a(x,\xi)=-\gamma_1\mathcal{F}^{-1}(x)(x-\xi) \quad (7)$$

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

$$\mathcal{F}(x)=\Big(\big(f_{j_1}(x)\big)_{j_1\in S_1}\ \ \big([f_{j_1},f_{j_2}](x)\big)_{(j_1,j_2)\in S_2}\Big),$$

and the control gain $\gamma_1>0$ to be defined later in the proof of the main result. Such a choice of u_i 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 $\mathcal{F}(x)$ for any $x\in D$. In (5b), $g(y,t)$ is the extremum seeking component. Here e_j denotes the unit vector in $\mathbb{R}^n$ with non-zero j-th entry 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 to satisfy the relation

$$[g_j(z),g_{j+n}(z)]=-\gamma_2,\ \gamma_2>0,\ j=\overline{1,n}.$$

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

$$g_j(z)=r_j(z)\sin\phi_j(z),\ g_{j+n}(z)=r_j(z)\cos\phi_j(z),\quad \text{with } r_j,\phi_j \text{ such that } r_j^2(z)\phi_j'(z)\equiv\gamma_2. \quad (8)$$

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

Next, the inputs $v_j^\mu(t)$ are given by

$$v_j^\mu(t)=\begin{cases}\sqrt{\dfrac{4\pi k_j}{\mu}}\cos\dfrac{2\pi k_j t}{\mu} & \text{for } j=\overline{1,n},\\ \sqrt{\dfrac{4\pi k_{j-n}}{\mu}}\sin\dfrac{2\pi k_{j-n} t}{\mu} & \text{for } j=\overline{n+1,2n},\end{cases}$$

where $\mu>0$, $k_j\in\mathbb{N}$, $k_{j_1}\neq k_{j_2}$ for all $j_1\neq j_2$.

Remark 2. Although the choice of g_j,g_{j+n} in (8) may look rather artificial, there are many extremum seeking systems whose control vector fields satisfy this relation. For example, the functions $g_j(z)=z$, $g_{j+n}(z)=1$ have been exploited by Dürr et al. (2013a); Dürr et al. (2017); $g_j(z)=\sin z$, $g_{j+n}(z)=\cos z$ by Scheinker and Krstić (2017); $g_j(z)=\sqrt{z}\sin(\ln z)$, $g_{j+n}(z)=\sqrt{z}\cos(\ln z)$ by Suttner and Dashkovskiy (2017); $g_j(z)=\sqrt{\frac{1-e^{-z}}{1+e^z}}\sin(e^z+2\ln(e^z-1))$, $g_{j+n}(z)=\sqrt{\frac{1-e^{-z}}{1+e^z}}\cos(e^z+2\ln(e^z-1))$ by Grushkovskaya et al. (2018). One more example will be given in Section 4.

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

$$\sigma_{11}\|x-x^*\|^2\leq J(x)-J^*\leq\sigma_{12}\|x-x^*\|^2, \quad (9)$$
$$\sigma_{21}(J(x)-J^*)\leq\|\nabla J(x)\|^2\leq\sigma_{22}(J(x)-J^*),\ \left\|\frac{\partial^2 J(x)}{\partial x^2}\right\|\leq\sigma_3,$$

with $x^*\in D$ and some positive constants σ_{11}, σ_{12}, σ_{21}, σ_{22}, σ_3. The main result of this paper is as follows.

Theorem 1. Given system (3) and a function $J\in C^2(D;\mathbb{R})$ satisfying (9), assume that:
– the vector fields $f_i\in C^2(D;\mathbb{R}^n)$ in (3) satisfy (4) in D, and there is an $\alpha>0$ such that $\|\mathcal{F}^{-1}(x)\|\leq\alpha$ for all $x\in D$;
– $g_j(J(\cdot))\in C^2(D\backslash\{x^*\};\mathbb{R})$, $L_{g_j}g_i(J(\cdot))$, $L_{g_l}L_{g_j}g_i(J(\cdot))\in C(D;\mathbb{R})$ for all $i,j,l=\overline{1,2n}$;
– for any compact $D'\subseteq D$, there are $L_g,L_{2g},M_{3g}\geq 0$ s.t.

$$\|g_i\big(J(x)\big)-g_i\big(J(\xi)\big)\|\leq L_g\|x-\xi\|,$$
$$\|L_{\left(g_{j_2}(J(x))-g_{j_2}(J(\xi))\right)}g_{j_1}\big(J(\xi)\big)\|\leq L_{2g}\|x-\xi\|,$$
$$\|L_{g_{j_3}(J(x))}L_{g_{j_2}(J(\xi))}g_{j_1}\big(J(\xi)\big)\|\leq M_{3g},\quad x,\xi\in D',\ i,j,l=\overline{1,2n}.$$

Then, for any $\delta\in\Big(0,\sqrt{\sigma_{11}/\sigma_{12}}\operatorname{dist}(x^*,\partial D)\Big)$ and any $\rho>0$, there exist $\bar\mu>0$, $\bar\gamma_1(\mu)>0$, and $\bar\varepsilon(\gamma_1,\mu)>0$ such that, for any $\mu\in(0,\bar\mu]$, $\gamma_1\in[\bar\gamma_1(\mu),\infty)$, and any $\varepsilon\in(0,\bar\varepsilon(\gamma_1,\mu)]$, each π_ε solution of (5) with $u_i=\varphi_i^\varepsilon(x,\xi,t)$ defined by (6) and the initial conditions from $\overline{B_\delta(x^*)}$ satisfies

$$\|x(t)-x^*\|\leq\beta\|x^0-x^*\|e^{-\lambda t}+\rho \text{ for all } t\in[0,\infty), \quad (10)$$

with some $\beta,\lambda>0$.

The proof of this theorem is given in Appendix B.

Remark 3. The proof of Theorem 1 represents a constructive procedure for choosing $\bar\mu$, $\bar\gamma_1(\mu)$, $\bar\varepsilon(\gamma_1,\mu)$, and clarifies the relation between these parameters and the coefficients β and λ. We would like to underline that the proposed bounds are quite conservative. The crucial assumption is $\varepsilon<\mu$, which means that subsystem (5a) oscillates faster than subsystem (5b). To simplify the proof, we also suppose that $\frac{\mu}{\varepsilon}\in\mathbb{N}$ and $x(0)=\xi(0)$, however the assertion of Theorem 1 can also be obtained without these assumptions.

In order to have γ_1 independent on μ, one may introduce an additional parameter η which will ensure a "slow" dynamics of (5b) (similarly to, e.g., Dürr et al. (2017)). This, however, will result in a slower convergence rate of the overall system to the optimal point. Namely, by taking $\tilde v_j^\mu(t):=\frac{1}{\eta}v_j^\mu\Big(\frac{t}{\eta}\Big)$ in (5b) and keeping the conditions of Theorem 1, one can prove the following statement:
For any $\delta\in\Big(0,\sqrt{\frac{\sigma_{11}}{\sigma_{12}}}\operatorname{dist}(x^*,\partial D)\Big)$ and any $\rho>0$, there exist $\bar\mu>0$, $\bar\varepsilon(\mu)>0$, and $\bar\eta(\varepsilon,\mu)>0$ such that, for any $\mu\in(0,\bar\mu]$, $\varepsilon\in(0,\bar\varepsilon(\mu)]$ and $\eta\in[\bar\eta(\varepsilon,\mu),\infty)$, each π_ε-solution of (5) with $u_i=\varphi_i^\varepsilon(x,\xi,t)$ defined by (6) and the initial conditions from $\overline{B_\delta(x^*)}$ satisfies $\|x(t)-x^*\|\leq\beta\|x^0-x^*\|e^{-\frac{\lambda t}{\eta}}+\rho$ for all $t\in[0,\infty)$, $\beta,\lambda>0$.
Similarly to Grushkovskaya et al. (2018), the behavior of the solutions of (3) can be improved by generating g_j vanishing at x^*. We will illustrate this feature with an example in the next section.

4. EXAMPLE

As an example, consider the well-known Brockett integrator (Brockett (1983)):

$$\dot x_1=u_1,\ \dot x_2=u_2,\ \dot x_3=u_1x_2-u_2x_1. \quad (11)$$

It is easy to see that, for all $x\in\mathbb{R}^3$, the vector fields $f_1=(1,0,x_2)^\top$ and $f_2=(0,1,-x_1)^\top$ of system (11) satisfy the rank condition (4) with $S_1=\{1,2\}$, $S_2=\{(1,2)\}$: $\operatorname{span}\{f_1(x),\ f_2(x),\ [f_1,f_2](x)\}=\mathbb{R}^3$ for all $x\in\mathbb{R}^3$; thus, we may apply the control algorithm proposed in Section 3.1. Namely, we take

$$u_1=a_1(x,\xi)+\sqrt{4\pi\kappa_{12}\big|a_{12}(x)\big|/\varepsilon}\ \operatorname{sign}\big(a_{12}(x,\xi)\big)\cos(2\pi\kappa_{12}t/\varepsilon),$$
$$u_2=a_2(x,\xi)+\sqrt{4\pi\kappa_{12}\big|a_{12}(x)\big|/\varepsilon}\ \sin(2\pi\kappa_{12}t/\varepsilon), \quad (12)$$
$$a(x,\xi)=\big(a_1(x,\xi),a_2(x,\xi),a_{12}(x,\xi)\big)^\top=-\gamma_1\mathcal{F}^{-1}(x)(x-\xi)$$
$$=-\gamma_1\left(x_1-\xi_1,x_2-\xi_2,\tfrac{1}{2}\big(-x_2\xi_1+x_1\xi_2-x_3+\xi_3\big)^\top\right),$$

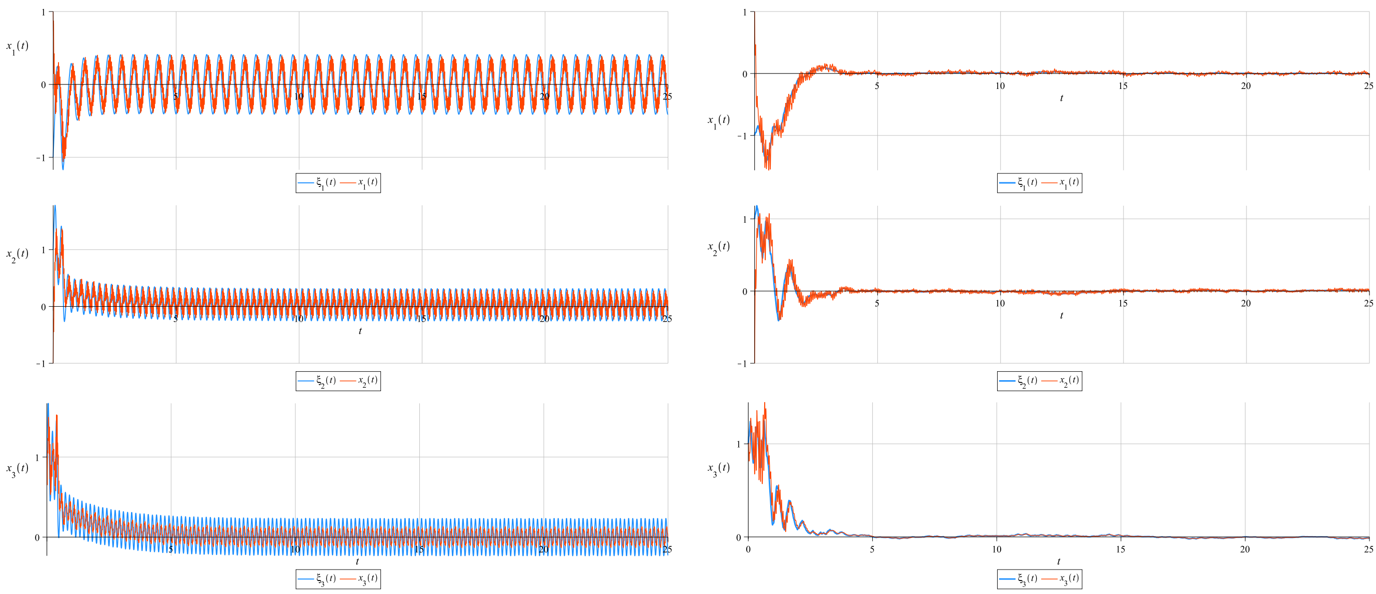

Fig. 1. Time-plots of the trajectories of system (11)–(13) with the generating functions (14) (left) and (15) (right).

$$\dot{\xi}_j = \sqrt{4\pi k_j/\mu}\Big(g_1(y)\cos(2\pi k_j t/\mu) + g_2(y)\sin(2\pi k_j t/\mu)\Big)e_j, \quad (13)$$

$j = 1, 2, 3$. In this example, we take $y{=}J(x){=}\|x\|^2$, $\gamma_1{=}20$, $\gamma_2{=}1$, $\kappa_{12}{=}4$, $k_1{=}1$, $k_2{=}2$, $k_3{=}3$, and consider two types of functions g_1, g_2. The results of numerical simulations with the functions from Dürr et al. (2017),

$$g_1(z) = z,\; g_2(z) = 1, \quad (14)$$

are depicted on Fig. 1 (left). Here $\varepsilon = 0.1$ 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.,

$$g_1(z){=}\sqrt{\tanh z/2}\sin\Big(2\ln(e^z{-}1){-}z\Big), \quad (15)$$
$$g_2(z){=}\sqrt{\tanh z/2}\cos\Big(2\ln(e^z{-}1){-}z\Big) \text{ if } z>0, g_1(0){=}g_2(0){=}0.$$

In this case, we took $\varepsilon{=}0.25$, $\mu{=}1$. 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 $x(0) = (1, -1, 1)^\top$, $\xi(0) = (-1, 1, 1)^\top$ 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

Astolfi, A. (1994). On the stabilization of nonholonomic systems. In *Proc. 33rd IEEE Conference on Decision and Control*, volume 4, 3481–3486.

Benosman, M. (2016). *Learning-Based Adaptive Control: An Extremum Seeking Approach–Theory and Applications.* Butterworth-Heinemann.

Brockett, R.W. (1983). Asymptotic stability and feedback stabilization. *Differential Geometric Control Theory*, 181–191.

Clarke, F.H., Ledyaev, Y.S., Sontag, E.D., and Subbotin, A.I. (1997). Asymptotic controllability implies feedback stabilization. *IEEE Tran on Automatic Control*, 42(10), 1394–1407.

Dürr, H.B., Krstić, M., Scheinker, A., and Ebenbauer, C. (2017). Extremum seeking for dynamic maps using Lie brackets and singular perturbations. *Automatica*, 83, 91–99.

Dürr, H.B., Stanković, M.S., Ebenbauer, C., and Johansson, K. (2013a). Lie bracket approximation of extremum seeking systems. *Automatica*, 49, 1538–1552.

Dürr, H.B., Zeng, C., and Ebenbauer, C. (2013b). Saddle point seeking for convex optimization problems. *Proc. 9th IFAC Symposium on Nonlinear Control Systems*, 540–545.

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

Feiling, J., Labar, C., Grushkovskaya, V., Garone, E., Kinnaert, M., and Ebenbauer, C. (2019). Extremum seeking algorithms based on non-commutative maps. *IFAC-PapersOnLine*, 52(16), 688–693.

Fu, L. and Özgüner, Ü. (2011). Extremum seeking with sliding mode gradient estimation and asymptotic regulation for a class of nonlinear systems. *Automatica*, 47(12), 2595–2603.

Ghaffari, A., Krstić, M., and Nešić, D. (2012). Multivariable newton-based extremum seeking. *Automatica*, 48(8), 1759–1767.

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

Grushkovskaya, V. and Zuyev, A. (2018). Obstacle avoidance problem for second degree nonholonomic systems. In *Proc. 57th IEEE Conf. on Decision and Control*, 1500–1505.

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

Guay, M. and Atta, K.T. (2018). Dual mode extremum-seeking control via Lie-bracket averaging approximations. In *Proc. 2018 Annual American Control Conference*, 2972–2977.

Guay, M. and Dochain, D. (2015). A time-varying extremum-seeking control approach. *Automatica*, 51, 356–363.

Haring, M., Van De Wouw, N., and Nešić, D. (2013). Extremum-seeking control for nonlinear systems with periodic steady-state outputs. *Automatica*, 49(6), 1883–1891.

Haring, M.A. and Johansen, T.A. (2017). Asymptotic stability of perturbation-based extremum-seeking control for nonlinear systems. *IEEE Trans. Autom. Control*, 62(5), 2302–2317.

Krstić, M. and Ariyur, K.B. (2003). *Real-Time optimization by Extremum Seeking Control*. Wiley-Interscience.

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

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

Liu, S.J. and Krstić, M. (2012). *Stochastic averaging and stochastic extremum seeking*. Springer Science & Business Media.

Mandić, F., Mišković, N., and Lončar, I. (2019). Underwater acoustic source seeking using time-difference-of-arrival measurements. *IEEE Journal of Oceanic Engineering*.

Nešić, D., Tan, Y., Moase, W.H., and Manzie, C. (2010). A unifying approach to extremum seeking: Adaptive schemes based on estimation of derivatives. In *Prc. 49th IEEE Conf. on Decision and Control*, 4625–4630.

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

Scheinker, A. and Krstić, M. (2017). *Model-free stabilization by extremum seeking*. Springer.

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

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

Zuyev, A. (2016). Exponential stabilization of nonholonomic systems by means of oscillating controls. *SIAM J. on Control and Optimization*, 54(3), 1678–1696.

Zuyev, A. and Grushkovskaya, V. (2017). Motion planning for control-affine systems satisfying low-order controllability conditions. *International Journal of Control*, 90(11), 2517–2537.

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 $\dot\xi=\sum_{i=1}^l h_i(\xi)w_i(t)$, and let the vector fields h_i be Lipschitz continuous in D with the Lipschitz constant L. Then $\|\xi(t)-\xi(0)\|\le t\nu\max_{1\le i\le l}\|h_i(\xi(0))\|e^{\nu Lt}$, $t\in[0,\tau]$, with $\nu=\max\limits_{t\in[0,\tau]}\sum_{i=1}^l|w_i(t)|$.*

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 $h_i\in C^2(D\setminus\Xi;\mathbb{R})$, where $\Xi=\{\xi\in D{:}h_i(\xi)=0$ for $1\le i\le l\}$, and $L_{h_j}h_i, L_{h_l}L_{h_j}h_i\in C(D;\mathbb{R}^n)$ for all $i,j,l=\overline{1,l}$. If $\xi(t)\in D$, $t\in[0,\tau]$, is a solution of $\dot\xi=\sum_{i=1}^l h_i(\xi)w_i(t)$ with $u\in C([0,\tau];\mathbb{R}^m)$ and $x(0)=x^0\in D$, then $\xi(t)$ can be represented by the Chen–Fliess series:*

$$\begin{aligned}\xi(t)=\xi^0+\sum_{i_1=1}^{l}h_{i_1}(\xi^0)\int\limits_0^t w_{i_1}(v)dv+\sum_{i_1,i_2=1}^{l}L_{h_{i_2}}h_{i_1}(\xi^0)\\ \times\int\limits_0^t\int\limits_0^v w_{i_1}(v)w_{i_2}(s)dsdv+R(t),\ t\in[0,\tau],\end{aligned}\tag{A.1}$$

$$R(t)=\sum_{i_1,i_2,i_3=1}^{l}\int\limits_0^t\int\limits_0^v\int\limits_0^s L_{h_{i_3}}L_{h_{i_2}}h_{i_1}(\xi(p))w_{i_1}(v)w_{i_2}(s)w_{i_3}(p)dpdsdv$$

is the remainder of the Chen–Fliess series expansion.

Lemma 3. (follows from Grushkovskaya et al. (2018)). *Let the conditions of Lemma 1 be satisfied and let $\xi^*\in D$. Assume that there exist $M_1,M_3\ge0$, $m\ge1$, $\varpi\in\{0\}\cup[1,\infty)$ such that*

$$\max_{1\le i_1\le l}\|h_{i_1}(\xi(0))\|\le M_1\|\xi(0)-\xi^*\|^m,$$
$$\max_{1\le i_1,i_2,i_3\le l}\|L_{h_{i_3}}L_{h_{i_2}}h_{i_1}(\xi)\|\le M_3\|\xi-\xi^*\|^\varpi\ \text{for all}\ \xi\in D.$$

Then, for all $t\in[0,\tau]$, the remainder $R(t)$ of the Chen–Fliess expansion (A.1) *of $x(t)$ satisfies the estimate*

$$\begin{aligned}\|R(t)\|\le\tfrac{2^{\varpi-2}}{3}(t\nu)^3\|\xi(0)-\xi^*\|^\varpi M_3\\ \times\Big(1+M_1(\tau\nu)^\varpi e^{\nu L\varpi\tau}\|\xi(0)-\xi^*\|^{\varpi(m-1)}\Big).\end{aligned}$$

Lemma 4. (Grushkovskaya et al. (2018)). *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:*

$$\sigma_{11}\|x-x^*\|^{2m}\le W(x)\le\gamma_{12}\|x-x^*\|^{2m},$$
$$\sigma_{21}W(x)^{2-\frac1m}\le\|\nabla W(x)\|^2\le\sigma_{22}W(x)^{2-\frac1m},$$
$$\left\|\tfrac{\partial^2W(x)}{\partial x^2}\right\|\le\sigma_3W(x)^{1-\frac1m},$$

where $m\ge1$ and $\sigma_{11},\sigma_{12},\sigma_{21},\sigma_{22},\sigma_3$ are positive constants. Then, for any $x^0=x(0)\in D\setminus\{x^*\}$ and any function $x:[0,\varepsilon]\to D$ satisfying the conditions

$$x(0)=x^0,\ x(\varepsilon)=x^0-\gamma\varepsilon\nabla W(x^0)+r_\varepsilon,\ \gamma>0,\ r_\varepsilon\in\mathbb{R}^n,$$

the function W satisfies the estimate:

$$W(x(\varepsilon))\le W(x^0)\Big(1-\frac{\varepsilon\varkappa_1}{m}W^{1-\frac1m}(x^0)+\frac{\varepsilon^2\varkappa_2}{2m^2}W^{2-\frac2m}(x^0)\Big)^m,$$

where $\varkappa_1=\gamma\sigma_{21}-\sqrt{\sigma_{22}}\|r_\varepsilon\|W^{\frac{1}{2m}-1}(x^0)/\varepsilon$, $\varkappa_2=((m-1)\sigma_{22}+m\sigma_{12})\Big(\gamma\sqrt{\sigma_{22}}+\|r_\varepsilon\|W^{\frac{1}{2m}-1}(x^0)/\varepsilon\Big)^2$.

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: γ_1, ε, and μ, assuming that $\varepsilon<\mu$. In the proof, we will determine big enough $\gamma_1=\gamma_1(\mu)$, small enough $\varepsilon=\varepsilon(\gamma_1,\mu)$, and small enough μ. It can be seen from the proof that such a choice is always possible. We use the following notations in the proof: for any $\tau\in[0,\varepsilon]$,

$$U(x(\varepsilon),\xi(\varepsilon),\tau)=\max_{\varepsilon\le t\le\varepsilon+\tau}\sum_{i=1}^{m}|\varphi_i^{\varepsilon}(x(\varepsilon),\xi(\varepsilon),t)|,$$

$$W(\tau)=\max_{0\le t\le\tau}\sum_{j=1}^{2n}|v_j^{\tau}(t)|\le\tfrac{c_w}{\sqrt{\mu}},\ c_w=2\sum_{j=1}^{n}\sqrt{2\pi k_j}.$$

Recall that the state-dependent control coefficients are defined by (7), which implies that, for any $x(0)=x^0\in D$, $\xi(0)=\xi^0\in D$, $\|a(x^0,\xi^0)\|\le\gamma_1\alpha\|x^0-\xi^0\|$. The Hölder inequality implies that, for any $\varepsilon>0$ and all $\tau\in[0,\varepsilon]$,

$$\begin{aligned}U(x^0,\xi^0,\tau)\varepsilon&\le\varepsilon\sum_{i\in S_1}|a_i(x^0,\xi^0)|+2\sqrt{2\pi\varepsilon}\\&\times\sum_{(i_1,i_2)\in S_2}\sqrt{\kappa_{i_1i_2}|a_{i_1i_2}(x^0,\xi^0)|}\le c_u\sqrt{\gamma_1\varepsilon\|x^0-\xi^0\|},\end{aligned}\tag{B.1}$$

$c_u=\sqrt{\alpha}(\sqrt{\gamma_1\alpha\varepsilon\|x^0-\xi^0\||S_1|}+2\sqrt{2\pi})\Big(\sum_{(j_1,j_2)\in S_2}\kappa_{j_1j_2}^{2/3}\Big)^{3/4}$ is strictly monotonically increasing w.r.t. ε. For any $\delta\in\Big(0,\sqrt{\frac{\sigma_{11}}{\sigma_{12}}}\mathrm{dist}(x^*,\partial D)\Big)$, let $\delta_x\in\Big(\sqrt{\frac{\sigma_{12}}{\sigma_{11}}}\delta,\mathrm{dist}(x^*,\partial D)\Big)$, and let D' be compact, $D_x=\overline{B_{\delta_x}(x^*)}\subset D'\subseteq D$. If D is compact, then we take $D'=D$. By the conditions of Theorem 1, there exist $M_f,M_g,M_{3g}>0$ such that, for all $x,\xi\in D_x$,

$$\|f_i(x)\|\le M_f,\ \|g_j(J(x))\|\le M_g,\ i=\overline{1,m},\ j=\overline{1,2n}\tag{B.2}$$

$$\|L_{f_{j_1}}f_{j_2}(x)\|\le M_{2f},\ \big\|L_{f_{j_3}}L_{f_{j_2}}f_{j_1}(x)\big\|\le 6M_{3f},\ j_1,j_2,j_3=\overline{1,m}.$$

If (B.2) and inequalities from the fourth condition of Theorem 1 hold globally in D, then we take $D'=D$.

Step 1. *At this step we construct some a priori estimates which will be exploited further in the proof.*

It is easy to see that the π_ε-solutions of system (5) satisfy

$$\|x(t)-x(0)\|\le M_fc_u\sqrt{\gamma_1\varepsilon\|x(0)-\xi(0)\|}\text{ for all }t\in[0,\varepsilon].\tag{B.3}$$

Let $\rho_1>0$ be given, $\nu=c_wM_g$, $\varsigma>0$, $\rho_0=\rho_1\mu^\varsigma\sqrt{\mu}/3$, $\delta_\xi=\delta_x+\nu\sqrt{\mu}$, $D_\xi=\overline{B_{\delta_\xi}(x^*)}$, $d=\mathrm{dist}(x^*,\partial D')-\delta_x>0$, and let μ_0 be the smallest positive root of the equation

$$\sqrt{\mu}\Big(2\rho_1\mu^\varsigma/3+\nu\Big)=d.\tag{B.4}$$

Obviously, for any $\mu\in(0,\mu_0]$, $\nu\sqrt{\mu}<d$, so that $\delta_\xi<\delta_x+d$ and $D_\xi\subseteq D'$. We will also assume that

$$\gamma_1>\bar\gamma_1(\mu)=\tfrac{3\nu}{\rho_1\mu^{\varsigma+1}}.\tag{B.5}$$

Such a choice of γ_1 will be motivated in Step 2.
Next, we take

$$\varepsilon_0(\gamma_1,\mu)=\tfrac{1}{\gamma_1}\min\Big\{1,\tfrac{\rho_1\mu^\varsigma\sqrt{\mu}}{3M_f^2c_u^2}\Big\},\tag{B.6}$$

and observe that $\varepsilon_0(\gamma_1,\mu)\le\frac{\rho_1\mu^{\varsigma+1}}{3\nu}$ because of (B.5). From (B.3) and (B.6) we obtain that, for each $\mu\in(0,\mu_0]$, $\gamma_1\in(\bar\gamma_1,\infty)$, $\varepsilon\in(0,\varepsilon_0(\gamma_1,\mu)]$, and for any $x(0)\in D_x$, $\xi(0)\in\overline{B_{\rho_0}(x(0))}$, if $\|\xi(t)-\xi(0)\|\le\frac{\nu\varepsilon}{\sqrt{\mu}}$ with $t\in[0,\varepsilon]$ then

$$\begin{aligned}\|x(t)-x(0)\|&\le M_fc_u\sqrt{\gamma_1\varepsilon\|x(0)-\xi(0)\|}\le M_fc_u\sqrt{\gamma_1\varepsilon\rho_0}\\&\le M_fc_u\sqrt{\frac{\gamma_1\varepsilon\rho_1\mu^\varsigma\sqrt{\mu}}{3}}\le M_fc_u\sqrt{\frac{\gamma_1\varepsilon\rho_1\mu^\varsigma\sqrt{\mu}}{3}}\le\frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3},\\\|x(t)-\xi(t)\|&\le\|x(t)-x^0\|+\|x^0-\xi^0\|+\|\xi(t)-\xi^0\|\\&\le M_fc_u\sqrt{\frac{\gamma_1\varepsilon\rho_1\mu^\varsigma\sqrt{\mu}}{3}}+\frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}+\frac{\nu\varepsilon}{\sqrt{\mu}}\le\rho_1\mu^\varsigma\sqrt{\mu}.\end{aligned}$$

If, additionally, $\|\xi^0-\xi^*\|\in D_\xi$ then

$$\begin{aligned}\|x(t)-x^*\|&\le\|x(t)-x^0\|+\|x^0-\xi^0\|+\|\xi^0-\xi^*\|\\&\le\frac{2\rho_1\mu^\varsigma\sqrt{\mu}}{3}+\delta_\xi\le\mathrm{dist}(x^*,\partial D').\end{aligned}$$

This proves the following intermediate statement.

Statement 1. For any $\mu\in(0,\mu_0]$, $\gamma_1\in(\bar\gamma_1,\infty)$, $\varepsilon\in(0,\varepsilon_0(\gamma_1,\mu)]$, $x^0\in D_x$, the π_ε-solutions of system (5) with the initial conditions $x(0)=x^0,\xi(0)=\xi^0$ satisfy the following property:

$$\|x^0-\xi^0\|\le\frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}\Rightarrow\|x(t)-\xi(t)\|\le\rho_1\mu^\varsigma\sqrt{\mu}\ \textit{for}\ t\in[0,\varepsilon].$$

Furthermore, if $\xi^0\in D_\xi\subseteq D'$ then $x(t)$ is well-defined in D' for $t\in[0,\mu]$.

Step 2. *Our next goal is to ensure that the x-component of the π_ε solution of system (5) is in a sufficiently small neighborhood of the ξ-component.*

For this, we apply Lemma A.1. Namely, assume that $x(t)\in D_x$ for $t\in[0,\varepsilon]$, $\xi(0)=\xi^0\in B_{\rho_1}(x^0)$. Then

$$\|\xi(t)-\xi(0)\|\le\frac{\nu\varepsilon}{\sqrt{\mu}},\tag{B.7}$$

and the π_ε-solution $x(t)$ of system (5) with controls (6) can be represented my means of the Chen–Fliess series:

$$x(\varepsilon)=x^0-\varepsilon\gamma_1(x^0-\xi^0)+R_1(\varepsilon)+R_2(x^0,\xi^0,\varepsilon),\tag{B.8}$$

where $R_1(\varepsilon)$ is defined from Lemma A.1, and

$$\begin{aligned}R_2(x^0,\xi^0,\varepsilon)&=\varepsilon^{3/2}\sum_{j_1\in S_1}\sum_{j_2=1}^{m}[f_{j_1},f_{j_2}](x^0)a_{j_1}(x^0,\xi^0)\\&\times\sum_{q:(q,j_2)\in S_2}\sqrt{\frac{|a_{qj_2}(x^0,\xi^0)|}{\pi K_{qj_2}}}\\&+\frac{\varepsilon^2}{2}\sum_{j_1,j_2\in S_1}L_{f_{j_2}}f_{j_1}(x^0)a_{j_1}(x^0,\xi^0)a_{j_2}(x^0,\xi^0).\end{aligned}$$

Denote $R(x^0,\xi^0,\varepsilon)=R_1(\varepsilon)+R_2(x^0,\xi^0,\varepsilon)$. Using (B.1) and notations from (B.2), we get $\|R_1(\varepsilon)\|\le M_{3f}c_u^3\big(\varepsilon\|x^0-\xi^0\|\big)^{3/2}$ for all $t\in[0,\varepsilon]$, and

$$\|R(x^0,\xi^0,\varepsilon)\|\le\zeta_1\big(\varepsilon\|x^0-\xi^0\|\big)^{3/2},\tag{B.9}$$

$\zeta_1=M_{3f}c_u^3+\frac{M_{2f}}{2}\sqrt{\nu\varsigma\alpha}(\gamma_1\alpha)^{3/2}+2(\gamma_1\alpha)^{3/2}M_{2f}\sqrt{|S_1|}\times\sum_{j_1=1}^{m}\Big(\sum_{(j_2,j_1)\in S_2}\kappa_{j_2j_1}^{-2/3}\Big)^{3/4}$. Combining (B.9), (B.7), and (B.8), we come to the following estimate:

$$\|x(\varepsilon)-\xi(\varepsilon)\|\le(1-\varepsilon\gamma_1)\|x^0-\xi^0\|+\zeta_1\big(\varepsilon\|x^0-\xi^0\|\big)^{3/2}+\frac{\nu\varepsilon}{\sqrt{\mu}}.$$

For any $\gamma_1>\bar\gamma_1$, let $\lambda_1\in[\bar\gamma_1,\gamma_1)$ and define

$$\varepsilon_1(\gamma_1,\mu)=\min\Big\{\varepsilon_0(\mu),\Big(\frac{\gamma_1-\lambda_1}{\zeta_1\sqrt{\delta_x}}\Big)^2\Big\}.\tag{B.10}$$

Recall that $\varepsilon\lambda_1<\varepsilon\gamma_1<1$. Then, for any $\varepsilon\in(0,\varepsilon_1(\gamma_1,\mu_1))$,

$$\|x(\varepsilon)-\xi(\varepsilon)\|<(1-\varepsilon\bar\gamma_1)\|x^0-\xi^0\|+\frac{\nu\varepsilon}{\sqrt{\mu}}.$$

Recall that $\bar\gamma_1$ is given by (B.5), which implies $\frac{\nu\varepsilon}{\sqrt{\mu}}=\frac{\gamma_1\rho_1\mu^\varsigma\sqrt{\mu}}{3}$. This together with Statement 1 gives us the next intermediate result.

Statement 2. Assume that $x(t)\in D'$ for all $t\in[0,\varepsilon_0]$, $x(0)\in D_x$. Then, for any $\mu\in(0,\mu_0]$, $\gamma_1\in(\bar\gamma_1,\infty)$, $\varepsilon\in(0,\varepsilon_1(\gamma_1,\mu)]$, the following properties hold:

$$\textit{if}\ \|x^0-\xi^0\|\le\frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}\ \textit{then}\ \|x(\varepsilon)-\xi(\varepsilon)\|\le\frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3},$$
$$\textit{and}\ \|x(t)-\xi(t)\|\le\rho_1\mu^\varsigma\sqrt{\mu}\ \textit{for all}\ t\in[0,2\varepsilon].$$

Step 3. Now let us put $x(0)=x^0=\xi^0=\xi(0)$, $x^0\in D_x$. Then $x(t)\equiv x^0\in D_x$ for all $t\in[0,\varepsilon]$, and

$$\|\xi(t)-\xi^0\|\le\|\xi(t)-\xi^0\|+\|\xi^0-\xi^*\|\le\nu\sqrt{\mu}+\delta_x=\delta_\xi,$$

i.e. $\xi(t) \in D_\xi \subset D'$ for $t \in [0,\varepsilon]$. Besides, Statement 2 implies $\|x(\varepsilon)-\xi(\varepsilon)\| < \frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}$.
From Statements 1 and 2, the x-component of the π_ε-solution of system (5) is also well-defined in D' for $t \in [\varepsilon, 2\varepsilon]$. Again, it is easy to see that $\|\xi(t)-\xi^0\| \le \nu\sqrt{\mu}+\delta_x$ for $t \in [0,2\varepsilon]$, i.e. $\xi(2\varepsilon) \in D_\xi$ and $\|x(2\varepsilon)-\xi(2\varepsilon)\| < \frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}$. Without loss of generality, we may assume that $\frac{\mu}{\varepsilon} = \mathcal{N}_1$, with some $\mathcal{N}_1 \in \mathbb{N}$. Repeating Steps 1–2 until $t = \mathcal{N}\varepsilon$, we come to the following statement.

Statement 3. For any $\mu \in (0,\mu_0]$, $\gamma_1 \in (\bar\gamma_1,\infty)$, $\varepsilon \in (0,\varepsilon_1(\gamma_1,\mu)]$, the π_ε-solutions $(x(t),\xi(t))$ of system (5) with the initial conditions $x(0)=\xi(0) \in D_x$ are well-defined in $D'\times D'$ for all $t \in [0,(\mathcal{N}_1+1)\varepsilon]$, $\|x(t)-\xi(t)\| \le \rho_1\mu^\varsigma\sqrt{\mu}$ for all $t \in [0,\mu]$, $\|x(\mu)-\xi(\mu)\| \le \frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}$.

Thus, for any $\mu \in (0,\mu_0]$, we can take $\gamma_1(\mu)$, $\varepsilon(\gamma_1(\mu),\mu)$, such that $x(t),\xi(t) \in D'$ for $t \in [0,\mu]$. In the next steps, we will find sufficiently small μ independently on ε and γ_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 μ.
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–Fliess series expansion of the ξ-component of solution of system (5) on the interval $[0,\mu]$:

$$\xi(\mu) = \xi^0 - \mu\gamma_2\nabla J(\xi^0) + R_3(\mu), \tag{B.11}$$

where

$$R_3(\mu) = \sum_{j=1}^{2n}\int_0^\mu \Big(g_j\circ J(x(s_1)) - g_j\circ J(\xi(s_1))\Big)e_j v_j^\mu(s_1)ds_1$$
$$+\sum_{j_1,j_2=1}^{2n}\int_0^\mu\int_0^{s_1} L_{e_{j_2}\left(g_{j_2}\circ J(x(s_2)) - g_{j_2}\circ J(\xi(s_2))\right)}$$
$$\times g_{j_1}\circ J(\xi(s_2))e_{j_1}v_{j_2}^\mu(s_2)v_{j_1}^\mu(s_1)ds_2ds_1$$
$$+\sum_{j_1,j_2,j_3=1}^{2n}\int_0^\mu\int_0^{s_1}\int_0^{s_2} v_{j_3}^\mu(s_3)v_{j_3}^\mu(s_2)v_{j_1}^\mu(s_1)$$
$$\times L_{e_{j_3}g_{j_3}\circ J(x(s_3))}L_{e_{j_2}g_{j_2}\circ J(\xi(s_3))}g_{j_1}\circ J(\xi(s_3))e_{j_1}ds_3ds_2ds_1.$$

Under the assumptions of Theorem 1, we conclude that

$$\|R_3(\mu)\| \le c_w\sqrt{\mu}(L_g+\sqrt{\mu}L_{2g}c_w)\max_{0\le t\le\mu}\|x(t)-\xi(t)\| + \mu^{3/2}M_{3g}.$$

Thus, applying Statement 3 we get $\|R_3(\mu)\| \le \zeta_2\mu^{1+\tilde\varsigma}$, where $\tilde\varsigma = \min\{\varsigma, 1/2\}$, $\zeta_2 = c_w\mu^{\max\{0,\varsigma-1/2\}}\rho_1(L_g + \sqrt{\mu}L_{2g}c_w) + \mu^{\max\{0,1/2-\varsigma\}}M_{3g}$.
Using Taylor's formula for the function $J(\xi)$,

$$J(\xi(t)) = J(\xi^0) + \nabla J(\xi^0)(\xi(t)-\xi^0) + \frac{1}{2}\sum_{i,j=1}^{2n}\frac{\partial^2 J(x)}{\partial x_i\partial x_j}\Big|_{x=\xi^0+\theta\xi(t)}(\xi_i-\xi_i^0)(\xi_j-\xi_j^0),$$

and exploiting (9), we obtain

$$\begin{aligned}J(\xi(\mu)) &\le J(\xi^0) - \mu\gamma_2\sigma_{21}J(\xi^0) + \mu^{1+\varsigma}\zeta_2\sqrt{\sigma_{22}J(\xi^0)}\\ &\quad+\sigma_3\left(\mu^2\gamma_2^2\sigma_{22}^2J(\xi^0)+\zeta_2^2\mu^{2+2\varsigma}\right)\\ &= J(\xi^0)\Big(1-\mu\gamma_2\big(\sigma_{21}-\mu\gamma_2\sigma_3\sigma_{22}^2\big)\Big)\\ &\quad+\mu^{1+\varsigma}\zeta_2\Big(\sqrt{\sigma_{22}J(\xi^0)}+\sigma_3\zeta_2\mu^{1+\varsigma}\Big).\end{aligned}$$

Let $\mathcal{L}_c = \{x \in D : J(x) \le c\}$, $c_J = \sigma_{11}\delta_x^2$. Then

$$\overline{B_\delta(x^*)} \subseteq \mathcal{L}_{c_J} \subseteq D_x. \tag{B.12}$$

For any $\rho_2 \in (0,c_J]$, $\lambda_2 \in (0,\gamma_2\sigma_{21})$, we define

$$\mu_1 = \min\{\mu_0, 1/\lambda_2, \hat\mu_1\}, \tag{B.13}$$

where $\hat\mu_1$ is the smallest positive root of the equation

$$\rho_2\mu\gamma_2\sigma_3\sigma_{22}^2 + \mu^\varsigma\zeta_2\left(\sqrt{\sigma_{22}\rho_2}+\sigma_3\zeta_2\mu^{1+\varsigma}\right) = \rho_2(\gamma_2\sigma_{21}-\lambda_2).$$

Then, for any $\mu \in (0,\mu_1)$, the following two scenarios are possible:
S1) If $J(\xi^0)\le\rho_2$ then $J(\xi(\mu))\le\rho_2(1-\mu\lambda_2)<\rho_2$. In this case, $\xi(\mu) \in D_x$, Additionally, Statement 3 implies that $\|x(\mu)-\xi(\mu)\| \le \frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}$. Repeating the above argumentation, we get $\xi(N\mu) \in D_x$ for all natural numbers N.
S2) If $J(\xi^0) > \rho_2$ then $J(\xi(\mu)) < J(\xi^0)(1-\mu\lambda_2) < J(\xi^0)$.
Consider S2). If $\xi^0 = x^0 \in \overline{B_\delta(x^*)}$ then $\xi(\mu) \in \mathcal{L}_{c_J} \subseteq D_x$. Again, Statement 3 gives $\|x(\mu)-\xi(\mu)\| \le \frac{\rho_1\mu^\varsigma\sqrt{\mu}}{3}$. Thus, we may repeat all the steps for $t \in [\mu, 2\mu]$.
Summarizing all the above, we arrive at the following conclusion: there exists an $\mathcal{N}_2 \in \mathbb{N}\cup\{0\}$ such that

$$J(\xi(t)) \le J(\xi^0)e^{-\lambda_2 t} \text{ for } t = 0,\mu,\dots,(\mathcal{N}_2-1)\mu,$$
$$J(\xi(t)) \le \rho_2 \text{ for } t = \mathcal{N}_2\mu, (\mathcal{N}_2+1)\mu,\dots.$$

Consequently, $\|\xi(t)-x^*\| \le \sqrt{\frac{\sigma_{12}}{\sigma_{11}}}\|x^0-x^*\|e^{-\lambda_2 t}$ for $t = 0,\mu,\dots,(\mathcal{N}_2-1)\mu$, $\|\xi(\mathcal{N}_2\mu)-x^*\| \le \sqrt{\frac{\rho_2}{\sigma_{11}}} \le \delta_x$ for $t = \mathcal{N}_2\mu, (\mathcal{N}_2+1)\mu,\dots$. For an arbitrary $t \in [0,\mathcal{N}_2\mu]$, we denote the integer part of $\frac{t}{\mu}$ as $\left[\frac{t}{\mu}\right]$ and observe that $0 < t - \left[\frac{t}{\mu}\right]\mu < \mu$. Then

$$\|\xi(t)-x^*\| \le \left\|\xi\left(\left[\frac{t}{\mu}\right]\mu\right)-x^*\right\| + \left\|\xi(t)-\xi\left[\frac{t}{\mu}\right]\right\|$$
$$\le \sqrt{\frac{\sigma_{12}}{\sigma_{11}}}\|x^0-x^*\|e^{-\lambda_2\left[\frac{t}{\mu}\right]\mu} + \nu\sqrt{\mu} \le \beta\|x^0-x^*\|e^{-\lambda_2 t}+\nu\sqrt{\mu},$$

where $\beta = \sqrt{\frac{\sigma_{12}}{\sigma_{11}}}e^{\lambda_2\mu}$. This yields the following result.

Statement 4. For any $\mu \in (0,\mu_1]$, $\gamma_1 \in (\bar\gamma_1,\infty)$, $\varepsilon \in (0,\varepsilon_0(\gamma_1,\mu)]$, the π_ε-solutions $(x(t),\xi(t))$ of system (5) with the initial conditions $x(0)=\xi(0) \in D_x$ are well-defined in $D'\times D'$ for all $t \in [0,\infty)$, and the following estimates hold:

$$\|\xi(t)-x^*\| \le \beta\|x^0-x^*\|e^{-\lambda t}+\nu\sqrt{\mu} \text{ for } t \in [0,\mathcal{N}_2\mu],$$
$$\|\xi(t)-x^*\| \le \sqrt{\frac{\rho_2}{\sigma_{11}}}+\nu\sqrt{\mu} \text{ for } t \in [\mathcal{N}_2\mu,\infty).$$

Furthermore, $\|x(t)-\xi(t)\| \le \rho_1\mu^\varsigma\sqrt{\mu}$ for all $t \in [0,\infty)$.

Step 5. Finally, we estimate $\|x(t)-x^*\|$ for $t \in [0,\infty)$. Applying the triangle inequality together with Statement 4, we get the following:

$$\|x(t)-x^*\| \le \beta\|x^0-x^*\|e^{-\lambda t}+\rho_1\mu^\varsigma\sqrt{\mu}+\nu\sqrt{\mu} \text{ for } t \in [0,\mathcal{N}_2\mu],$$
$$\|x(t)-x^*\| \le \rho_1\mu^\varsigma\sqrt{\mu}+\sqrt{\frac{\rho_2}{\sigma_{11}}}+\nu\sqrt{\mu} \text{ for } t \in [\mathcal{N}_2\mu,\infty).$$

Since ρ_1, ρ_2 are arbitrary and μ can be chosen small enough, the above inequalities imply the assertion of Theorem 1. In particular, for an arbitrary $\rho > 0$, one can take $\rho_1 > 0$ and $\mu > 0$ such that

$$\rho_1\mu^\varsigma\sqrt{\mu}+\nu\sqrt{\mu} \le \frac{\rho}{2}, \tag{B.14}$$

and $\rho_2 \le \frac{1}{4}\rho^2\sigma_{11}$. Then $\|x(t)-x^*\| \le \beta\|x^0-x^*\|e^{-\lambda t}+\rho$ for all $t \in [0,\infty)$. Note that the choice of μ does not depend on ε,η, and the choice of γ_1 does not depend on ε. Namely, given $\delta,\rho,\rho_1,\rho_2$, one can choose a $\bar\mu > 0$ satisfying (B.4), (B.13) and (B.14), and take any $\hat\mu \in (0,\bar\mu]$. The next step is to determine $\bar\gamma_1(\hat\mu)$ satisfying (B.5), and take any $\hat\gamma_1 \in (\bar\gamma_1,\infty)$. Finally, $\bar\varepsilon(\hat\gamma_1,\hat\mu)$ has to be specified according to (B.6) and (B.10).