Nonsmooth Extremum Seeking Control With User-Prescribed Fixed-Time Convergence

Jorge I. Poveda and Miroslav Krstić, Fellow, IEEE

Abstract—This article introduces a new class of nonsmooth extremum seeking controllers (ESCs) with convergence bounds given by class-$\mathcal{KL}$ functions that have a uniformly bounded settling time. These ESCs are characterized by nominal average systems that render uniformly globally fixed-time stable (UGFxTS), the set of minimizers of the response map of a stable nonlinear plant. Given that, under suitable tuning of the parameters of the controllers, the ESCs inherit the convergence properties of their average systems, the proposed dynamics can achieve a better transient performance compared to the traditional ESCs based on gradient descent or Newton flows. Moreover, for the case when the plant is a static map, the convergence time of the proposed algorithms can be prescribed a priori by the users for all initial conditions without the need of retuning the gain of the learning dynamics of the ESC. Since autonomous feedback controllers with fixed-time convergence properties are necessarily non-Lipschitz continuous, standard averaging and singular perturbation tools, traditionally used in ESC, are not applicable anymore. We address this issue by using averaging and singular perturbation tools for nonsmooth and set-valued systems, which further allows us to consider ESCs modeled by discontinuous vector fields that are typical in fixed-time and finite-time optimization problems.

Index Terms—Adaptive control, extremum seeking, optimization.

I. INTRODUCTION

Extremum seeking control (ESC) has shown to be a powerful technique for the solution of model-free optimization problems in dynamical systems [1]–[4]. Stability, convergence, and robustness guarantees for different types of constrained and unconstrained smooth ESCs have been extensively studied in the literature [5]–[8]. Recently, ESC has also been extended to nonsmooth and hybrid settings that are able to overcome some of the intrinsic limitations of smooth feedback controllers [9], [10]. An early use of nonsmooth ESC can also be found in [11, Sec. 5].

In this article, we extend and generalize some of these results by introducing a new class of nonsmooth ESCs that have stability properties characterized by class-$\mathcal{KL}$ functions with the “fixed-time convergence property,” namely, there exists a continuous settling time function $r \mapsto T(r)$ and a positive number $T^* > 0$ such that $\lim_{r \to T(r)} \beta(r, s) = 0$ and $T(r) < T^*$, for all $r \geq 0$. Functions with this attribute are also said to be of class-$\mathcal{KL}$ [12], and they characterize the convergence properties of systems whose solutions converge to a particular set in a finite time $T(r)$ that can be upper bounded by a constant $T^*$ that is independent of the distance $r$ of the initial conditions to the set. This powerful property has motivated the development of new algorithms in the context of regulation, optimization, and estimation problems, see [13]–[16] and [17]. Nevertheless, ESCs with fixed-time convergence properties remain completely unexplored in the literature, and, as we will show in this article, they have the potential of inducing dramatic improvements in the transient performance of the closed-loop system for certain classes of plants having response maps with strong monotonicity properties.

In order to design ESCs with $\mathcal{KL}$ convergence bounds, which are called fixed-time extremum seeking controllers (FxTESCs), our starting point is the averaging-based paradigm considered in [3], [5], [18], and [19], which, due to its modular approach, can accommodate different types of optimization algorithms, making a direct connection between the bounds that characterize the convergence of the trajectories of the nominal average system and the actual control signal generated by the ESCs. However, since most of the existing results in the literature rely on averaging and singular perturbation tools for Lipschitz continuous systems, they are not suitable for the design and analysis of ESCs with $\mathcal{KL}$ convergence bounds. Instead, in this article, we use averaging and singular perturbation tools for nonsmooth and set-valued systems [9], [20]–[22], which allows us to additionally consider ESCs based on differential inclusions that may be related to discontinuous optimization algorithms. Moreover, we also develop Newton-like fixed-time ESCs that remove from the convergence bound $T^*$ the dependence on the unknown parameters of the Hessian matrix of the response map. The best of our knowledge, the ESCs for dynamical systems presented in this article are the first that have convergence bounds characterized by class-$\mathcal{KL}$ functions. For static maps and specific continuous ESC algorithms, preliminary results with sketches of the proofs were presented in the conference papers [1] and [2]. In contrast to these works, this article addresses the general ESC problem in dynamic plants, considers a general family of possibly discontinuous ESCs (which subsume those considered in [1] and [2]), derives tighter convergence bounds for the algorithms, new auxiliary averaging results for nonsmooth systems, and also presents the complete stability analysis. Moreover, in contrast to [1] and [2], the results of this article are also applicable to ESC algorithms with nominal average systems having only finite-time convergence properties, thus, addressing another existing gap in the literature of ESC.

The rest of this article is organized as follows. Section II introduces the notation and preliminaries. Section III presents the problem statement. Section IV presents the main results for gradient-based ESCs,
Consider a nonlinear dynamic plant with input \( u \in \mathbb{R}^n \), output \( y \in \mathbb{R} \), and state \( x \in \Xi \subset \mathbb{R}^p \), modeled by the equations
\[
\dot{x} = f(x, u), \quad y = h(x, u)
\] (3)
where \( f \) is locally Lipschitz, and \( h \) is continuously differentiable. In this article, we assume that the operational space \( \Xi \) for the states of the plant (3) is closed and bounded. In practice, boundedness of \( \Xi \) can be related to the physical limitations of the plant, or to operational sets that are rendered forward invariant by using internal feedback controllers that implement mechanisms such as Lipschitz projections or barrier functions. Compactness of \( \Xi \) is also guaranteed when the dynamics (3) have the bounded-input bounded-state property and \( u \) is uniformly bounded (as will be the case in our results).

In order to have a well-defined ESC problem, we also make the following standard stability assumption on system (3).

Assumption 1: There exists a continuous function \( \ell_x : \mathbb{R}^n \to \mathbb{R}^p \), such that for each compact set \( K_u \subset \mathbb{R}^n \) the constrained dynamical system \( (x, u) \in \Xi \times K_u, \dot{x} = f(x, u), \dot{u} = 0 \), renders UGAS the compact set \( M_{K_u} := \{(x, u) \in \Xi \times K_u : x = \ell(x, u)\} \).

The existence of a UGAS quasi-steady-state manifold \( \ell_x \) for the plant (3) is a standard assumption in ESC, see [3], [5], and [6].

The response map of the plant (3) is defined as \( \phi(u) := h(\ell_x(u), u) \), which is assumed to satisfy the following.

Assumption 2: The function \( u \mapsto \phi(u) \) is twice continuously differentiable, and the set \( A_{\phi} := \text{argmin}_{u \in \mathbb{R}^n} \phi(u) \) is nonempty and compact.

Based on these assumptions, the ESC problem that we study in this article consists on regulating the input \( u \) of system (3) toward the set \( A_{\phi} \), by using only output measurements of \( y \), without any knowledge of the mathematical form of \( f, h \), or \( \phi \).

### A. Transient Limitations of Gradient-Descent-Based ES

To motivate the FxTESCs considered in this article, we first review the convergence properties of the standard gradient descent-based extremum seeking controller (GDESC) studied in [3], [6], and [5]. This controller is characterized by the feedback law and the dynamics

\[
u := a + \nabla \mu, \quad \dot{\mu} = -k_3 \xi \]

where \( k_3 := \varepsilon_0 k \) and \(( k, \varepsilon_0, a) \) are positive tunable parameters. The auxiliary states \(( \xi, \mu) \) of (4) have dynamics

\[
\dot{\xi} = -k_2 (\xi - F_G(y, \mu)) \quad \dot{\mu} = -k_3 R_\theta \mu, \quad \mu \in \mathbb{T}^n
\]

(5)

The mapping \( F_G \) in (5) is defined as \( F_G(y, \mu) := yM(\mu) \), with \( M(\mu) := \frac{1}{\mu} \nabla \mu \). We study the transient performance of this controller under the following additional assumption on the response map \( \phi \).

Assumption 3: There exists \( k > 0 \) such that for all \( u \in \mathbb{R}^n \),

\[
|\nabla \phi(u)|^2 \leq 2(\nu(\phi(u) - \phi(u')))^2 \]

for all \( u \in \mathbb{R}^n \). Moreover, \( \nabla \phi \) is \( L \)-globally Lipschitz, and \( A_{\phi} = \{u^*\} \).

Typical examples of response maps that satisfy the inequality of Assumption 3 include strongly convex functions, such as positive definite quadratic functions, which are ubiquitous in the literature of ESC, e.g., [3], [8], and [24]. However, Assumptions 1–3 do not necessarily ask for convexity of \( h \) or linearity of (3). Instead, response maps \( \phi \) satisfying Assumptions 2 and 3 can be generated by different classes of plants (3) having linear/nonlinear dynamics \( f \) and nonlinear/linear outputs \( h \). Nevertheless, it is important to note that in the context of ESC, the goal is to achieve model-free optimization; thus, the mathematical forms of the functions \(( f, h )\) are generally unknown.

To guarantee convergence to \( A_{\phi} \), a time-scale separation is induced between the dynamics of the plant (3) and the dynamics of the controller (4), (5). In particular, by introducing the new time scale \( \tau := t\varepsilon_0 \), and by using the definition of \( D \), the dynamics of the closed-loop system
can be written as follows:

\[
\begin{align*}
\frac{d\hat{u}}{d\tau} &= -kF_u(\hat{u}, \xi), \\
\frac{d\xi}{d\tau} &= -\frac{1}{\varepsilon_2} \left( \xi - F_C(y, \mu) \right) \tag{7a}
\end{align*}
\]

\[
\begin{align*}
\frac{d\mu}{d\tau} &= -2\pi R\alpha \mu, \\
\varepsilon_1 \frac{d\tilde{\tau}}{d\tau} &= f(x, \hat{u} + aD\mu) \tag{7b}
\end{align*}
\]

where according to (4), the learning dynamics of the GDESC are characterized by the mapping \( F_u(\hat{u}, \xi) = \xi \). When \( \varepsilon_0 \) is sufficiently small, system (7) is a singularly perturbed system with fast “boundary layer dynamics” corresponding to the dynamics of the plant, and slow “reduced dynamics” corresponding to the dynamics of the states \((\hat{u}, \xi, \mu)\). By Assumption 1, the plant has a well-defined quasi-steady-state manifold \( \xi(\cdot) \). Thus, the reduced dynamics are as follows:

\[
\frac{d\hat{u}}{d\tau} = -kF_u(\hat{u}, \xi), \quad \frac{d\xi}{d\tau} = -\frac{1}{\varepsilon_2} (\xi - F_C(\phi(\hat{u} + aD\mu), \mu))
\]

\[
\varepsilon_1 \frac{d\tilde{\tau}}{d\tau} = f(x, \hat{u} + aD\mu), \quad \mu \in \mathbb{T}^n
\]

where we used the definitions of \( \phi \) and \( F_C \). Let \( \tilde{\mu} := D\mu \). For small values of \( \alpha \), we can perform a Taylor expansion of \( \phi(\hat{u} + aD\mu) \) around the point \( \hat{u} \), leading to \( \phi(\hat{u} + a\tilde{\mu}) = \phi(\hat{u}) + a\tilde{\mu} \nabla \phi(\hat{u}) \nabla (a^2) \). By using the definitions of \( \tilde{\mu} \), the fact that the solutions of the oscillator are given by sinusoids with unitary amplitude, and [2, 6, 6], we can average the dynamics of the states \((\hat{u}, \xi)\) along the trajectories \( \tilde{\mu} \). Since \( F_u \) is independent of \( \tilde{\mu} \), the resulting average system is

\[
\frac{d\hat{u}}{d\tau} = -kF_u(\hat{u}, \xi), \quad \frac{d\xi}{d\tau} = -\xi + \tilde{F}_C(\hat{u}, \alpha)
\]

where the function \( \tilde{F}_C \) is given by

\[
\tilde{F}_C(\hat{u}, \alpha) := \nabla \phi(\hat{u}) + \alpha(1)
\]

For each \( \varepsilon_2 > 0 \), system (8) is a \( O(\alpha) \)-perturbed version of a nominal average system with \( O(\alpha) = 0 \). In turn, since \( \varepsilon_2 \ll 1/k \), this nominal average system is also a singularly perturbed system with exponentially stable boundary layer \( \xi^A \)-dynamics with equilibrium point \( \xi^A = \nabla \phi(\hat{u}) \), and reduced nominal average dynamics with state \( \hat{u} \), given by

\[
\frac{d\hat{u}}{d\tau} = -kF_u(\hat{u}, \xi), \quad \frac{d\hat{\xi}}{d\tau} = -\xi + \tilde{F}_C(\hat{u}, \alpha)
\]

which is a gradient descent (GD) flow with gain \( k > 0 \). Therefore, by using the Lyapunov function \( V(\hat{u}, \xi) = \phi(\hat{u}) - \phi(u) \), it can be shown that under Assumption 3, all solutions of the GD flow (10) satisfy the bound \( |\hat{u}(\tau)|_{\alpha} \leq \sqrt{\frac{L}{\varepsilon_2}} \hat{u}(0)|_{\alpha} e^{-\kappa \tau} \), for all \( \tau \geq 0 \). This establishes a USGAR result for system (10) with an exponential class-\( KL \) function \( \beta(r, s) = \sqrt{\frac{\kappa}{2}} r e^{-ks} \). After some manipulations (see proof of Th. 1), we can now repeatedly apply singular perturbation and averaging theory [18, Th. 1], as well as structural robustness results for smooth ODEs [25, Lemma 7.20], to conclude that for each pair \( \Delta > \nu > 0 \) there exists \( \varepsilon_2^* > 0 \) such that when each \( \varepsilon_2 \in (0, \varepsilon_2) \), in the sense of the closed-loop system (7) with an ESC implementing learning dynamics characterized by the function (13), the mapping \( F_v \) is not Lipschitz continuous without further conditions on \( (\alpha_1, \alpha_2) \). However, even in cases when \( F_v \) is discontinuous at the point \( \xi = 0 \), the existence of generalized solutions in the sense of Krasovskii can always be guaranteed [25, Lemma 5.26]. Functions of this form have been extensively studied in the literature of fixed-time stabilization by using the notion of homogeneity (in the bilimit), see for instance [26, Sec. 5.1], [14, Sec. 1], [23, Sec. 4-A], [27, Ex. 1], or [28, Lemma 2.1]. However, in the context of ESC, they remained unexplored.

The closed-loop system of Fig. 1 can be studied by following similar steps as in the previous section. In this case, instead of (10), the reduced nominal average dynamics are given by

\[
\frac{d\hat{u}}{d\tau} = -k \left( \frac{\nabla \phi(\hat{u})}{|\nabla \phi(\hat{u})|^{\alpha_1}} + \frac{\nabla \phi(\hat{u})}{|\nabla \phi(\hat{u})|^{\alpha_2}} \right)
\]

which can be analyzed by using the smooth Lyapunov function \( V_G(\hat{u}) = \frac{1}{2} ((\phi(\hat{u}) - \phi(u))^2) \), which, under Assumption 3, is radially unbounded and positive definite with respect to \( u \). The time derivative of \( V_G \) along the solutions of system (14) satisfies

\[
\Delta V_G(\hat{u}(\tau)/\nu) \leq -k (c_1 V_G(\hat{u}(\tau)^{1/2}) + c_2 V_G(\hat{u}(\tau)^{1/2}) \text{ for all } \hat{u} \neq u^*, \text{ where}
\]

\[
\varepsilon_2 \frac{d\hat{\tau}}{d\tau} = f(x, \hat{u} + aD\mu)
\]

\[
\frac{d\hat{u}}{d\tau} = -k \left( \frac{\nabla \phi(\hat{u})}{|\nabla \phi(\hat{u})|^{\alpha_1}} + \frac{\nabla \phi(\hat{u})}{|\nabla \phi(\hat{u})|^{\alpha_2}} \right)
\]

which can be analyzed by using the smooth Lyapunov function \( V_G(\hat{u}) = \frac{1}{2} ((\phi(\hat{u}) - \phi(u))^2) \), which, under Assumption 3, is radially unbounded and positive definite with respect to \( u^* \). The time derivative of \( V_G \) along the solutions of system (14) satisfies

\[
\Delta V_G(\hat{u}(\tau)/\nu) \leq -k (c_1 V_G(\hat{u}(\tau)^{1/2}) + c_2 V_G(\hat{u}(\tau)^{1/2}) \text{ for all } \hat{u} \neq u^*, \text{ where}
\]

\[
\varepsilon_2 \frac{d\hat{\tau}}{d\tau} = f(x, \hat{u} + aD\mu)
\]

\[
\frac{d\hat{u}}{d\tau} = -k \left( \frac{\nabla \phi(\hat{u})}{|\nabla \phi(\hat{u})|^{\alpha_1}} + \frac{\nabla \phi(\hat{u})}{|\nabla \phi(\hat{u})|^{\alpha_2}} \right)
\]

which can be analyzed by using the smooth Lyapunov function \( V_G(\hat{u}) = \frac{1}{2} ((\phi(\hat{u}) - \phi(u))^2) \), which, under Assumption 3, is radially unbounded and positive definite with respect to \( u^* \). The time derivative of \( V_G \) along the solutions of system (14) satisfies

\[
\Delta V_G(\hat{u}(\tau)/\nu) \leq -k (c_1 V_G(\hat{u}(\tau)^{1/2}) + c_2 V_G(\hat{u}(\tau)^{1/2}) \text{ for all } \hat{u} \neq u^*, \text{ where}
\]
is UGFxTS for system (14). Moreover, if we set $\alpha_0 = -\alpha_1$, the assumptions of Lemma 1 hold with $\alpha = 2/\alpha_1$, and $\sqrt{d}b = 4k\kappa$, and we obtain the following estimate on the bound for the settling time $T_\alpha$ of (14):

$$T_\alpha := \frac{\pi}{2k\kappa_1\kappa}. \quad (15)$$

Finally, note that all the previous computations hold when $\xi$ and $\dot{u}$ are vectors in $\mathbb{R}^n$, and an explicit function $\beta_{G,k}(\cdot)$ can be computed by using Lemma 3 in the Appendix. Therefore, for the ESC of Fig. 1, the dynamics (14) can be seen as general fixed-time gradient (FxTG) flows in $\mathbb{R}^n$ with the UGFxTS property.

Remark 1: FxTG flows have been recently studied in the context of model-based optimization and fixed-time stabilization in [28, Lemma 2.1], [15], and [16]. Continuity of the function (13) can be guaranteed as in [1] and [15].

The key implication of the fixed-time bound (15), which is independent of the initial conditions of system (14), is that for positive values of $(k, \kappa)$ and $p = (\nu, \alpha_1)$, with $\alpha_1 \in (0, 1)$, the value of $\tau_\alpha$ obtained from (12) will be larger than $T_\alpha$ whenever $\dot{u}(0) \in \Omega_{G,\nu} := \{u_0 \in \mathbb{R}^n : |u_0|_{\alpha_1} > 0.5v\exp(\frac{\pi}{2l})\}$. If, for any pair $\alpha > 0$, the parameters $(\varepsilon_0, \varepsilon_1, a, \varepsilon_2)$ of the FxTESC shown in Fig. 1 can be selected such that the bound (11) holds with the KL function $\beta_{G,k}(\cdot)$ (a result that we will establish in the following section), then the convergence time of the FxTESC will outperform the convergence time of the GDESC for all initial conditions $u_0(0)$ in the set $\Omega_{G,\nu} \cap (A_\kappa + \Delta B)$, which is nonempty when $\Delta$ is sufficiently large or $\nu$ is sufficiently small.

Remark 2: Since the bound (15) is independent of the initial conditions of the system, a universal gain $k$ can now be used to induce a desired convergence time $T_\alpha$ via the class KL function $\beta_{G,k}(\cdot)$. This property is fundamentally different from the asymptotic or exponential (semiglobal practical) convergence properties of the smooth ESCs considered in [3], [5], [18], and [19]. Note, however, that the parameters $(\varepsilon_0, \varepsilon_1, a, \varepsilon_2)$ will still depend on $\nu$ and $\Delta$, since their role is to guarantee that the ESC approximates the behavior (on compact sets) of its reduced nominal average dynamics (14).

Example 1: In order to illustrate the previous discussion, let us consider a simple plant in $\mathbb{R}^2$ with $f(x, u) = 10 \times [-x_1 + u_1 - x_2 + u_2]^2$ and output $y = (x_1 - 1)^2 + (x_2 - 5)^2$. Since $f(x, u)$ describes a stable linear system that generates bounded states under bounded inputs, Assumption 1 holds. We set $\Delta = 100$, and we simulate the closed-loop system (7) using the following parameters: $a = 0.01, \varepsilon_0 = 4 \times 10^{-5}, \varepsilon_1 = 1 \times 10^{-4}, \varepsilon_2 = 5 \times 10^{-1}, \text{and } k = 0.02$, which satisfy the relations of (6). For the oscillator (2), we used $2\pi \theta_1 = 3.5, 2\pi \theta_2 = 4$, and $\varepsilon_0(0) = 0.1, 0, 1, 0^\top$. For the FxTESC with learning dynamics (13), we used $\alpha_1 = 0.5 = -\alpha_2$. Since in this case $\phi(u) = (u_1 - 1)^2 + (u_2 - 5)^2$, it follows that Assumption 3 holds with $\kappa = L = 2$. Using (15), we obtain $T_\alpha = 78.53$. Fig. 2 compares the behavior of the trajectories of the GDESC and the FxTESC. We emphasize that both algorithms used the same parameters $(a, k, \varepsilon_0, \varepsilon_1, \varepsilon_2)$. In the left figures, we have numerically approximated the reachable set of both algorithms from initial conditions with $\dot{u}(0) \in [-100,100] \times [-100,100]$, and $x(0) = (1, 0)^\top$ by running $1 \times 10^3$ simulations with random initializations on this set. The insets show that all solutions generated by the FxTESC converge to a $\nu$-neighborhood (with $\nu = 1 \times 10^{-7}$) of the optimal point $w^* = (1, 5)^\top$ before the time $T_\alpha$. On the other hand, the right plots show the time history of one solution of the GDESC and FxTESC, respectively, with identical initialization. As expected, and as shown by the dashed lines, the trajectories generated by both ESCs are almost identical to the trajectories of their reduced nominal average dynamics (10) and (14). In particular, as highlighted in the upper logarithmic plot, the trajectory generated by the FxTESC approximately inherits the “fixed-time convergence property” of system (14).

Remark 3 (Finite-Time Versus Fixed-Time Stability in ESC): In contrast to the property of UGFxTS, the property of finite-time stability is characterized by a class-$\mathcal{K}$ function that satisfies $\lim_{\tau \to \tau^*} \beta(r, s) = 0$, where $\tau^*$ is not necessarily uniformly bounded, see [29]. The ESC shown in Fig. 1 can induce this weaker property by using $\alpha_1 = \alpha_2 = 1$, which generates reduced nominal average dynamics given by the discontinuous flow $\dot{u}_n = -2k\nabla \phi | \nabla \phi |$, studied in [30] using generalized solutions in the context of differential inclusions. Under $\kappa$-strong convexity of $\phi$ and L-globally Lipschitz of $\nabla \phi$, the lower bound $\tau_{\kappa_L}$ of the convergence time of this system satisfies $k\tau_{\kappa_L}^* \leq (2\kappa)^{-1} L\|u_n(0)\|_{\alpha_1}$, which grows linearly with $|u_n(0)|_{\alpha_1}$. Thus, in the ESC case, the bound on $\tau_{\kappa_L}$ would also grow linearly with $\Delta$, which is a weaker property compared to the constant bound (15). Other optimization flows with finite-time convergence properties are presented in [30] and [17]. Note that when $\alpha_1 = \alpha_2 = 0$ in (13), the FxTESC of Fig. 1 reduces to the standard GDESC.

Next, we formalize and generalize the previous discussion by characterizing an entire family of FxTESCs.

IV. GRADIENT-BASED FIXED-TIME ES CONTROLLERS

Consider the closed-loop system (7) with general learning dynamics in (4) now modeled as follows:

$$\dot{\hat{u}} \in -k_1 F_u(\hat{u}, \xi) \quad (16)$$

where $F_u: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a set-valued map, and $k_1 = \varepsilon_0 k$.

Assumption 4: The set-valued mappings $F_u(\cdot, \cdot)$ and $F_\xi(\cdot) := F_u(\cdot, \nabla \phi(\cdot))$ satisfy the Basic Conditions.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on December 04,2021 at 01:29:14 UTC from IEEE Xplore. Restrictions apply.  

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on December 04,2021 at 01:29:14 UTC from IEEE Xplore. Restrictions apply.
The regularity properties of Assumption 4 are standard in the analysis of nonsmooth systems, and they also hold when $F_0$ and $F_G$ are single-valued continuous functions. However, by working with differential inclusions, we will be able to consider ESCs with learning dynamics that are not necessarily continuous. In this case, solutions must be understood in a generalized sense by considering the set-valued map $F_G^K(\hat{u}) := \bigcap_{k > 0} \mathfrak{K} F_G(\hat{u} + \delta k)$, where $\mathfrak{K} F_G$ stands for the closed convex hull, see [9, Sec. 6.1], for examples, in the context of ESC. When $F_G$ is LB, the set-valued map $F_G^K$ satisfies the Basic Conditions [25, Lemma 5.16].

The following stability assumption characterizes the ESCs considered in this article.

Assumption 5: For each $k > 0$, system $\dot{\hat{u}} = -k F_G(\hat{u})$ renders the set $A_0 \cap UFGXs$ with some $\beta_{G,k} \in K C^L$, with continuous settling time function satisfying $T_k(\tau) \leq T'_k$ for all $r \geq 0$.

Remark 4: Assumption 5 can be verified via Lyapunov functions (c.f. Lemma 1 or [26, Th. 5.8] for differential inclusions), or by studying the homogeneity properties in the bi-limit of the mapping $F_G$, see [14], [26], [31], and [15], [16] for different examples of gradient-based optimization dynamics $\dot{\hat{u}} \in -k F_G(\hat{u})$ that satisfy Assumption 5.

We are now ready to state the main result of this article.

Theorem 1: Suppose that Assumptions 1, 2, 4, and 5 hold. Then, $\forall k > 0$ and $\forall \Delta > 0$, $\exists \delta > 0$ such that $\forall \varepsilon_2 \in (0, \delta)$, $\exists \alpha^* > 0$ such that $\forall \varepsilon_1 \in (0, \alpha^*)$, $\exists \alpha > 0$ such that $\forall \varepsilon_0 \in (0, \epsilon_0)$, all solutions of (7) with learning dynamics (16), and $\hat{u}(0) \in \Delta$, $\xi(0) \leq \Delta$, the bound

$$\|u(\tau)\|_{A_0} \leq \beta_{G,k}(\hat{u}(0)\|A_0\|, \tau) + \nu, \quad \forall \tau \geq 0$$

and $\beta_{G,k}(\hat{u}(0)\|A_0\|, \tau) = 0$ for all $\tau \geq T_k'$.

Proof: Let $k$, $\Delta$, and $\nu$ be given. Without loss of generality we assume $\nu \in (0, 1)$. Let Assumption 5 generate the function $\beta_{G,k} \in K C^L$. We define the set $\bar{K} := \{u \in \mathbb{R}^n : \|u\|_{A_0} \leq \beta_{G,k}(\|y(A_0 u, 0)\| + 1)\}$. By construction, this set is compact since without loss of generality $\beta_{G,k}$ can be taken to be continuous or to be upper bounded by a continuous class $K C^L$ function [25, pp. 69]. Thus, there exists $M > 0$ such that $\bar{K} \subset M \mathbb{B}$. By continuity of $\varphi$, there exists $M > M$ such that $\{F_G(u, a) + \nu \mathbb{B} \}$ for all $u \in \bar{U}_{\mathbb{B}}$. By definition of $M$, we define in (9). Using this construction, we divide the proof in two main steps.

Step 1: Stability: The closed-loop system (7) is in standard form for the application of singular perturbation theory for nonsmooth systems (see [20] and [23]). The boundary layer dynamics are $\dot{x} = f(x, \hat{u} + a\mu)$, $\hat{u} = 0$, $\xi = 0$, $\mu = 0$. By Assumption 1, the plant dynamics (3) have a well-defined quasi-steady-state manifold $x^* = \ell(x, a + a\mu)$. Therefore, the closed-loop system has a well-defined reduced system, given by the following:

$$\begin{align*}
\dot{\hat{u}} &\in -F_G(\hat{u}, \xi), \quad \varepsilon_1 \frac{d\hat{u}}{d\tau} = -2\pi R_{\beta} \mu, \quad \mu \in \mathbb{T}^n, \\
\dot{\xi} &\in -F_{\xi}(\hat{u}, a + \mu), \quad \varepsilon_2 \frac{d\xi}{d\tau} = -\xi + F_G(\hat{u}, a)
\end{align*}$$

with $F_G$ given by (9). When $a = 0$, this system is also in the standard form for the application of singular perturbation theory, with $\varepsilon_2$ acting as small parameter. Using the definition of $F_G$ and the exponential stability properties of the low-pass filter in (17), we obtain the reduced nominal average dynamics $\frac{d\varepsilon_1}{d\tau} = -k F_G(\hat{u}, \xi) = -k F_G(\hat{u}_0)$. By Assumption 5, this system renders the set $\bar{A}_0 \cap UFGXs$ with pair $(\beta_{G,k}, T'_k)$.

Now, for the purpose of analysis, let us restrict the dynamics (18) to evolve in the compact flow set $S = K \times M \mathbb{B}$, where $M := M + 1$; and the dynamics (17) and (7) to evolve in the compact flow set $S = K \times M \mathbb{B} \times \mathbb{T}^n \times \Xi$. Applying [21, Th. 2] and [25, Th. 7.21] to the restricted system (18), we immediately obtain that the compact set $A'_k := A_0 \times M \mathbb{B}$ is GAS as $(a, e_2) \to 0^+$ with $\beta_{G,k} \in K C^L$. Since by the definition of solutions to (1), we have that $\{\xi(\tau)\}_{\tau > 0} = \emptyset$ for all $\tau \in \mathbb{T}^n$, we also have that $\{\xi(\tau)\}_{\tau} = \{\hat{u}(\tau)\}_{\tau}$, Thus, for each $\nu \in (0, 0.5\nu)$ there exists $e_2' > 0$ such that for all $e_2' \in (0, e_2')$ there exists $\alpha^* > 0$ such that for each $a \in (0, \alpha^*)$ every solution of the restricted system (18) satisfies the bound

$$\xi(\tau) \leq \beta_{G,k}(\|\xi(0)\|, \tau) + 0.5\nu.$$
\[ F_0 := 0 \text{ whenever } \xi_2 = 0. \] As shown in [2], the parameters \((\alpha_1, \alpha_2)\) can be selected again as in Remark 1 to guarantee continuity of \(F_0\). The input \(u\), the dynamic oscillator, the mappings \((F_G, M)\), and the constants \((k_1, k_2, k_3, a)\) are defined again as in Section III. However, system (21) has an extra state \(\xi_1\) with dynamics depending on the mapping \(F_H\), defined as \(F_H(y, \mu) := y N(\mu)\), where \(N: \mathbb{R}^n \to \mathbb{R}^{n \times n}\) is matrix-valued function with entries \(N_{ij}\) satisfying \(N_{ij} = N_{ji}\), \(N_{ij} = \frac{\partial^2}{\partial \mu_i \partial \mu_j} \xi_2 \) for \(i \neq j\), and \(\mu_i = \mu_j\) for \(i = j\), where \(\mu_i\) is the \(i\)th entry of the vector \(\mu\).

Remark 6: In (21), the state \(\xi_1\) is a matrix of dimension \(n \times n\).
Therefore, the dynamics of \(\xi_1\) must be understood as a matrix differential equation. This notation, which is used to simplify our presentation, is consistent with the notation used in the Newton-based ESCs of [24].

By using [2, Lemma 6], we can analyze system (21) via singular perturbation and averaging theory for nonsmooth systems. In particular, for the reduced dynamics of (21), we can carry out Taylor expansion of \(\phi(\tilde{u} + a\mu)\) around the point \(\tilde{u}\) for small values of \(a\), where we now retain the second-order terms: \(\phi(\tilde{u} + a\mu) = \phi(\tilde{u}) + a^2\mu^T \nabla^2 \phi(\tilde{u}) + O(a^3)\). Using this expansion, the definitions of the mappings \(M, N, F_G, F_H\), and [2, Lemma 6], we obtain an average system with state \(\xi^A = (\tilde{u}^A, \xi^A)\) and dynamics

\[
\frac{d\tilde{u}^A}{dr} = F_u(\tilde{u}^A, \xi^A) \quad \text{and} \quad \frac{d\xi^A}{dr} = F_{\xi}(\tilde{u}^A, \xi^A, \mu) + O(a^3),
\]

which is also a singularly perturbed system. When \(a = 0\), and for fixed-values of \(\tilde{u}^A\), the fast dynamics render locally exponentially stable [24, pp. 1761] the quasi-steady-state manifold \(\xi^A(\tilde{u}^A) = (\nabla^2 \phi(\tilde{u}^A)\xi^A_1 + \xi^A_1 + O(a))\). Therefore, the reduced nominal average dynamics of (23) correspond to

\[
\dot{\tilde{u}} = -k\nabla^2 \phi(\tilde{u})^{-1}(\nabla^2 \phi(\tilde{u}))(\tilde{u}) + O(a^3).
\]

Using again \(V_N\), we obtain that \(\dot{V}_N(\tilde{u}) = -(k1 \alpha \xi_1^2)\), and by using the Comparison Lemma and Assumptions 3 and 6, we obtain that the solution of (20) satisfies \(\|\tilde{u}(\tau)\|_A \leq \frac{2}{k1} e^{-k1 \alpha \tau}(\tilde{u}(0))_A\) for all \(\tau \geq 0\). Now, considering the fixed-time Newton-like ESC (FxTNECS) of Fig. 3 that generates the following closed-loop system in the \(\tau\)-time scale:

\[
\begin{align*}
\frac{d\tilde{u}}{d\tau} &= -k\nabla^2 \phi(\tilde{u})^{-1}\nabla \phi(\tilde{u}), \\
\frac{d\xi}{d\tau} &= -(\xi F_H(y, \mu) - \xi_1),
\end{align*}
\]

where \(\mu \in T^n\), and where the learning dynamics \(F_u\) are now defined as follows:

\[
F_u(\tilde{u}, \xi) := \xi_1 \left(\frac{\xi_2}{|\xi_2|^{a_1}} + \frac{\xi_2}{|\xi_2|^{a_2}}\right)
\]
Proof: The proof is almost identical to the proof of Theorem 1, with the difference that the results hold only locally due to the local stability properties of the boundary layer dynamics of system (23). In particular, define $A_0 := (\nabla^2 \phi(A_0))^{-1} \times \{0\}$, and note that by the stability properties of the boundary layer dynamics of (23), and by the smoothness properties of $\phi$, there exists $\delta > 0$ and $a^* > 0$ such that for all $\xi(0) \in ((\nabla^2 \phi(A_0))^{-1}) \times \{0\} + \delta \mathbb{B}$, $a \in (0, a^*)$ and all $|\hat{u}^A|_{A_0} \leq \delta$, every solution $\xi^A$ of the boundary layer dynamics of system (23) is complete and satisfies $|\xi^A(\tau)|_{A_0} \leq \delta^\prime \mathbb{B}$ for all time $\tau \geq 0$, and $\delta^\prime > 0$. In turn, since $\beta_{N,k}$ is a class $\mathcal{K}\mathcal{L}_T$ function, there exists $\Delta > 0$ sufficiently small such that $\beta_{N,k}(\Delta, 0) > 0.5 \delta < \delta$, which implies that for $\Delta$ sufficiently small the set $\tilde{K}$ used in the proof of Theorem 1 can be constructed such that $\tilde{K} \subset A_0 + \delta B$. From here we can repeat the same Steps 1 and 2 of the proof of Theorem 1.

Remark 7: In the above-mentioned discussion, the parameter $\varepsilon_2$ was the same for the dynamic of $\varepsilon_1$ and $\varepsilon_2$. However, this was done only to simplify the presentation, and in practice they can be different in order to simplify the tuning of the algorithm. As in Theorem 1, the convergence result of Theorem 2 covers a variety of Newton-based ESCs that go beyond the one presented in Fig. 3, having fixed-time or finite-time (in this case $T_N^* = T_N(\Delta)$) convergence properties, including ESCs with discontinuous vector fields.

For constant values of $k > 0$, $\alpha \in (0, 1)$, $p = (\nu, \alpha)$, and by using the structure of the exponential $\mathcal{K}\mathcal{L}_T$ bound of (20), the value of $T_N^*$ in (25) will be smaller than the convergence time $\tau_N^*$ of the NESC whenever the initial conditions $\bar{u}(0)$ of the controllers are in the set $\Omega_{Np} := \{u_n, \bar{u} \in \mathbb{R}^n : \Delta \geq |u_n, \bar{u}|_{A_0} > \varepsilon \exp(\pi/2\alpha_1)\}$. Since Theorem 2 is a local result, in this case $\Delta$ cannot be selected arbitrarily large. However, this does not necessarily imply that the set $\Omega_{Np}$ is empty, as $\nu \to 0^+$.

Example 2: We consider the same plant and cost function of Example 1, but this time we simulate the closed-loop system using the FxTNESC of Fig. 3. We set $\alpha_2 = 0.5 = -\alpha_2$, and $k = \pi/100$, which assigns $T_N^* = 100$ via (25). To guarantee that the NESC behaves (on compact sets) as its average system in the slowest time scale, we use again $\alpha = 0.01, \varepsilon_0 = 5 \times 10^{-5}, \varepsilon_1 = 1 \times 10^{-4}, \varepsilon_2 \varepsilon_1 = 1 \times 10^2, \varepsilon_2 \varepsilon_2 = 2 \times 10^{-1}$. For the oscillator we used $2\pi \theta_1 \approx 5, 2\pi \theta_2 \approx 3.5,$ and $\mu(0) = [0, 1, 0, 1]^T$. We further used a low-pass filter to smooth the Hessian estimator. This filter is not necessary for the simulation, but it can simplify the tuning of the Newton-based ESCs (see [24]). We computed again a numerical approximation of the reachable set (for the state $\hat{u}$) of the NESC and the FxTNESC from initial conditions satisfying $\bar{u} \in A_0 + 10^{3} \mathbb{B}$, $\xi(0) = 1, \xi_1(0) = [0.25, -0.1, -0.1, 0.25]$, via $1 \times 10^9$ simulations with random initialization in this set. The result is shown in the left plots of Fig. 4. The fixed-time convergence property of the proposed FxTNESC is further illustrated in the logarithmic scale of the upper right plots, shown in Fig. 4, which also shows the trajectories of the reduced average nominal dynamics (20) and (24) using the same gains $k$ and with identical initialization. As shown in the lower right plot of Fig. 4, the trajectories of the ESCs remain close to the trajectories of their respective reduced average nominal dynamics.

Remark 8: When the plant (3) is a static map, i.e., $y = h(u)$, one can take $\phi(u) = h(u)$ and $\varepsilon_0 = 1$. In this case, Theorems 1 and 2 recover the results of [1] and [2], which are specialized for the learning dynamics (13) and (22), now with sharper bounds $T_G$ and $T_N^*$ given by (15) and (25), respectively. In this case, the convergence time $T_N^*$ can be completely prescribed a priori by the user, without the need of retuning the gain $k$ for different initial conditions. A similar observation holds for $T_G^*$ if a lower bound on $\kappa$ is known a priori. When the plant is dynamic, the bounds hold in the $t$-time scale with $T_G^*/\varepsilon_0$ and $T_N^*/\varepsilon_0$.

VI. CONCLUSION

In this article, we introduced a novel class of nonsmooth ESCs with convergence bounds characterized by class-$\mathcal{K}\mathcal{L}_T$ functions that confer suitable transient performance. Our main results can be used for the design and analysis of different averaging-based ESCs that go beyond those considered in this article, and which are not necessarily Lipschitz continuous, or even continuous. In the latter case, the ESCs must be analyzed using the framework of differential inclusions. When the plant is a static map, the convergence time of the algorithms can be prescribed a priori by the users without retuning the gain of the learning dynamics for different initial conditions. Two numerical examples were presented to illustrate our theoretical results. Future research directions will focus on FxTESCs for multiagent systems.

APPENDIX

The following Lemma is a minor extension of [21, Th. 2], for the case when the average dynamics have a compact set that is SGPS instead of UGAS. The proof follows directly by using [21, Th. 1], and the same steps of the proof of [22, Th. 7], and therefore, it is omitted due to space limitations.

Lemma 2: Consider the singularly perturbed system with state $(x_1, x_2, y) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ and dynamics

$$\dot{x}_1 \in F_0(x), \quad \dot{x}_2 \in G_0(x, y), \quad \varepsilon \dot{y} = H(x, y) \quad (26)$$

where $X_2 \subset \mathbb{R}^n$ and $X_2 \subset \mathbb{R}^n$ are compact sets, $x = [x_1, x_2]^T$, and for each $\delta > 0$ the set-valued mapping $F_0 : \mathbb{R}^{2n} \mapsto \mathbb{R}^n$ satisfies the Basic Conditions, and the mappings $G_0 : \mathbb{R}^{2n} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $H : \mathbb{R}^{2n} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ are Lipschitz continuous. Suppose the following holds.

1) Existence of Average: There exists $\delta^* > 0$ such that for each $\delta \in (0, \delta^*)$ there exists a continuous function $G^\delta : \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$ such that for each compact set $K \subset \mathbb{R}^n$ there exists a class-$\mathcal{L}$ function $\phi^\delta : \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$.
such that for each $L > 0$, and each solution $y_k : [0, L] \to X_y$ of the system: $(x, y_k) \in K \times X_{x,y}$, $\dot{x} = 0$, $y_k = H(x, y_k)$, the following holds $\frac{1}{L} \int_0^L G^T_\lambda(x) - G(x, y_k(s)) \, ds \leq \sigma_K(L)$. The $G$-PAS of Average System: There exists a compact set $A_\lambda \subset \mathbb{R}^n$ such that the system $\dot{x}^\lambda \in F_\lambda(x^\lambda)$, $\dot{x}^\lambda = G^\lambda_\beta(x^\lambda)$ renders the set $A_\lambda \times X_{x,y}$ G-PAS as $\lambda \to 0^+$ with $\beta \in K \mathcal{L}_\mathcal{F}$.

Lemma 3: Suppose that $V : \mathbb{R}^n_{>0} \to \mathbb{R}^n_{>0}$ satisfies the Assumptions of Lemma 1, and there exist $\lambda_1, \lambda_2, p > 0$ such that $\lambda_1|\lambda|^p \leq V(x) \leq \lambda_2|\lambda|^p$. Then, system (26) renders the compact set $A_\lambda \times X_{x,y}$ G-PAS as $(\varepsilon, \delta) \to 0^+$ with $\beta \in K \mathcal{L}_\mathcal{F}$. □

Proof: Let $\dot{y} = -a \gamma_1 y_1 - b \gamma_2 y_2$, with $\gamma_1, \gamma_2$ as in Lemma 1. Using steps as in the proof of [23, Lemma 2], and the fact that $y^\tau = 0$ is an equilibrium point, it follows that every solution $y$ satisfies $\frac{\sqrt{a}}{\sqrt{b}} \arctan(\sqrt{a/b} y^\tau(t)) = \max(0, -t + 2\pi \arctan(\sqrt{a/b} y^\tau(t)))$ for all $t \geq 0$. Solving for $y$ and using the generalized Comparison Lemma of [34, Lemma 1], we obtain $V(x(t)) \leq (a/b)^{\alpha} \max(0, -b/2\alpha t + \arctan(\sqrt{b/2\alpha V(x(t))}^{1/2}))^{2\alpha}$ for all $t \geq 0$. The result now follows by using the upper and lower bounds of $V$, and the continuity and monotonicity properties of the functions arctan : $\mathbb{R}_{>0} \to (0, \pi/2]$ and tan : $[0, \pi/2] \to \mathbb{R}$.

Acknowledgment

The first author would like to thank A. R. Teel for fruitful discussions on singularly perturbed nonsmooth systems.

References