MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

A Neuro-Adaptive Architecture for Extremum Seeking Control Using Hybrid Learning Dynamics

Benosman, M.

TR2017-050 April 2017

Abstract

This paper presents a novel approach to achieve online multivariable hybrid optimization of response maps associated to set-valued dynamical systems, without requiring the use of averaging theory. In particular, we propose a prescriptive framework for the analysis and design of a class of adaptive control architectures based on neural networks (NN) and learning dynamics described by hybrid dynamical systems (HDS). The NNs are used as model-free gradient approximators that are online tuned in order to obtain an arbitrarily precise estimation on a compact set of the gradient of the response map of the system under control. For the closed-loop system a semi-global practical asymptotic stability result is obtained, and the results are illustrated via numerical examples.

American Control Conference (ACC)

A Neuro-Adaptive Architecture for Extremum Seeking Control Using Hybrid Learning Dynamics

Jorge I. Poveda, Kyriakos G. Vamvoudakis, and Mouhacine Benosman

***Abstract*— This paper presents a novel approach to achieve online multivariable hybrid optimization of response maps associated to set-valued dynamical systems, without requiring the use of averaging theory. In particular, we propose a prescriptive framework for the analysis and design of a class of adaptive control architectures based on neural networks (NN) and learning dynamics described by hybrid dynamical systems (HDS). The NNs are used as model-free gradient approximators that are online tuned in order to obtain an arbitrarily precise estimation on a compact set of the gradient of the response map of the system under control. For the closed-loop system a semi-global practical asymptotic stability result is obtained, and the results are illustrated via numerical examples.**

I. INTRODUCTION

The problem of online optimization of response maps in nonlinear dynamical plants, also called extremum seeking control (ESC), has received significant attention during the last years, e.g., [1], [2], [3], [4], [5], [6]. Though most of the existing extremum seeking algorithms are based on averaging theory for smooth systems [7], [8], or hybrid systems [9], other type of Lipschitz continuous extremum seeking architectures that are not based on averaging theory have shown potential advantages in terms of transient performance, e.g., [10], [11], [12], [13]. Because of this, and motivated by recent developments in averaging-based hybrid extremum seeking control [5], we present in this paper a novel class of *hybrid* extremum seeking architectures that do not rely on averaging theory. In particular, we propose a prescriptive framework for the design and analysis of hybrid ESCs based on neural networks (NN) with provable stability guarantees. The design of this type of systems is also motivated by results in the area of neuro-adaptive control [14], and neuro-adaptive optimal control [15], which have been successfully applied in several contexts where the model of the plant is absent or partially known. Because of the use of NN techniques to approximate the gradient of the cost function from cost measurements, the class of hybrid ESCs proposed here is termed *neuro-adaptive hybrid extremum seeking control* (NHESC). Since the approach presented in this paper is of modular nature, the hybrid optimization/learning dynamics can be designed independently of the NN block. As shown in [5], allowing for the presence of hybrid dynamics in the optimizing/learning block opens the door for the design of novel ESCs not considered so far in the literature, such as discontinuous dynamics for "finite-time" convergence, set-valued dynamics for game theoretical applications, or learning dynamics for multi-agent systems with switching graphs, to just name a few. Moreover, the framework of HDS allows us to consider discontinuous or set-valued plants. The results presented in this paper thus parallel those presented in [5] for HESCs based on averaging theory, offering an alternative approach based on NN, with the potential practical advantage of obtaining better transient performance by using a large number of neurons, as well as minimizing the impact of the choice of dither signal on the performance of the closed-loop system by only requiring a standard persistency of excitation (PE) condition. Indeed, removing the need for averaging allows us to eliminate one of the time-scales that emerge in the closed-loop system, as well as to substitute the constraint of using dithers with "sufficiently" small amplitudes and frequencies, by the condition of using a sufficiently large number of neurons. This observation is relevant since in many practical applications the excitation signal is turned off after an initial learning phase [15].

This paper is organized as follows: Section II presents some mathematical preliminaries. Section III characterizes the type of plants that we study. Section IV characterizes the main block components of the NHESC. Section V presents our main stability result. Section VI shows numerical examples, and Section VII ends with some conclusions.

J. I. Poveda is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA, 93106-9560, USA. E-mail: jipoveda@ece.ucsb.edu. This work was done in part while J. I. Poveda was an intern at Mitsubishi Electric Research Laboratories, Cambridge, MA, 02139, USA.

K. G. Vamvoudakis is with the Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA, 24061-0203 USA. E-mail: kyriakos@vt.edu.

M. Benosman is with Mitsubishi Electric Research Laboratories, Cambridge, MA, 02139, USA. E-mail: m.benosman@ieee.org.

II. PRELIMINARIES

We denote by $\mathbb{R}$ ($\mathbb{R}_{\geq 0}$) the set of real numbers (resp. non-negative real numbers), and by $\mathbb{Z}_{\geq 0}$ the set of non-negative integers. For any $\rho > 0$ the closed ball of appropriate dimension in the Euclidean norm, of radius ρ, is denoted by $\rho\mathbb{B}$. We denote by $a \cdot \mathbb{S}$ the circle in $\mathbb{R}^2$ with radius a. Given a compact set $\mathcal{A} \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$, we use $|x|_{\mathcal{A}} := \inf_{y \in \mathcal{A}} \|x - y\|$ to denote the distance of x to $\mathcal{A}$. A set-valued mapping $M : \mathbb{R}^p \rightrightarrows \mathbb{R}^n$ is outer semicontinuous (OSC) at x if for each $(x_i, y_i) \mapsto (x, y) \in \mathbb{R}^p \times \mathbb{R}^n$ satisfying $y_i \in M(x_i)$ for all $i \in \mathbb{Z}_{\geq 0}$, we have $y \in M(x)$. A mapping M is locally bounded (LB) at x if $M(x)$ is bounded. The mapping M is OSC and LB relative to a set $K \in \mathbb{R}^p$ if M is OSC for all $x \in K$ and $M(K) := \cup_{x \in K} M(x)$ is bounded.

In this paper we will consider dynamical systems with continuous and discrete-time dynamics, called *hybrid dynamical systems* (HDS). A HDS with state $x \in \mathbb{R}^n$ is represented by its data $\mathcal{H} := \{C, F, D, G\}$, and the evolution of x is characterized by the equations

$$\dot{x} \in F(x), \quad x \in C \tag{1a}$$
$$x^+ \in G(x), \quad x \in D, \tag{1b}$$

where the set-valued mappings $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, called the flow map and the jump map, respectively, describe the evolution of the state x when it belongs to the flow set C or/and the jump set D, respectively. We will always impose the following conditions on the data of the system:

(C1) The sets C and D are closed.
(C2) F is OSC and LB relative to C, $F(x)$ is convex for every $x \in C$, and $C \subset \text{dom}(F)$.
(C3) G is OSC and LB relative to D, and $D \subset \text{dom}(G)$.

Solutions of (1) are defined on *hybrid time domains*. We recall [16, Chap. 2] that a compact hybrid time domain is a subset of $\mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ of the form $\bigcup_{j=0}^{J}([t_j, t_{j+1}] \times \{j\})$ for some nonnegative integer J and some real numbers $0 = t_0 \leq t_1 \leq \ldots \leq t_{J+1}$. A hybrid time domain is a set $E \subset \mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ such that, for each $(T, J) \in E$ the set $E \cap ([0, T] \times \{0, \ldots, J\})$ is a compact time domain. A *hybrid arc* is a mapping $\phi : E \to \mathbb{R}^n$ such that E is a hybrid time domain and, for each $j \in \mathbb{Z}_{\geq 0}$, $t \to \phi(t, j)$ is locally absolutely continuous. A hybrid arc x is a *solution* to (1) from $x_0 \in \mathbb{R}^n$, denoted by $x \in \mathcal{S}(x_0)$, if: **1)** $x(0,0) = x_0$. **2)** If $(t_1, j), (t_2, j) \in \text{dom}(x)$ with $t_1 < t_2$ then, for almost every $t \in [t_1, t_2]$, $x(t, j) \in C$ and $\dot{x} \in F(x(t, j))$. **3)** If $(t, j), (t, j+1) \in \text{dom}(x)$, then $x(t, j) \in D$ and $x(t, j+1) \in G(x(t, j))$.

The following two stability definitions would be used in this paper:

Definition 2.1: Let $\mathcal{H}$ be a hybrid system of the form (1), and $\mathcal{A} \subset \mathbb{R}^n$ be a compact set. The set $\mathcal{A}$ is uniformly globally asymptotically stable (UGAS) for $\mathcal{H}$ if there exists a $\mathcal{KL}$ function β such that any solution x to $\mathcal{H}$ satisfies $|x(t, j)|_{\mathcal{A}} \leq \beta(|x(0, 0)|_{\mathcal{A}}, t + j)$, for all $(t, j) \in \text{dom}(x)$. ■

Definition 2.2: For a HDS parametrized by $\varepsilon > 0$, denoted $\mathcal{H}_\varepsilon := \{C_\varepsilon, F_\varepsilon, D_\varepsilon, G_\varepsilon\}$, a compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be semi-globally practically asymptotically stable (SGP-AS) as $\varepsilon \to 0^+$ if there exists a function $\beta \in \mathcal{KL}$ such that the following holds: For each $\Delta > 0$ and $\nu > 0$ there exists $\varepsilon^* > 0$ such that for each $\varepsilon \in (0, \varepsilon^*)$ each solution x of $\mathcal{H}_\varepsilon$ that satisfies $|x(0, 0)|_{\mathcal{A}} \leq \Delta$ also satisfies $|x(t, j)|_{\mathcal{A}} \leq \beta(|x(0, 0)|_{\mathcal{A}}, t + j) + \nu$, for all $(t, j) \in \text{dom}(x)$. ■

Remark 2.1: If the sets C_ε and D_ε are compact, SGP-AS is equivalent to *global practical asymptotic stability* (GP-AS), since Δ can be selected sufficiently large to encompass every initial condition where solutions of $\mathcal{H}_\varepsilon$ are defined. ■

III. PROBLEM STATEMENT

Consider a constrained dynamical system described by the differential inclusion

$$\dot{\theta} \in P(\theta, u), \quad (\theta, u) \in \Lambda_\theta \times \mathbb{U}, \quad y = \varphi(\theta), \tag{2}$$

where $\theta \in \mathbb{R}^m$, $u \in \mathbb{R}^n$, and y is a scalar output. System (2) is characterized by the set-valued mapping $P : \mathbb{R}^m \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and the output function $\varphi : \mathbb{R}^m \to \mathbb{R}$. The state θ is constrained to evolve in the compact set $\Lambda_\theta := \lambda_\theta \mathbb{B}$, where $\lambda_\theta \in \mathbb{R}_{>0}$ is a parameter that can be selected arbitrarily large in order to encompass any *complete* solution of interest, i.e., any solution of interest with an unbounded time domain. The input u is constrained to evolve in the set $\mathbb{U} \subset \mathbb{R}^n$. System (2) must satisfy the following regularity assumption:

Assumption 3.1: $P(\cdot, \cdot)$ satisfies condition (C2) relative to $\Lambda_\theta \times \mathbb{U}$, and $\varphi_i(\cdot)$ is continuous. ■

Equation (2), together with the dynamics $\dot{u} = 0$, i.e., u kept constant, describe an open-loop system that can be modeled as a HDS of the form (1) with no jumps, i.e.,

$$\mathcal{H}_{\theta,u} := \{\Lambda_\theta \times \mathbb{U}, P \times \{0\}, \emptyset, \emptyset\}. \tag{3}$$

We impose the following stability assumption on (3).

Assumption 3.2: There exists a set-valued mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ that is OSC and LB relative to $\mathbb{U}$, such that for each $\rho \in \mathbb{R}_{>0}$ the compact set $\mathbb{M}_\rho := \{(\theta, u) : \theta \in H(u), u \in \mathbb{R}^n \cap \rho\mathbb{B}\}$ is UGAS for system (3) with restricted flow set $\Lambda_\theta \times (\mathbb{U} \cap \rho\mathbb{B})$. ■

In order to have a well defined optimization problem the following two assumptions are also required.

Assumption 3.3: Let $H(\cdot)$ be given by Assumption 3.2. For each $u \in \mathbb{R}^n$ and each pair $\theta, \theta' \in H(u)$, we have that $\varphi(\theta) = \varphi(\theta')$. ■

The response map of system (2) is then defined as

$$J(u) := \{\varphi(\theta) : \theta \in H(u)\}, \tag{4}$$

where $H(\cdot)$ is given in Assumption 3.2. Note that under Assumption 3.3 the mapping $J(\cdot)$ is well-defined. The optimization problem is then characterized by the following assumption.

Assumption 3.4: $J : \mathbb{R}^n \to \mathbb{R}$ is smooth and has a global minimum $u^* \in \mathbb{U}$. ■

Assumptions 3.1-3.3 are the generalization for differential inclusions of the standard ESC assumptions used when the plant is a Lipschitz continuous ODE, e.g., [2]. Thus, as in the standard ESC, our main goal is to design model-free feedback mechanisms that guarantee convergence of u to u^* by using only measurements of y, assuming no knowledge of J, P, and φ.

IV. NEURO-ADAPTIVE HYBRID EXTREMUM SEEKING CONTROL

A conceptual scheme of the proposed neuro-adaptive hybrid extremum seeking architecture for the minimization of J is presented in Figure 1. This scheme is comprised of four main blocks: (A) the neural network-based model-free gradient approximator; (B) the differential inclusion generating the dither signal μ needed to provide the persistency of excitation to the system; (C) the hybrid learning dynamics $\hat{\mathcal{H}}_\delta$; and d) the plant, which was already characterized in the previous section. We proceed now to the describe in detail the components (A), (B) and (C).

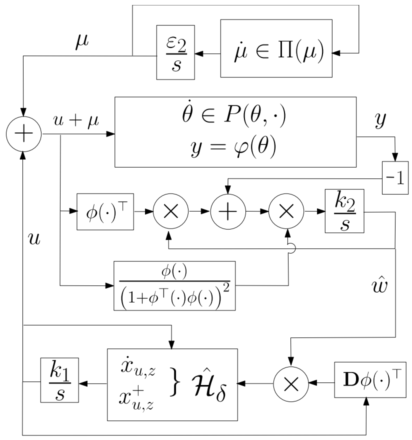

Fig. 1: Modular scheme of NHESC

A. NN-based Model-Free Approximator

Under Assumption 3.4 the response map $J(\cdot)$ is smooth, which under the High-Order Weierstrass Approximation Theorem and the results in [17], imply the existence of a complete independent basis set of functions $\{\phi_i(u)\}$ such that J and ∇J are uniformly approximated on compact sets, i.e., there exist scalar coefficients c_i such that

$$J(u) = \sum_{i=1}^{N} c_i \phi_i(u) + \sum_{i=N+1}^{\infty} c_i \phi_i(u), \tag{5}$$

$$\frac{\partial J(u)}{\partial u_j} = \sum_{i=1}^{N} c_i \frac{\partial \phi(u)}{\partial u_j} + \sum_{i=N+1}^{\infty} c_i \frac{\partial \phi(u)}{\partial u_j}, \tag{6}$$

where $\phi(u) = [\phi_1(u), \dots, \phi_N(u)]^\top : \mathbb{R}^n \to \mathbb{R}^N$, and the last terms in (5)-(6) converge uniformly to 0 as $N \to \infty$. Therefore, given $N \in \mathbb{Z}_{\geq 1}$, (5) and (6) can be written as

$$J(u) = \phi(u)^\top w^* + \epsilon(u) \tag{7}$$

$$\nabla J(u) = \mathbf{D}\phi(u)^\top w^* + \nabla\epsilon(u), \tag{8}$$

where $w^* \in \mathbb{R}^N$, and $\mathbf{D}\phi(u)$ is the Jacobian matrix of ϕ evaluated at u. The mapping $\phi : \mathbb{R}^n \to \mathbb{R}^N$ is called the NN activation function vector, N the number of neurons in the hidden layer, and $\epsilon(\cdot)$ the NN approximation error.

Assumption 4.1: The function $\phi(\cdot)$ is continuous. ■

The following lemma corresponds to [15, Lemma 1], and establishes the approximation properties of the neurons in the hidden layer.

Lemma 4.1: Let $\Omega \subset \mathbb{R}^n$ be compact. Then as $N \to \infty$ the approximation errors $\epsilon(u)$ and $\nabla\epsilon(u)$ satisfy $\epsilon(u) \to 0$ and $\nabla\epsilon(u) \to 0$, uniformly. Moreover, for each fixed N, there exist $(\bar{\epsilon}_1, \bar{\epsilon}_2, \bar{w}) \in \mathbb{R}^3_{>0}$ such that $\|\epsilon(u)\| \leq \bar{\epsilon}_1$, $\|\nabla\epsilon(u)\| \leq \bar{\epsilon}_2$, and $\|w^*\| \leq \bar{w}$, for all $u \in \Omega$. ■

The main challenge in equations (7) and (8) is that the ideal weights w^* are unknown. Therefore, they have to be estimated online by using only measurements of J. In order to do this, let u be fixed, and define the output of the NN as $\hat{J}(u) = \hat{w}^\top \phi(u)$, where $\hat{w}$ is the estimated value of the NN weights w^*. Define the weight approximation error as

$$\tilde{w} = \hat{w} - w^*, \tag{9}$$

and the estimation error of J as

$$e = \hat{J} - J, \quad \text{or} \quad e = \tilde{w}^\top \phi(u) + \epsilon(u). \tag{10}$$

We aim to select $\hat{w}$ to minimize the squared residual error $E = \frac{1}{2} e^\top e$. To do this, we consider the following learning dynamics based on the modified Levenberg-Marquardt gradient descent algorithm

$$\dot{\hat{w}} = -\Gamma \frac{\phi}{(1 + \phi^\top \phi)^2} e, \tag{11}$$

where $\Gamma \in \mathbb{R}_{>0}$, which has a normalization term $(1+\phi^\top\phi)^2$ instead of the standard $(1+\phi^\top\phi)$, see also [18]. The learning dynamics (11) will be constrained to evolve in a compact set Ω_c that will be defined in the next section.

B. PE Signal Generator

Similar to standard adaptive architectures [18], in order to obtain convergence of the learning dynamics (11) to their correct values, a persistency of excitation condition is needed in ϕ. To achieve this, a dither signal $\mu : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ needs to be injected to u for all $t \geq 0$. This signal will be generated as the solution of the time-invariant differential inclusion

$$\dot{\mu} \in \Pi(\mu), \quad \mu \in \Psi, \tag{12}$$

where $\Pi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, and $\Psi \subset \mathbb{R}^n$ is the flow set describing the points in the space where μ is allowed to evolve.

Assumption 4.2: The set-valued mapping $\Pi(\cdot)$ satisfies (C2) relative to Ψ, and Ψ is compact. ■

Note that Assumption 4.2 is not restrictive, and it is satisfied if, for instance, $\Pi(\cdot)$ is single-valued and continuous. However, using set-valued mappings as signal generators allows us to consider a broader class of excitation signals compared to those generated by differential equations.

The following stability and completeness assumption is also imposed on system (12).

Assumption 4.3: For system (12) there exists a UGAS compact set $\mathcal{A}_\mu \subset \Psi$, and for each initial condition in Ψ there exists at least one complete solution $\mu(\cdot)$. ■

For system (12) the dimension of the state μ can be enlarged to generate enough dither signals with different excitation frequencies. A simple example of this case is presented as follows:

Example 4.1: Consider the time-invariant oscillator given by the dynamics

$$\dot{\mu} = \Phi_\omega \cdot \mu, \quad \mu \in a \cdot \mathbb{S}^n, \quad a > 0, \tag{13}$$

where $\Phi_\omega : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a block matrix of the form

$$\Phi_\omega := \begin{bmatrix} \Omega_{\omega_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Omega_{\omega_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Omega_{\omega_3} \end{bmatrix}, \tag{14}$$

and where the block components Ω_{ω_i} are defined as

$$\Omega_{\omega_i} := \begin{bmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{bmatrix}, \quad \omega_i \in \mathbb{R}_{>0}, \quad \forall \ \ i \in \{1, \dots, n\}.$$

From Definition 2.1 it is easy to see that system (13) renders the set $a\mathbb{S}^n$ UGAS. Then, one can take as dither signals linear combinations of the odd components of the solutions of (13), which are given by $\mu_i(t) = \cos(\omega_i t)\mu_{2i-1}(0) + \sin(\omega_i t)\mu_{2i}(0)$, for $i \in \{1, \ldots, n\}$. Similar sinusoidal excitation signals have been used by averaging-based hybrid extremum seekers [5], as well as by algorithms not based on averaging [12]. ■

Although the stability condition imposed by Assumption 4.3 is critical for our stability analysis of the closed-loop system, so it is the PE property associated to the signal ϕ in the learning dynamics (11):

Assumption 4.4: Define $\phi(t) := \phi(u + \mu(t))$. Then, there exists constants $(\beta_1, \beta_2, T) \in \mathbb{R}^3_{>0}$ such that for each $u \in \mathbb{R}^n$ and each solution $\mu(\cdot)$ of (12), the normalized time varying signal $\bar{\phi}(t) := \frac{\phi(t)}{1+\phi(t)^\top \phi(t)}$ satisfies

$$\beta_1 I \leqslant \int_t^{t+T} \bar{\phi}(\tau)\bar{\phi}^\top(\tau) d\tau \leqslant \beta_2 I$$

for all $t \geq 0$. ■

Remark 4.1: Note that under Assumptions 4.1 and 4.3, and by the compactness of Ψ, for each fixed u, the signal $\phi(t)$ is uniformly bounded. Moreover, there will always exist a signal ϕ with unbounded time domain, such that the PE condition can actually be evaluated. ■

Remark 4.2: One can ensure that the signal $\bar{\phi}(t)$ is persistently exciting by adding exploration noise formed by sinusoids of different frequencies, see [18, Chapter 4]. Such sinusoids can be obtained as in Example 4.1. ■

Using the PE property of $\bar{\phi}$, as well as the learning dynamics (11), and the change of variable (9), for each pair of positive numbers $(\rho, c) \in \mathbb{R}^2_{>0}$ we can study the stability properties of the system

$$\left.\begin{array}{l} \dot{u} = 0 \\ \dot{\mu} \in \Pi(\mu) \\ \dot{\tilde{w}} = \dfrac{-\Gamma \cdot \phi(u+\mu)}{\left(1 + \phi(u+\mu)^\top \phi(u+\mu)\right)^2} e \end{array}\right\}, \; (u,\, \mu,\, \tilde{w}) \in C_{\rho,c} \tag{15}$$

with $e = \tilde{w}^\top \phi(u+\mu) + \epsilon(u+\mu)$, and with flow set

$$C_{\rho,c} := (\rho\mathbb{B} \cap \mathbb{U}) \times \Psi \times \Omega_c, \tag{16}$$

where $\Omega_c := \{\tilde{w} \in \mathbb{R}^N : \frac{1}{2}\tilde{w}^\top \Gamma^{-1} \tilde{w} \leq c\}$. The following proposition establishes a stability result for this system with frozen input.

Proposition 4.2: Consider system (15) with flow set (16), and suppose that Assumptions 4.1 - 4.4 hold. Then, for each pair $(\rho, c) \in \mathbb{R}^2_{>0}$, and each $\bar{\epsilon}$ such that $\bar{\epsilon}\mathbb{B} \subset \text{int}(\Omega_c)$, there exists a number N^* of NN such that for all $N \geq N^*$ there exists a compact set $\mathcal{A}_c \subset \bar{\epsilon}\mathbb{B}$ such that the compact set

$$\mathcal{M}_\rho := \{(u, \mu, \tilde{w}) : u \in \rho\mathbb{B} \cap \mathbb{U}, (\mu, \tilde{w}) \in \mathcal{A}_\mu \times \mathcal{A}_c\} \tag{17}$$

is UGAS.

Proof: Since u is constant and μ is generated by the uncoupled PE signal generator (12) which renders the set $\mathcal{A}_\mu$ UGAS, it suffices to study the behavior of the dynamics of $\hat{w}$. Moreover, since u is constrained to lie in the compact set $\rho\mathbb{B} \cap \mathbb{U}$, and μ is uniformly bounded, we can take $\phi(t)$ in (15) as an independent absolutely continuous function of time. Now, consider the radially unbounded Lyapunov function

$$V_{\tilde{w}} = \frac{1}{2}\tilde{w}^T \Gamma^{-1} \tilde{w}, \tag{18}$$

and note that $\dot{\tilde{w}} = \dot{\hat{w}}$. Taking the derivative of (18) one gets

$$\begin{aligned} \dot{V}_{\tilde{w}} &= \dot{\tilde{w}}^T \Gamma^{-1} \tilde{w} \\ &= \left(-\Gamma\bar{\phi}\left(\tilde{w}^\top \bar{\phi} + \frac{\epsilon}{1+\phi^\top\phi}\right)\right)^\top \Gamma^{-1}\tilde{w} \\ &= -\tilde{w}^T \bar{\phi}\bar{\phi}^T \tilde{w} + \epsilon \cdot \bar{\phi}^T \frac{1}{(1+\phi^T\phi)} \tilde{w}. \end{aligned}$$

When $\epsilon = 0$ we have that $\dot{V}_w$ reduces to

$$\dot{V}_{\tilde{w}} = -\tilde{w}^T \bar{\phi}\bar{\phi}^T \tilde{w}, \tag{19}$$

which can be rewritten as

$$\dot{V}_{\tilde{w}} = -\frac{e^2}{(1+\phi^\top\phi)^2} \leq 0. \tag{20}$$

This implies that $\sup_{t\geq 0} |V_{\tilde{w}}(t)| < \infty$ and $\sup_{t \geq 0} \|\tilde{w}(t)\| < \infty$. Moreover, (20) and (18) imply that for each $c \in \mathbb{R}_{>0}$ the compact set $\Omega_c := \{\tilde{w} \in \mathbb{R}^N : V_{\tilde{w}}(\tilde{w}) \leq c\}$, is positively invariant under the dynamics (11), which guarantees the existence of complete solutions for system (15). Using the PE condition on $\bar{\phi}$, and equation (19), it follows by [18, Theorem 4.3.2] or [15, Thm. 1] that $\tilde{w}$ converges to zero exponentially fast. Now, if $\epsilon \neq 0$, using the PE condition, one has that $\dot{V}_{\tilde{\omega}}$ is negative if [15, Thm. 1],

$$\|\tilde{\omega}^\top \bar{\phi}\| \geq \frac{\|\epsilon\|}{\|1 + \phi^T\phi\|}. \tag{21}$$

Note that $\|1 + \phi^T\phi\| \geq 1$. Let $\bar{\epsilon} > 0$ be such that $\bar{\epsilon}\mathbb{B} \subset \text{int}(\Omega_c)$, and let $\bar{\epsilon}_1 < \bar{\epsilon}$. Then, since u is constrained to the compact set $\rho\mathbb{B} \cap \mathbb{U}$ and μ to the compact set Ψ, using Lemma 4.1 the condition $\|\epsilon\| \leq \bar{\epsilon}_1$ can be guaranteed by taking N sufficiently large. Therefore, there exists a N^* such that for any $N \geq N^*$ equation (21) is satisfied if $\|\tilde{w}^\top \bar{\phi}\| \geq \bar{\epsilon}_1$. Then, since $\|\bar{\phi}(t)\| < 1$ for all $t \geq 0$, using the Cauchy-Schwarz inequality we have that $\dot{V}_{\tilde{\omega}}$ will be negative outside the set $\tilde{\Omega}_{\bar{\epsilon}} := \{\tilde{w} \in \mathbb{R}^N : \|\tilde{w}\| \leq \bar{\epsilon}_1\}$. Finally, define $B_{\bar{\epsilon}} := \bar{\epsilon}\mathbb{B} \subset \mathbb{R}^N$ and note that for sufficiently small $\bar{\epsilon}_1$ one has that $\tilde{\Omega}_{\bar{\epsilon}_1} \subset B_{\bar{\epsilon}} \subset \text{int}(\Omega_c)$. Thus $\tilde{w}(t)$ converge exponentially fast and in finite time to $B_{\bar{\epsilon}}$. Since the flow set in (15) is compact, the previous arguments, and Assumption 4.3, imply by [16, Corollary 7.7] and [5, Lemma 1] that there exists a compact set $\mathcal{A}_c \subset w^* + \bar{\epsilon}\mathbb{B}$ such that the set $(\rho\mathbb{B} \cap \mathbb{U}) \times \mathcal{A}_\mu \times \mathcal{A}_c$ is UGAS for system (15). ■

Proposition 4.2 will allows us to make use of singular perturbation arguments once the dynamics of $\dot{u}$ are taken into consideration. This approach makes the design of the NHESC of *modular nature*, allowing us to design the learning dynamics of u independently of the dynamics of the NN.

Remark 4.3: The proof of Proposition 4.2 exploits the stability properties of μ, as well as the PE Assumption 4.4 on the *solutions* of the generator (12), instead of using the weaker and more common PE assumption on $\phi(\cdot)$. However, several classes of dithers can be generated by system (12), including the standard sinusoid signals. ■

C. Hybrid Learning Dynamics

The optimization block $\hat{\mathcal{H}}$ in Figure 1, is modeled as a HDS of the form

$$\dot{x}_{u,z} \in \hat{F}_\delta\left(x_{u,z}, \mathbf{D}\phi(u)^\top \hat{w}\right), \quad x_{u,z} \in C_u \times C_z, \tag{22a}$$

$$x_{u,x}^+ \in \hat{G}_\delta\Big(x_{u,z}\Big), \quad x_{u,z} \in D_u \times D_z, \tag{22b}$$

where $x_{u,z} := (u^\top, z^\top)^\top \in \mathbb{R}^{n+r}$ and $n + r = \ell$. The state z is an auxiliary state of dimension[1] $r \in \mathbb{Z}_{\geq 0}$, which can be used to model timers, logic states, etc. The sets $C_u, D_u \subset \mathbb{R}^n$ and $C_z, D_z \subset \mathbb{R}^r$ define the flow and jump sets for u and z, respectively, and $\delta \in \mathbb{R}_{>0}$ is a tunable parameter that gives flexibility for the design of the set-valued mappings $\hat{F}_\delta : \mathbb{R}^\ell \times \mathbb{R}^n \rightrightarrows \mathbb{R}^\ell$ and $\hat{G}_\delta : \mathbb{R}^l \rightrightarrows \mathbb{R}^\ell$. The hybrid learning dynamics (22) are based on the following two assumptions:

Assumption 4.5: There exists a $\delta^* \in \mathbb{R}_{>0}$ such that for all $\delta \in (0, \delta^*]$ the following holds:

- $C_u \times C_z$ and $D_u \times D_z$ satisfies (C1).
- $\hat{F}_\delta(\cdot, \cdot)$ satisfies (C2) relative to $(C_u \times C_z) \times \mathbb{R}^n$, and $\hat{G}_\delta(\cdot)$ satisfies (C3) relative to $D_u \times D_z$.
- For each compact set $K \subset \mathbb{R}^l$ there exists an $\epsilon^* \in \mathbb{R}_{>0}$ such that for each $\epsilon \in (0, \epsilon^*)$ and each $x_{u,z}(0, 0) \in K$ the HDS (22) with flow map $\hat{F}_\delta(x_{u,z}, \nabla J(u) + \epsilon\mathbb{B})$ generates at least one complete solution. ■

Assumption 4.6: The set $\mathbb{U}$ satisfies $C_u \cup D_u = \mathbb{U}$, and there exists a compact set $\Upsilon \subset C_z \times D_z$ such that the set $\mathcal{A}_{u,z} := \{u^*\} \times \Upsilon$ is SGP-AS as $\delta \to 0^+$ for system (22) with $\mathbf{D}\phi(u)^\top \hat{w} = \nabla J(u)$. ■

If $\hat{F}_\delta$ and $\hat{G}_\delta$ are independent of any parameter δ, Assumption 4.6 is just a UGAS assumption on (22). Different examples of hybrid learning dynamics satisfying Assumptions 4.5 and 4.6 can be found in [5].

V. Main Result

The closed-loop system is obtained by combining the plant (2), the PE generator (12), the NN dynamics (11) with J replaced by $y = \varphi(\theta)$, and the hybrid dynamics (22). The resulting system is a HDS $\mathcal{H} := \{C, F, D, G\}$ with state $x := (x_{u,z}^\top, \hat{w}^\top, \mu^\top, \theta^\top)^\top$ and equations

$$C := (C_u \times C_z) \times \Omega_c \times \Psi \times \Lambda_\theta \tag{23a}$$

$$\dot{x} \in F(x) := \begin{pmatrix} k_1 \cdot \hat{F}_\delta\left(x_{u,z}, \mathbf{D}\phi(u)^\top \hat{w}\right) \\ -k_2 \cdot \bar{\bar{\phi}}(u+\mu)\left(\hat{w}^\top \phi(u+\mu) - \varphi(\theta)\right) \\ -\varepsilon_2 \cdot \Pi(\mu) \\ P(\theta, u+\mu) \end{pmatrix} \tag{23b}$$

$$D := (D_u \times D_z) \times \Omega_c \times \Psi \times \Lambda_\theta \tag{23c}$$

$$x^+ \in G(x) := \begin{pmatrix} \hat{G}_\delta\Big(x_{u,z}\Big) \\ \hat{w} \\ \mu \\ \theta \end{pmatrix}, \tag{23d}$$

[1]The case $r = 0$ indicates that the auxiliary state z is omitted.

where $k_1 := \varepsilon_1 \cdot \varepsilon_2$, $k_2 := \varepsilon_2 \cdot \Gamma$, $(\varepsilon_1, \varepsilon_2) \in \mathbb{R}^2_{>0}$ are tunable parameters, and where

$$\bar{\bar{\phi}}(u+\mu) := \frac{\bar{\phi}(u+\mu)}{(1 + \phi(u+\mu)^\top \phi(u+\mu))}, \tag{24}$$

with $\bar{\phi}$ defined as in Assumption 4.4. The following theorem is the main result of this paper. Its proof is based on the singular perturbation results for HDS presented in [19], and it is omitted due to space limitations.

Theorem 5.1: Suppose that all the assumptions of Sections III and IV hold. Then, for each compact set $\tilde{K} \subset \mathbb{R}^l$ satisfying $\mathcal{A}_{u,z} \subset \text{int}(\tilde{K})$, there exists a pair (c, λ_θ) such that the set $\mathcal{A}_{u,z} \times \{w^*\} \times \mathcal{A}_\mu \times \Lambda_\theta$ is GP-AS as $(\delta, \frac{1}{\lceil N \rceil}, \epsilon_2, \epsilon_1) \to 0^+$ for the HDS (23) with restricted flow and jump sets

$$C_{\tilde{K}} := [(C_u \times C_z) \cap \tilde{K}] \times \Omega_c \times \Psi \times \Lambda_\theta \tag{25}$$

$$D_{\tilde{K}} := [(D_u \times D_z) \cap \tilde{K}] \times \Omega_c \times \Psi \times \Lambda_\theta. \tag{26}$$

■

The proof of Theorem 5.1 is similar to the proof of [5, Thm. 1]. However, it has three main differences: First, using the NN model-free approximator allows us to eliminate one of the multiple time-scales that emerge in the closed-loop system. Second, the averaging-based step of [5, Thm. 1] is replaced by a singular-perturbation argument that makes use of Proposition 4.2. Third, the amplitude of the dither signal μ does not necessarily have to be small in order to obtain a good approximation of the gradient of the response map J. Instead, the number of neurons in the hidden layer should be sufficiently large. This is in contrast to results such as those in [10], where the number of neurons N do not significantly affect the performance of the system provided the amplitude of the dither signal is sufficiently small.

VI. Numerical Examples

We present two simple numerical examples:

A. SISO Case

Suppose that the plant is a static map such that $y = J(u) := u^2$, which has a global minimum at $u^* = 0$. We want to achieve convergence to u^* in "finite-time", and with an approximately constant evolution rate. To achieve this consider dynamics (22) given by the differential inclusion

$$\dot{u} \in k_1 \cdot \begin{cases} 1 & \text{if } \nabla J(u) < 0 \\ [-1, 1] & \text{if } \nabla J(u) = 0 \\ -1 & \text{if } \nabla J(u) > 0, \end{cases} \tag{27}$$

which simply corresponds to the Krasovskii regularization [20] of the discontinuous gradient method [21]. The dither signal is generated by an oscillator as the one considered in Example 4.1. The vector of basis functions is defined as $\phi = [u^2, u, 1]^\top$. Figure 2 shows the evolution of the vector of weights $\hat{w} = [\hat{w}_1, \hat{w}_2, \hat{w}_3]^\top$ converging to the actual values $w^* = [1, 0, 0]^\top$. The inset shows the evolution in time of the input u converging to the optimal value u^*. The amplitude of the dither signal was 2, and $\Gamma = 500$.

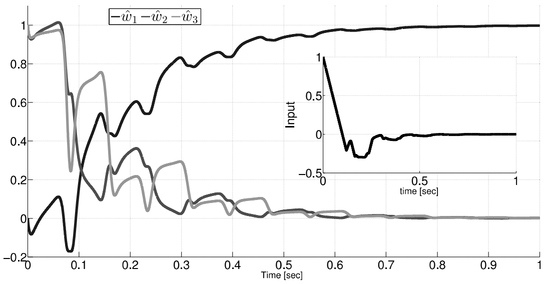

Fig. 2: Evolution in time of $\hat{w}$ and u.

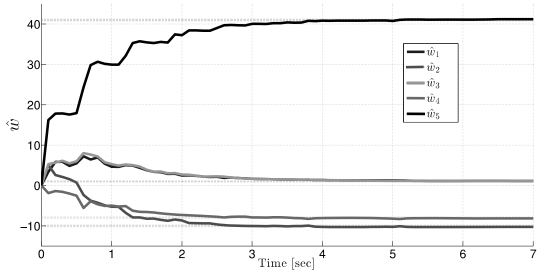

Fig. 3: Evolution in time of $\hat{w}$.

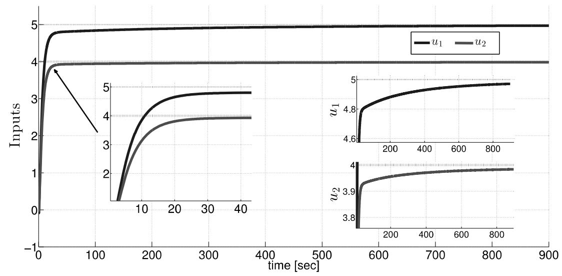

Fig. 4: Evolution in time of u.

B. MISO Case

Now, consider a system with state $\theta \in \mathbb{R}^2$, dynamics $\delta \cdot \dot{\theta} = A\theta + Bu$, and output $y = (\theta_1 - 5)^2 + (\theta_2 - 4)^2$, where A and B are just the identity matrix I, $u \in \mathbb{R}^2$, and $\delta = 5 \times 10^{-4}$. Note that the output is a scalar nonlinear function of both states θ_1 and θ_2. For this system Assumption 3.1 is easily satisfied. Moreover, Assumptions 3.2 and 3.3 are satisfied with $H(u) = u$. The response map is obtained as $J = (u_1 - 5)^2 + (u_2 - 4)^2$, which satisfies Assumption 3.4 with $u^* = [5, 4]^\top$. For this system we use a vector of basis functions given by $\phi = [u_1^2, u_1, u_2^2, u_2, 1]^\top$. Note that although here for simplicity we consider quadratic functions, one could consider other types of polynomial or not polynomial basis functions for non-quadratic response maps. The dither signals are generated by a dynamical system as the one considered in Example 4.1. The dynamics (22) are a standard gradient descent. Figure (3) shows the evolution of the weights w_i of the basis functions. Figure (4) shows the evolution of the control actions converging to the optimal point that minimizes the response map. The amplitude of the dither signals was 2.5, and $\Gamma = 200$.

VII. CONCLUSIONS

This paper presents a novel prescriptive framework for the design of hybrid extremum seeking controllers that used neural networks-based model-free approximators. The approach avoids the use of averaging theory, and does not require a small amplitude in the dither signal in order to estimate the gradient of the response map of the plant under control. The optimization/learning dynamics are allowed to be modeled by general set-valued hybrid systems, which allows for the implementation of discontinuous, set-valued, and hybrid optimization dynamics. Future directions include a rigorous computation and numerical study of the method, as well as its application in a multi-agent system setting.

REFERENCES

[1] K. Ariyur and M. Krstic, *Real-Time Optimization by Extremum-Seeking Control*. Hoboken, NJ: Wiley, 2003.

[2] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking controllers," *Automatica*, vol. 42, no. 6, pp. 889–903, 2006.

[3] C. Zhang and R. Ordóñez, *Extremum-Seeking Control and Applications. A Numerical Optimization-Based Approach*. NY: Springer, 2012.

[4] A. R. Teel and D. Popovíc, "Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming," *In Proc. of American Contr. Conf.*, pp. 2394–2399, 2001.

[5] J. I. Poveda and A. R. Teel, "A framework for a class of hybrid extremum seeking controllers with dynamic inclusions," *Automatica*, vol. 76, pp. 113–126, 2017.

[6] J. I. Poveda and A. R. Teel, "A robust event-triggered approach for fast sampled-data extremization and learning," *IEEE Trans. on Automatic Control, to appear.*, 2017.

[7] A. R. Teel, J. Peuteman, and D. Aeyels, "Semi-global practical asymptotic stability and averaging," *Systems and Control Letters*, vol. 37, pp. 329–334, 1999.

[8] H. Durr, M. S. Stankovic, C. Ebenbauer, and K. H. Johansson, "Lie bracket approximation of extremum seeking systems," *Automatica*, vol. 49, no. 6, pp. 1538–1552, 2013.

[9] W. Wang, A. R. Teel, and D. Nesic, "Analysis for a class of singularly perturbed hybrid system via averaging," *Automatica*, vol. 48, no. 6, pp. 1057–1068, 2012.

[10] M. Chioua, B. Srinivasan, M. Guay, and M. Perrier, "Model adequacy for a precise optimization using extremum seeking control," *in Proc. of European Control Conference*, pp. 513–518, 2016.

[11] A. Favache, D. Dochain, M. Perrier, and M. Guay, "Extremum seeking control of retention for a microparticulate system," *The Canadia Journal of Chemical Engineering*, vol. 86, pp. 815–827, 2008.

[12] M. Guay and D. Dochain, "A time-varying extremum-seeking control approach," *Automatica*, vol. 51, pp. 356–363, 2015.

[13] M. Guay, E. Moshksar, and D. Dochain, "A constrained extremum-seeking control approach," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 16, pp. 3132–3153, 2015.

[14] F. L. Lewis and D. Liu, *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. John Wiley/IEEE Press, Computational Intelligence Series, 2012.

[15] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, pp. 878–888, 2010.

[16] R. Goebel, R. Sanfelice, and A. R. Teel, *Hybrid Dynamical System*. Princeton, NJ: Princeton University Press, 2012.

[17] K. Hornik, S. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural Networks*, vol. 3, pp. 551–560, 1990.

[18] P. Ioannou and J. Sun, *Robust Adaptive Control*. Prentice Hall, 1996.

[19] R. G. Sanfelice and A. R. Teel, "On singular perturbations due to fast actuators in hybrid control systems," *Automatica*, 2011.

[20] A. F. Filippov, *Differential Equations with Discontinuous Right-hand Sides*. Kluwer, 1988.

[21] J. Cortés, "Finite-time convergent gradient flows with applications to network consensus," *Automatica*, vol. 42, no. 11, pp. 1993–2000, 2006.