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

# Data-Enabled Extremum Seeking: A Cooperative Concurrent Learning-Based Approach

Poveda, Jorge; Benosman, Mouhacine; Vamvoudakis, Kyriakos

TR2020-180 January 23, 2021

**Abstract**

This paper introduces a new class of feedback-based data-driven extremum seeking algorithms for the solution of model-free optimization problems in smooth continuous-time dynamical systems. The novelty of the algorithms lies on the incorporation of memory that enables the use of information-rich data sets during the optimization process, and allows to dispense with the time-varying dither excitation signal needed by standard extremum seeking algorithms that rely on a persistence of excitation (PE) condition. The model-free optimization dynamics are developed for single-agent systems, as well as for multi-agent systems with communication graphs that allow agents to share their state information while preserving the privacy of their individual data. In both cases, sufficient richness conditions on the recorded data, as well as suitable optimization dynamics modeled by ordinary differential equations are characterized in order to guarantee convergence to a neighborhood of the solution of the extremum seeking problems. The performance of the algorithms is illustrated via different numerical examples in the context of source seeking problems in multi-vehicle systems.

*International Journal of Adaptive Control and Signal Processing*

© 2021 MERL. This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

# Data-Enabled Extremum Seeking: A Cooperative Concurrent Learning-Based Approach †

Jorge I. Poveda*[1] | M. Benosman[2] | K. Vamvoudakis[3]

[1]Department of Electrical, Computer and Energy Engineering, University of Colorado, Boulder, CO, USA
[2]Mitsubishi Electric Research Laboratories, Cambridge, MA, USA
[3]School of Aerospace Engineering, Georgia Institute of Technology, GA, USA

**Correspondence**
*Jorge I. Poveda, 425 UCB, Boulder , CO 80309-0425, USA. Email: jorge.poveda@colorado.edu

**Present Address**
425 UCB, Boulder , CO 80309-0425, USA.

**Summary**

This paper introduces a new class of feedback-based data-driven extremum seeking algorithms for the solution of model-free optimization problems in smooth continuous-time dynamical systems. The novelty of the algorithms lies on the incorporation of memory that enables the use of information-rich data sets during the optimization process, and allows to dispense with the time-varying dither excitation signal needed by standard extremum seeking algorithms that rely on a persistence of excitation (PE) condition. The model-free optimization dynamics are developed for single-agent systems, as well as for multi-agent systems with communication graphs that allow agents to share their state information while preserving the privacy of their individual data. In both cases, sufficient richness conditions on the recorded data, as well as suitable optimization dynamics modeled by ordinary differential equations are characterized in order to guarantee convergence to a neighborhood of the solution of the extremum seeking problems. The performance of the algorithms is illustrated via different numerical examples in the context of source seeking problems in multi-vehicle systems.

**KEYWORDS:**
Extremum seeking, Data-driven optimization, Multi-agent systems, Concurrent learning.

## 1 | INTRODUCTION

The increasing availability of information-rich data sets and high-performance computing devices has motivated the development of new data-driven algorithms for the solution of estimation, optimization, and feedback control problems[1,2]. Several approaches based on linear and nonlinear parameterizations have successfully exploited these algorithms in applications that range from industrial production to autonomous cars[3]. However, while areas such as machine learning, reinforcement learning, and gradient-free optimization have made significant breakthroughs during the last years, the complex interactions that emerge between physical and digital components in cyber-physical systems have triggered an urgent need to develop novel data-driven algorithms with provable convergence, stability and robustness properties[4,5,6].

In the context of model-free optimization, extremum seeking control has emerged as a powerful technique for the real-time optimization of dynamical systems[7]. Traditionally, ES dynamics have been designed under the paradigm of *exploration vs exploitation*, using an external dither signal to guarantee enough exploration on the cost function, and to facilitate the optimization process via gradient approximations based on parameterizations, averaging techniques, or sampled-data reconstructions. In

†The material in this paper was partially presented at the 11th IFAC Symposium on Nonlinear Control Systems (NOLCOS), Vienna, Austria, 2019, and at the 58th IEEE Conference on Decision and Control, Nice, France, 2019.

all these approaches, the injection of the external signal is needed mainly to guarantee a uniform convergence property in the closed-loop system, i.e., to avoid closed-loop systems with rates of convergence that depend heavily on the initial conditions and which may not be able to recover from external disturbances and/or slows changes in the cost function[8,9]. Because of this, adaptive ES dynamics rely on persistence of excitation (PE) conditions that may be difficult to satisfy in practice, specially in applications where mechanical constraints restrict the persistent excitability of the system, or in high-performance applications where persistent vibrations are undesirable. These limitations have motivated an active line of research that aims to dispense with the PE condition by considering adaptive dynamics that relax the convergence properties[10,11,12], or by incorporating recorded past data into the feedback controller[13,14,15]. In the context of ES, this later approach is particularly appealing for applications where a large amount of recorded data already exists, and in large-scale network multi-agent systems where the nodes of the network are allowed to share some of their information with their neighboring agents. As a matter of fact, online learning and optimization dynamics that incorporate recorded data have become ubiquitous in the context of experience replay[16], iterative learning[17], machine learning[3], and reinforcement learning[18]. However, in the context of ES, they remain mostly unexplored.

## Literature Review

The first modern stability analysis of ES control was presented in[19] using averaging and singular perturbation theory for smooth ordinary differential equations. Later, semi-global practical results were developed in[20] and further extended in[21,22] for a broader class of multi-variable gradient-based optimization dynamics modeled as ODEs and set-valued hybrid dynamical systems with possibly non-unique solutions[23]. Other averaging-based ES algorithms have been studied in[24,25] using Lie bracket averaging, in[26] using systems with delays, and in[27] using discontinuous exploration signals. Sampled-data ES dynamics were initially developed in[28,29] using ideas from nonlinear programming, and were later generalized in[30] and[31] for periodic and aperiodic sampled-data systems with discrete-time set-valued optimization algorithms. In the context of multi-agent systems, different types of distributed ES dynamics have been presented in[32],[33],[34], and[35], to just name a few. The idea of harnessing multiple agents in order to cooperatively and efficiently minimize a common cost function was developed in[35] under the assumption of persistent exploration[35, Assumption 1]. Distributed averaging-based ES dynamics for multi-agent systems were presented in[36] for static graphs, and in[37] for arbitrarily switching graphs.

Following the seminal work of[38], different extremum seeking controllers based on persistence of excitation (PE) conditions have been developed during the last 20 years, including the works[39,40,41] and[42] for single-agent smooth systems, and[43] for multi-variable hybrid systems. Recently, the works[44,45] have developed distributed multi-agent ES dynamics based on PE conditions, and a set-point-based relaxed PE condition with sinusoids was presented in[46].

While all these approaches have provided significant insight into the design of adaptive ES dynamics in different settings, guaranteeing *a priori* the satisfaction of the PE condition along the trajectories of a single-agent or multi-agent system remains a persistent challenge in many applications, e.g., optimization of wind farms, resource allocation in network and wireless systems, and ES problems in mechanical systems with constraints. Moreover, in some cases the persistent excitation induced by the dither signals can also generate undesirable wear and damage in critical mechanical components of the system. Additionally, the existing ES architectures are not designed to exploit information-rich data sets that are now available in several engineering systems such as intelligent transportation systems, healthcare systems, energy systems, heating, ventilation, and air conditioning systems (HVAC), and manufacturing systems, to just name a few. As shown in[13] and[15], the information provided by these data sets can be instrumental in the design of adaptive dynamics that dispense with the traditional PE conditions without sacrificing the robustness properties of the controller. Moreover, as shown in[47], it is possible to further relax restrictive individual excitation conditions in multi-agent systems with communication networks by incorporating cooperative terms into the learning dynamics of the agents. Nevertheless, to the best of our knowledge, the development and analysis of ES dynamics that use current and past data concurrently during the learning process, as well as cooperation between agents in multi-agent systems, are absent in the literature.

## Contributions and Organization of the Paper

Motivated by the previous background, this paper presents a novel class of adaptive ES algorithms for dynamical systems. The ES dynamics dispense with the classic PE condition by exploiting information-rich data sets and cooperation in multi-agent systems (MAS). The proposed algorithms combine ideas from concurrent learning and robust gradient-based optimization dynamics in order to solve real-time optimization problems in network MAS with unknown mathematical models. In particular, the main contributions of this paper are fourfold:

First, we present a new class of data-enabled extremum seeking (DES) dynamics that integrate data into the closed-loop system in order to solve, in a model-free way, a general class of extremum seeking problems formulated as steady state variational inequalities with compact constraints. Since the DES dynamics are data-driven, we provide sufficient conditions on the "richness" of the data and the optimization dynamics in order to guarantee uniform convergence of the trajectories of the algorithm to a neighborhood of the set of solutions of the optimization problem.

Next, since it is unrealistic to assume that in large-scale network multi-agent systems *every* agent satisfies a richness condition on the individual data, we introduce a new class of cooperative data-enabled extremum seeking (CODES) dynamics that integrate data and cooperation between the agents in order to dispense with the standard individual PE assumption. In order to solve the ES problem, we characterize a sufficient condition on the richness of the data of the *overall* network MAS, which can be seen as a spatio-temporal relaxation of the standard persistence of excitation conditions. This spatio-temporal condition merges together past data, i.e., temporal information; and cooperation between the agents, i.e., spatial information.

Third, instead of focusing our attention in one particular optimization algorithm, we characterize a general class of optimization dynamics that can be safely interconnected with the data-driven dynamics in order to solve different types of constrained extremum seeking problems. Moreover, suitable robustness results with respect to noise and small bounded disturbances are established for the closed-loop system.

Finally, we show that the data-enabled dynamics considered in this paper are suitable for applications in multi-vehicle autonomous systems in the context of source seeking. Preliminary results for static maps were reported in the conference papers[48] and[49]. The current paper extends those results by addressing the complete extremum seeking problem in dynamical systems, and by presenting the complete stability and convergence proofs of the algorithms, as well as novel robustness results and a more general formulation of the multi-agent extremum seeking problem.

The rest of this paper is organized as follows: In Section 2 we present some definitions and preliminaries in dynamical systems. In Section 3 we introduce the DES dynamics for single agent systems, as well as the sufficient richness conditions on the data to guarantee convergence to a neighborhood of the set of optimizers. After this, in Section 4 we present the CODES dynamics for network MAS. Section 5 presents the convergence analysis. Section 6 presents some numerical applications, and finally Section 7 ends with the conclusions.

## 2 | PRELIMINARIES

The set of (nonnegative) real numbers is denoted by ($\mathbb{R}_{\geq 0}$) $\mathbb{R}$. The set of (nonnegative) integers is denoted by ($\mathbb{Z}_{\geq 0}$) $\mathbb{Z}$. We use $\mathbb{B}$ to denote a closed unit ball of appropriate dimension, $\rho\mathbb{B}$ to denote a closed ball of radius $\rho > 0$, and $\mathcal{X} + \rho\mathbb{B}$ to denote the union of all sets obtained by taking a closed ball of radius $\rho$ around each point in the set $\mathcal{X}$. We use $\overline{\text{co}}\ \mathcal{X}$ to denote the closed convex hull of $\mathcal{X}$, $\overline{\mathcal{X}}$ to denote its closure, and $\text{int}(\mathcal{X})$ to denote its interior. We use $I_n$ to denote the identify matrix of dimension $n \times n$. Given a vector $x \in \mathbb{R}^n$ and a compact set $\mathcal{A} \subset \mathbb{R}^n$, we use $|x|_{\mathcal{A}} := \inf_{y\in\mathcal{A}} |x - y|$ to denote the minimum distance of $x$ to $\mathcal{A}$, where $|\cdot|$ is the standard Euclidean norm. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be of class $C^k$ if its $k^{th}$ derivative is continuous. A function $\beta(\cdot,\cdot)$ is said to be of class $\mathcal{KL}$ if it is nondecreasing in its first argument, non-increasing in its second argument, $\lim_{r\to 0^+} \beta(r, s) = 0$ for each $s \in \mathbb{R}_{\geq 0}$, and $\lim_{s\to\infty} \beta(r, s) = 0$ for each $r \in \mathbb{R}_{\geq 0}$. The dynamics considered in this paper are modeled by constrained $\varepsilon$-parameterized ODEs of the form

$$x \in C, \quad \dot{x} = F_\varepsilon(x), \tag{1}$$

where $x \in \mathbb{R}^n$ is the overall state, $C \subset \mathbb{R}^n$ is called the flow set, $F_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ is called the flow map, and $\varepsilon > 0$ is a tunable parameter which can be a vector. Throughout this paper we will consider systems of the form (1) with a Lipschitz continuous function $x \mapsto F_\varepsilon(x)$ and a compact set $C$. Following the notation of[23], a continuously differentiable function $x : \text{dom}(x) \to \mathbb{R}^n$ is said to be a solution of (1) if: 1) $x(0) \in C$; and 2) $x(t) \in C$ and $\frac{dx(t)}{dt} = F_\varepsilon(x(t))$ for all $t \in \text{dom}(x)$. System (1) is said to render a compact set $\mathcal{A} \subset \mathbb{R}^n$ uniformly globally asymptotically stable (UGAS) if there exists a $\mathcal{KL}$ function $\beta$ such that $|x(t)|_{\mathcal{A}} \leq \beta(|x(0)|_{\mathcal{A}}, t)$, for all $t \in \text{dom}(x)$ and all $x(0) \in C$. System (1) is said to render a compact set $\mathcal{A} \subset \mathbb{R}^n$ semi-globally practically asymptotically stable (SGPAS) as $\varepsilon \to 0^+$ if there exists a $\mathcal{KL}$ function $\beta$ such that for each pair $\Delta > \nu > 0$ there exists $\varepsilon^* \in \mathbb{R}_{>0}$ such that for each $\varepsilon \in (0, \varepsilon^*)$ every solution of (1) with $|x(0)|_{\mathcal{A}} \leq \Delta$ also satisfies $|x(t)|_{\mathcal{A}} \leq \beta(|x(0)|_{\mathcal{A}}, t) + \nu$, for all $t \in \text{dom}(x)$. If $\text{dom}(x) = [0, \infty)$, the solution $x$ is said to be complete. A set $K \subset \mathbb{R}^n$ is said to be forward invariant for (1) if for all $x(0) \in K$ every solution $x$ of (1) satisfies $x(t) \in K$ for all $t \geq 0$.

## 3 | DATA-ENABLED EXTREMUM SEEKING IN SINGLE-AGENT SYSTEMS

We start by considering the standard extremum seeking problem in single-agent systems, i.e., multivariable systems where the flow of information between states is not restricted by a communication graph. We consider a dynamical system modeled by the following equation

$$\dot{\theta} = f(\theta, z) \tag{2a}$$

$$y = h(\theta, z), \tag{2b}$$

where $f : \mathbb{R}^s \times \mathbb{R}^n \to \mathbb{R}^s$ is a Lipschitz continuous function, $h : \mathbb{R}^s \times \mathbb{R}^n$ is a $C^2$ output function, $\theta \in \mathbb{R}^s$ describes the states of the plant, which are restricted to evolve in a compact set $\Theta \subset \mathbb{R}^s$, and $z \in \mathbb{R}^n$ is the input, which is also restricted to evolve in a pre-defined compact set $\mathcal{F} \subset \mathbb{R}^n$. The sets $\Theta$ and $\mathcal{F}$ can model physical constraints or feasible operational sets, and the functions $f$ and $h$ are assumed to be unknown.

Since the goal in extremum seeking is to optimize the steady-state input-to-output mapping of system (2) using output measurements $y$, we make the following standard stability assumption on the open loop plant dynamics, see also [19], [20], and [22].

**Assumption 1.** There exists a continuous function $\ell : \mathbb{R}^n \to \mathbb{R}^s$ satisfying $\ell(\mathcal{F}) \subset \Theta$ such that the restricted open loop plant

$$(\theta, z) \in \Theta \times \mathcal{F}, \quad \begin{cases} \dot{\theta} = f(\theta, z) \\ \dot{z} = 0 \end{cases} \tag{3}$$

generates complete solutions from every initial condition and renders UGAS the set $H := \{(\theta, z) \in \mathbb{R}^{s+n} : \theta = \ell(z),\ z \in \mathcal{F}\}$.

We define the response map of system (2) as the mapping $\phi : \mathbb{R}^n \to \mathbb{R}$ given by

$$\phi(z) := h(\ell(z), z), \tag{4}$$

where $h$ is the output function of (2) and $\ell$ is the mapping generated by Assumption 1. The mapping $\phi$ is assumed to be unknown. However, for the purpose of analysis we make the following regularity assumption on the pair $(\phi, \mathcal{F})$.

**Assumption 2.** The set $\mathcal{F}$ is compact, convex, and nonempty. The function $\phi$ is convex and continuously differentiable on an open set $D \supset \mathcal{F}$.

Based on Assumptions 1-2, the extremum seeking problem is characterized by the following constrained steady-state optimization problem:

$$\begin{aligned} &\text{minimize} \quad \phi(z) \\ &\text{subject to} \quad z \in \mathcal{F}, \end{aligned} \tag{5}$$

where $z$ is the input to system (2). Since the mathematical models of the plant dynamics (2) and the mapping $\phi$ are unknown, problem (5) needs to be solved in a model-free way.

*Remark 1.* While the smoothness and compactness conditions of Assumption 2 are fundamental for our results, in certain cases it is possible to relax the convexity assumption. However, to simplify our presentation in this paper we assume that (5) is a well-posed convex optimization problem. Extensions to non-convex problems will be developed in the future by using hybrid extremum seeking dynamics [50].

Under Assumption 2, every solution $z^\star$ of problem (5) is also a solution to the following variational inequality (VI):

$$(z - z^\star)^\top \nabla\phi(z^\star) \geq 0, \qquad \forall\, z \in \mathcal{F}, \tag{6}$$

where $\nabla\phi$ is the gradient of $\phi$. Moreover, by [51, Thm 6.12], every point $z^\star$ that satisfies (6) is also a solution of (5). Thus, the constrained convex extremum seeking problem (5) can be equivalently cast as regulating the input of system (2) towards the set

$$\mathcal{A} := \{z^\star \in K : (z - z^\star)^\top \nabla\phi(z^\star) \geq 0,\ \forall\, z \in \mathcal{F}\}, \tag{7}$$

by using only measurements of the output of the plant dynamics. By Assumption 2 and [52, Corolloray 2.2.5], the set $\mathcal{A}$ is nonempty and compact. Moreover, by continuity and Assumptions 1-2, we have that $\lim_{t\to\infty} z(t) \to \mathcal{A} \implies y(t) \to y^\star := \phi(\mathcal{A})$. Thus, regulating the input $z$ towards $\mathcal{A}$ also maximizes the output of the system at steady state.

*Remark 2.* Unlike traditional offline optimization problems, in extremum seeking control it is fundamental to design feedback-based optimization algorithms with suitable stability and robustness properties with respect to noisy measurements and small bounded external disturbances. Because of this reason, traditional numerical optimization approaches with no stability guarantees are not suitable for ES in dynamical systems.

### 3.1 | Uniform Approximation of the Response Map and Its Gradient

In order to solve problem (5), we consider the following approximation of the response map $\phi$ of system (2) on the compact set $\mathcal{F}$:

$$\phi(z) = b(z)^\top w^\star + \epsilon(z), \qquad \forall\, z \in \mathcal{F}, \tag{8}$$

where $\epsilon : \mathbb{R}^n \to \mathbb{R}$ is an approximation error, $w^\star \in \mathbb{R}^p$ is a vector of ideal weights, and $b : \mathbb{R}^n \to \mathbb{R}^p$ is a continuously differentiable pre-defined vector of admissible basis functions. The admissible basis function $b := [b_1, b_2, \dots, b_j, \dots, b_p]^\top$ should be selected such that the functions $b_j : \mathbb{R}^n \to \mathbb{R}$ define a complete independent basis set for $\phi$ in the set $\mathcal{F}$. Typical choices of $b_j$ include polynomial functions, radial basis functions, or sigmoid functions, see[53,54,55] for details on universal approximation properties of different types of functions.

We shall need the following technical assumption on the approximation (8).

**Assumption 3.** For any admissible basis function $b$ and any $p \in \mathbb{Z}_{>0}$ the approximation error $z \mapsto \epsilon(z)$ in (8) is continuously differentiable.

Using the smoothness of $\phi$, $b$, and $\epsilon$, we can compute the gradient of $\nabla\phi$ as follows:

$$\nabla\phi(z) = \nabla b(z)^\top w^\star + \nabla\epsilon(z), \quad \forall\, z \in \mathcal{F}, \tag{9}$$

where $\nabla b$ is the Jacobian matrix of $b$. Since $\mathcal{F}$ is compact and the mappings $\nabla\phi(z)$ and $\nabla\varepsilon(z)$ are continuous, by the Weierstrass high-order approximation theorem[56, Thm. 2.4.11], the approximation errors $\epsilon$ and $\nabla\epsilon$ converge to zero as the number of basis $p$ increases, i.e., $\epsilon(x) \to 0$ and $\nabla\epsilon(z) \to 0$ as $p \to \infty$, uniformly on $\mathcal{F}$. Moreover, due to Assumption 3, the compactness of $\mathcal{F}$ and the fact that $b \in C^2$, we have that $b(z)$, $\nabla b(z)$, $\epsilon(z)$, and $\nabla\epsilon(z)$ are all uniformly bounded in $\mathcal{F}$. Thus, for any $\delta > 0$ there exists $p \in \mathbb{Z}_{>0}$ such that the ideal weights $w^\star$ satisfy

$$\sup_{z\in\mathcal{F}} \left( |\phi(z) - b(z)^\top w^\star| + |\nabla\phi(z) - \nabla b(z)^\top w^\star| \right) \le \delta. \tag{10}$$

To simply our presentation, and without loss of generality, we assume that the optimal weight $w^\star$ in (8) is unique. However, extensions to settings where the optimal weights are not unique and characterized by an optimal compact set could also be considered in our framework.

*Remark 3.* The approximation errors $\epsilon$ and $\nabla\epsilon$ can be made arbitrarily small by increasing the number of basis functions $b_j$ in the approximation (8). However, in certain cases the number of basis functions needed to achieve small approximation errors may be prohibitively large. In order to deal with this limitation, it has been shown in[54] that composite multi-layer basis functions can generate suitable approximations with a smaller set of hyperparameters $w^\star$.

*Remark 4.* Approximations of the form (8) are usually referred to as single-layer *neural networks*, and they have become ubiquitous in neuro-adaptive control[57,55] and approximate reinforcement learning-based control[58,59,60]. Similar approximations have also been studied in the context of adaptive control[61] and extremum seeking control, see for instance[38].

In order to solve problem (5), let $\hat{\phi}$ be an approximation of the response map (8), defined as

$$\hat{\phi}(z) := b(z)^\top \hat{w}, \tag{11}$$

where $\hat{w} \in \mathbb{R}^p$ is an auxiliary state. Let $\tilde{w} := \hat{w} - w^\star$ be the parameter estimation error. Using equations (8) and (11) we can define the approximation error of the response map as

$$\begin{aligned} e_{ss}(z) :&= \hat{\phi}(z) - \phi(z) \\ &= b(z)^\top \hat{w} - b(z)^\top w^\star - \epsilon(z) \\ &= b(z)^\top \tilde{w} - \epsilon(z). \end{aligned} \tag{12}$$

Similarly, the approximation error of the gradient of the response map can be computed as

$$\begin{aligned}\nabla e_{ss}(z) &= \nabla\hat{\phi}(z) - \nabla\phi(z)\\ &= \nabla b(z)^\top \hat{w} - \nabla b(z)^\top w^\star - \nabla\epsilon(z)\\ &= \nabla b(z)^\top \tilde{w} - \nabla\epsilon(z).\end{aligned}$$

Thus, if $\tilde{w} = 0$, the approximation error in the response map and its gradient will be of order $\mathcal{O}(\delta)$, where $\delta > 0$ can be made arbitrarily small by increasing the dimension of the basis vector. However, while it is well-known that several estimation dynamics $\dot{\hat{w}}$ can be implemented to minimize the parameter estimation error, a persistence of excitation (PE) condition needs to be satisfied by the basis functions in order to achieve uniform convergence[62]. Namely, there must exist $T > 0$ and $\gamma > 0$ such that

$$\int_t^{t+T} b(\tau)b(\tau)^\top \mathrm{d}\tau \geq \gamma I, \tag{13}$$

for all $t \geq 0$. Nevertheless, in many practical applications it is difficult to certify a priori the satisfaction of the PE condition (13) along the trajectories of the system. Additionally, in system (2) the function $\phi$ is not available for online measurements.

### 3.2 | Data-Driven Approximation of Response Maps

To dispense with the PE condition (13), and motivated by the increasing number of available information-rich data sets in engineering systems, we consider a class of data-enabled extremum seeking dynamics (DES) that use concurrently real-time *and* recorded measurements of the input and output of system (2). The recorded data used by the DES dynamics is characterized by a finite sequence of inputs and outputs $\{(z_k, y_k)\}_{k=1}^J$, where $k \in \{1, 2, \dots, J\}$ denotes the time index of a data point, i.e., $z_k := z(t_k)$, and $\{t_k\}_{k=1}^J$ is a sequence of measurement times $t_0 \leq t_1 \leq \dots t_J$. Let $b(z_k)$ be the basis function evaluated at the point $z_k$. The estimation error of the *response map* induced by the input data collected at time $t_k$ is given by

$$\begin{aligned}e_{ss}(z_k) &:= \hat{\phi}(z_k) - \phi(z_k),\\ &= \tilde{w}^\top b(z_k) - \epsilon(z_k).\end{aligned} \tag{14}$$

By definition, the parameter estimation error $\tilde{w}$ depends on the current value of $\hat{w}$. Therefore, as a function of time, the response map's estimation error can be written as

$$e_{ss}(t_k, t) = \tilde{w}(t)^\top b\left(z(t_k)\right) - \epsilon\left(z(t_k)\right), \tag{15}$$

for all $t_k \in \{1, 2, \dots, J\}$ and all $t \geq 0$. In order to achieve extremum seeking, we will need input recorded data that is "sufficiently rich". This is formalized by the following definition.

**Definition 1.** A sequence of data $\left\{x_k\right\}_{k=1}^J$, with $x_k \in \mathbb{R}^p$, satisfying the inequality

$$\sum_{k=1}^J x_k x_k^\top \geq \gamma I_p, \tag{16}$$

is said to be $(\gamma, J)$-*sufficiently rich* (SR).

Based on Definition 1, the data $\left\{x_k\right\}_{k=1}^J$ is $(\gamma, J)$-SR if its elements form a basis in $\mathbb{R}^p$ during the window of discrete time $\{1, 2, \dots, J\}$. Indeed, by defining the following *matrix of data*:

$$D := \left[x_1, x_2, \dots, x_k, \dots, x_J\right] \in \mathbb{R}^{p\times J}, \tag{17}$$

we can write the left hand side of inequality (16) as $\sum_{k=1}^J x_k x_k^\top = DD^\top \in \mathbb{R}^{p\times p}$. Since $\mathrm{rank}(DD^\top) = \mathrm{rank}(D)$, it follows that inequality (16) holds for some $\gamma > 0$ if and only if $\mathrm{rank}(D) = p$. Thus, the matrix of data must have as many linearly independent columns as the dimension of the data points. For a given application with available information-rich data, this condition can be verified *a prori*.

*Remark 5.* Condition (16), originally presented in[13], and recently used in[48] for ES in static maps, can be seen as a *finite time* persistently exciting condition. A similar richness condition was used in[63] for dead-beat parameter estimation in model predictive control, and in[2] in the context of data-driven predictive control.

### 3.3 | Data-Enabled Extremum Seeking: Algorithms and Convergence Result

We are now ready to present the data-enabled extremum seeking (DES) dynamics for the solution of problem (5). The DES dynamics are modeled by the following ODE:

$$\dot{\hat{w}} = \frac{\varepsilon_1}{\varepsilon_2} F_{\hat{w}}(\hat{w}, z, y), \tag{18a}$$

$$\dot{z} = \varepsilon_1 F_z(g, z), \tag{18b}$$

where $\hat{w}$ is the auxiliary state used in the approximation (11), $z \in \mathbb{R}^n$ is the input of the plant dynamics (2), and $g$ is a place holder for the vector $g = \nabla b(z)^\top \hat{w}$. The function $F_w$ is defined as:

$$F_{\hat{w}} := -\alpha_1 \Psi(z)(\hat{\phi}(z) - y) - \alpha_2 \sum_{k=1}^{J} \Psi(z_k)(\hat{\phi}(z_k) - y_k), \quad \alpha_1, \alpha_2 > 0, \tag{19}$$

where the mapping $y \mapsto \Psi(y)$ is defined as

$$\Psi(y) := \frac{b(y)}{\left(1 + b(y)^\top b(y)\right)^2}. \tag{20}$$

The pairs $(z, y)$ and $(z_k, y_k)$ correspond to real-time and recorded input-output data of the plant dynamics (2), respectively.

The dynamics of $\hat{w}$, characterized by equation (19), have two main components: The first component is driven by real-time measurements of the output of the plant, and it can be seen as a normalized gradient descent aiming to minimize the square error $e^2$. The second component is driven by the recorded error $e(z_k)$, which depends on the sequence of inputs-outputs $\left\{(z_k, y_k)\right\}_{k=1}^{J}$. As noted in[12], when system (2) is a static map, the data-driven term of (19) can be seen as a type of $\sigma$-modification used to relax the PE condition, see also[8, Ch. 5]. Figure 1 shows a scheme illustrating the DES dynamics interconnected with the plant dynamics (2).

In order to solve the extremum seeking problem (5) in a data-driven way, the input and output data $\left\{(z_k, y_k)\right\}_{k=1}^{J}$ must satisfy the following assumption:

**Assumption 4.** For each $\tilde{\rho} > 0$ there exist input-output data $\{(z_k, y_k)\}_{k=1}^{J}$ of system (2), with $z_k = z(t_k)$ and $y_k = y(t_k)$, satisfying

$$\left| y(t_k) - h(\ell(z(t_k)), z(t_k)) \right| \leq \frac{\tilde{\rho}}{J}, \tag{21}$$

for all $k \in \{1, 2, \ldots, J\}$.

In words, Assumption 4 asks that the data $\{(z_k, y_k)\}_{k=1}^{J}$ used by the DES dynamics must be *consistent*, in the sense that for each $k \in \{1, 2, \ldots, J\}$ the data point $y_k$ is a measurement of the output of (2) that is $\tilde{\rho}$-close to a steady state condition induced by the input $z_k$. By Assumption 1, this type of data can always be collected during a training phase where only a *finite* amount of representative inputs $z_k$ are used to excite the plant dynamics (2) in order to collect output measurements at steady state.

The type of functions $F_z$ used by the DES dynamics in equation (18b) are application dependent and must be designed to stabilize the set of optimizers $\mathcal{A}$ under the assumption of having access to the gradient information of the response map $\phi$:

**Assumption 5.** The constrained ODE

$$z \in \mathcal{F}, \quad \dot{z} = F_z(\nabla\phi(z), z), \tag{22}$$

satisfies:

(a) The function $F_z(\cdot, \cdot)$ is Lipschitz continuous with respect to both arguments.

(b) The set $\mathcal{A}$ is UGAS.

(c) For each bounded continuous function $d : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, and each $z(0) \in \mathcal{F}$, the perturbed system

$$z \in \mathcal{F}, \quad \dot{z} = F_z(\nabla\phi(z) + d, z), \tag{23}$$

generates complete solutions.

In words, Assumption 5 asks that the constrained ODE (22) is a well-posed optimization algorithm with suitable stabilizing properties with respect to the compact set $\mathcal{A}$, and rendering forward invariant the set $\mathcal{F}$ under bounded disturbances on the gradient. While conditions (a) and (b) are common in averaging[21,22] and sampled-data based ES dynamics[30,31], condition (c) is somehow stronger since it requires forward invariance of the set $\mathcal{F}$ under non-necessarily small bounded disturbances on the

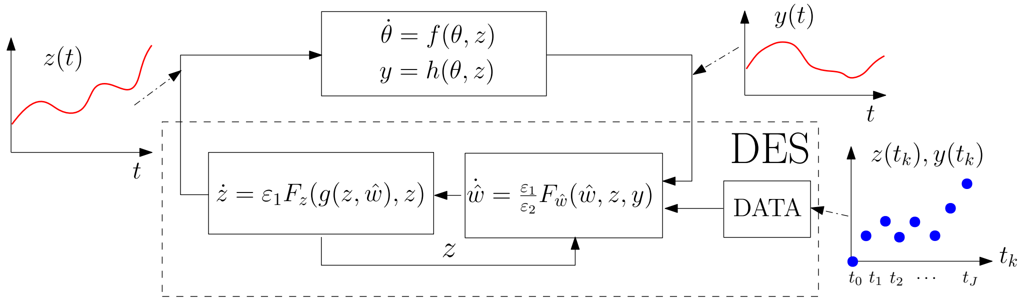


**FIGURE 1** Schematic representation of the data-enabled extremum seeking dynamics for static maps. The feedback-based optimization mechanism implements real-time and past recorded data concurrently during the seeking process.

gradient. Luckily, there exist several algorithms in the literature that satisfy the conditions of Assumption 5, including Lipschitz projected gradient systems that can handle different types of convex optimization problems modeled as VIs of the form (6).

**Example 1** (Lipschitz Projected Gradient Descent)**.** A class of optimization dynamics that satisfy Assumption 5 corresponds to the Lipschitz projected gradient descent[64,65], given by

$$z \in \mathcal{F}, \quad \dot{z} = -z + P_{\mathcal{F}}(z - \nabla\phi(z)). \tag{24}$$

where $P_{\mathcal{F}} : \mathbb{R}^n \to \mathcal{F}$ is the Euclieean Lipschitz projection operator

$$P_{\mathcal{F}}(z) := \arg\min_{y \in \mathcal{F}} |z - y|. \tag{25}$$

As shown in[64], when $\nabla\phi : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth and strongly monotone mapping, and Assumption 2 holds, system (24) renders exponentially stable the set of solutions of the VI (6), which is a singleton under strong monotonicity of $\phi$. Indeed, under Assumption 2, the dynamics (24) render UGAS the compact set $\mathcal{A}$ provided $\nabla\phi : \mathbb{R}^n \to \mathbb{R}^n$ is monotone[65, Thm. 3]. Forward invariance of $\mathcal{F}$ is guaranteed by the projection operator. Simple explicit forms for the projection operator (25) can be computed for common sets $\mathcal{F}$, such as those describing box and sphere constraints.

Optimization problems with coupled equality constraints can also be handled by systems of the form (22).

**Example 2** (Smooth Dynamics for Resource Allocation)**.** Consider a resource allocation problem where the goal is to optimally allocate a resource $R$ (e.g., traffic demand) into $n$ different subsystems (e.g., available routes). In this case, we can define $\mathcal{F} := \left\{z \in \mathbb{R}^n_{\geq 0} : \sum_{i=1}^n z_i = R\right\}$, and we can consider the following Lipschitz continuous dynamics that update the $i^{th}$ component of the state $z$ as follows:

$$z \in \mathcal{F}, \quad \dot{z}_i = \max\left\{0, z^\top \nabla\phi(z) - \frac{\partial\phi(z)}{dz_i}\right\} - \frac{z_i}{R}\sum_{j=1}^n \max\left\{0, z^\top \nabla\phi(z) - \frac{\partial\phi(z)}{dz_j}\right\}. \tag{26}$$

By the results in[66, Ch. 6], these dynamics render forward invariant the set $\mathcal{F}$ under bounded inputs on the gradient, and also render UGAS the set of solutions of the VI (6) whenever $\phi$ is *strictly* convex.

Other approaches that can be used to guarantee forward invariance of compact sets in ODEs include barrier functions, safety functions, or switched gradient flows.

Having characterized the data-driven parameter estimation function (19), as well as the optimization dynamics (22), we are ready to state the first main result of this paper.

**Theorem 1.** Suppose that:

(a) The plant dynamics (2) satisfy Assumptions 1, 2, and 3.

(b) There exists input-output data $\{z_k, y_k\}_{k=1}^J$ satisfying Assumption 4, and the sequence of data $\left\{\tilde{b}(z_k)\right\}_{k=1}^J$ with $k^{th}$ entry given by

$$\tilde{b}(z_k) := \frac{b(z_k)}{1 + b(z_k)^\top b(z_k)}, \tag{27}$$

is $(\gamma, J)$-SR.

(c) The optimization dynamics (18b) satisfy Assumption 5.

Then, for each pair $\Delta > \nu > 0$ there exists $p \in \mathbb{Z}_{>0}$ and $\varepsilon_2^\star \in \mathbb{R}_{>0}$ such that for each $\varepsilon_2 \in (0, \varepsilon_2^\star)$ there exists $\varepsilon_1^\star$ such that for each $\varepsilon_1 \in (0, \varepsilon_1^\star)$ there exists $T \in \mathbb{R}_{>0}$ such that all the trajectories of the closed-loop system given by equations (2) and (18), with $|\hat{w}(0) - w^\star| \leq \Delta$, satisfy

$$|z(t)|_{\mathcal{A}} \leq \nu, \quad |\hat{w}(t) - w^\star| \leq \nu, \quad |\theta(t)|_{\ell(\mathcal{A})} \leq \nu, \quad |y(t) - y^\star| \leq \nu, \tag{28}$$

for all $t \geq T$.

Theorem 1 establishes convergence in finite time to any arbitrarily small $\nu$-neighborhood of the set of optimizers $\mathcal{A}$ and the optimum $y^\star$, from initial conditions on $\Delta$-compact sets defined a priori, provided the parameters $(\varepsilon_1, \varepsilon_2)$ are sufficiently small and appropriately tuned.

*Remark 6.* Unlike the PE condition (13), which applies to the past and future behavior of the functions $b$ evaluated along the trajectories of the system, the richness condition (16) needs to be verified only for a finite *recorded* data $\left\{\tilde{b}(z_k)\right\}_{k=1}^{J}$. This feature allows to exploit information-rich data sets that are available in applications with periodic behaviors and repetitive patterns, e.g., transportation systems, health care systems, manufacturing systems, energy systems, etc. The data can also be obtained by performing repetitive experiments, or by exciting the system during an initial *finite* amount of time.

*Remark 7.* Unlike the standard adaptive ES architectures considered in the literature[38],[46], the DES dynamics do not require the injection of a time-varying dither signal, which, as shown later in Section 6, can generate trajectories with less oscillatory behavior in numerical experiments.

The following Corollary establishes suitable robustness properties for the DES dynamics.

**Corollary 1.** Suppose that all the conditions of Theorem 1 hold, and let $\Delta, \nu, p, \varepsilon_1$ and $\varepsilon_2$ be fixed such that the conclusion of Theorem 1 holds. Then there exists $\rho^\star > 0$ and $T' > 0$ such that for all measurable perturbations $\rho : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ satisfying $\sup_{t\geq 0} |\rho(t)| \leq \rho^\star$ and all solutions of the perturbed dynamics

$$\dot{\theta} = f(\theta + \rho, z + \rho) + \rho \tag{29a}$$

$$\dot{\hat{w}} = \frac{\varepsilon_2}{\varepsilon_1} F_{\hat{w}}(\hat{w} + \rho, b(z + \rho), y + \rho) + \rho, \tag{29b}$$

$$\dot{z} = \varepsilon_1 F_z(g + \rho, z + \rho) + \rho, \tag{29c}$$

with $|\hat{w}(0) - w^\star| \leq \Delta$, the trajectories of the closed-loop system satisfy

$$|z(t)|_{\mathcal{A}} \leq 2\nu, \quad |\hat{w}(t) - w^\star| \leq 2\nu, \quad |\theta(t)|_{\ell(\mathcal{A})} \leq 2\nu, \quad |y(t) - y^\star| \leq 2\nu, \tag{30}$$

for all $t \geq T'$.

Corollary 1 establishes the existence of a strictly positive margin of robustness with respect to noisy state measurements or perturbations on the DES dynamics. As noted in[9], these margins of robustness are critical for the safe implementation of feedback-based algorithms, and they may not exist unless the optimization dynamics satisfy certain regularity and stability properties.

## 4 | COOPERATIVE DATA-ENABLED EXTREMUM SEEKING FOR MULTI-AGENT SYSTEMS

The data-enabled extremum seeking dynamics considered in Section 3 depend on a sequence of data that satisfies the richness condition (16). While this richness condition can be easily verified in small-scale engineering systems such as engines, individual wind turbines, photovoltaic converters, mobile robots, etc, it may be difficult to guarantee the *individual* satisfaction of the richness condition (16) in large-scale multi-agent systems (MAS) comprised of several subsystems with no centralized agent. Motivated by these limitations, we now extend the results of Section 3 to MAS with communication networks characterized by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, N\}$ is the set of vertices or agents, and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of communication links between agents, i.e., the edges. For simplicity, we assume that the communication graph is time-invariant, undirected, connected, and unweighted, i.e., all the entries of the adjacency matrix of the graph satisfy $a_{ij} \in \{0, 1\}$, where $a_{ij} = 0$ if and only if there is no communication link between agents $i$ and $j$.

Each agent of the MAS represents a dynamical system

$$\dot{\theta}_i = f_i(\theta, z_i) \tag{31a}$$
$$y_i = h_i(\theta_i, z_i), \tag{31b}$$

where $f_i : \mathbb{R}^{sN} \times \mathbb{R}^n \to \mathbb{R}^s$ is a Lipschitz continuous function characterizing the dynamics of the $i^{th}$ agent, and $h_i : \mathbb{R}^s \times \mathbb{R}^n$ is a $\mathcal{C}^2$ output function characterizing its output. While the function $f_i$ is written as a general mapping that depends on the overall vector $\theta$, the dynamics of each agent only depend on its individual state $\theta_i$ and the state $\theta_j$ of its neighboring agents $j \in \mathcal{N}_i := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$, that is, the mapping $f_i$ already incorporates the communication graph $\mathcal{G}$.

In order to have a well-defined extremum seeking problem, and similar to Section 3, we assume that the dynamics of the agents are stable and have a well-defined quasi-steady state manifold. In particular, let $z = [z_1^\top, z_2^\top, \dots, z_N^\top]^\top \in \mathbb{R}^{nN}$ be the overall input of the MAS, which is now restricted to a compact set $\mathcal{F}_c \subset \mathbb{R}^{nN}$, and let $\theta := [\theta_1^\top, \theta_2^\top, \dots, \theta_N^\top]^\top$ be the overall state, which is restricted to a set $\Theta_c \subset \mathbb{R}^{nN}$. The following assumption establishes open-loop stability of the MAS under constant inputs implemented by the agents of the system.

**Assumption 6.** There exists a function $\ell_c : \mathbb{R}^{nN} \to \mathbb{R}^{sN}$ satisfying $\ell_c(z) = \ell_{c,1}(z_1) \times \ell_{c,2}(z_2) \times \ldots \times \ell_{c,N}(z_N)$ and $\ell_c(\mathcal{F}) \subset \Theta_c$ such that the open-loop MAS

$$(\theta, z) \in \Theta_c \times \mathcal{F}_c \begin{cases} \dot{\theta} = \tilde{f}(\theta, z) := f_1(\theta, z_1) \times f_2(\theta, z_2) \times \ldots \times f_N(\theta, z_n) \\ \dot{z} = 0 \end{cases} \tag{32}$$

renders UGAS the set $H := \left\{(\theta, z) \in \mathbb{R}^{N(s+n)} : \theta = \ell_c(z),\ z \in \mathcal{F}_c\right\}$.

Based on Assumption 6, the response map $\phi_i : \mathbb{R}^n \to \mathbb{R}$ of each agent $i$ can be defined as

$$\phi_i(z_i) := h_i(\ell_{c,i}(z_i), z_i). \tag{33}$$

The distributed extremum seeking problem considered in this paper is characterized by the following homogeneity assumption on the mappings (33).

**Assumption 7.** The response map $\phi_i$ satisfies $\phi_i = \phi$ for all $i \in \mathcal{V}$, where $\phi : \mathbb{R}^n \to \mathbb{R}$ is a smooth and convex function.

According to Assumption 7, all agents have a response map $\phi$ with identical mathematical form. However, this does not imply that all agents have the same dynamics (31) since different combinations of mappings $f_i$ and $h_i$ can generate the same response map $\phi$.

*Remark 8.* Examples of extremum seeking problems where agents have homogenous response maps include source seeking problems in multi-vehicle systems[35,67], cooperative surveillance with constraints[68], and resource allocation problems in energy systems with identical generators[69].

Let $\mathcal{T} : \mathbb{R}^N \to \mathbb{R}$ be an application-dependent real-valued function known by all agents of the system. The distributed extremum seeking problem is defined as

$$\begin{aligned} &\text{minimize} \quad \mathcal{T}\left(\phi(z_1), \phi(z_2), \dots, \phi(z_N)\right) \\ &\text{subject to} \quad z \in \mathcal{F}_c. \end{aligned} \tag{34}$$

Similar to Section 3, our standing assumption is that the mathematical form of $\phi$ and its gradient $\nabla\phi$ are unknown to all agents. However, agents have access to individual real-time measurements of $\phi$, and are also allowed to share information with their neighboring agents. We make the following regularity assumption on problem (34).

**Assumption 8.** The function $\mathcal{T} : \mathbb{R}^N \to \mathbb{R}$ is smooth and convex, and $\mathcal{F}_c \subset \mathbb{R}^{Nn}$ is compact and convex.

The formulation of problem (34) is quite general, and encompasses distributed coupled and uncoupled optimization problems.

**Example 3** (Source seeking with bounded navigation sets)**.** The problem of locating the source of a signal $\phi$ by using only measurements of its intensity has been studied in[67],[35],[32] and[70] using extremum seeking controllers for single and multi-vehicle systems. Since the signal $\phi$ is homogeneous to all the agents, the source seeking problem can be modeled as (34). Moreover, by defining the feasible set $\mathcal{F}_c$ as $\mathcal{F}_c = \mathcal{F}_1 \times \mathcal{F}_2 \times \ldots \times \mathcal{F}_N$, bounded individual navigation sets $\mathcal{F}_i$ can be assigned to each vehicle of the system. To achieve individual source seeking, the function $\mathcal{T}$ in (34) can be defined as $\mathcal{T}(s_1, s_2, \dots, s_n) := \sum_{i=1}^N s_i$, which guarantees that the solution of (34) is the same as the solution of $N$ uncoupled source seeking problems with individual navigation sets.

**Example 4** (Distributed Resource Allocation)**.** Let us consider again the resource allocation problem of Example 2. Suppose now that each road is an agent controlling its own traffic flow $z_i \in \mathbb{R}$, and interacting with neighboring roads $j \in \mathcal{N}_i$. In this case, we can define again $\mathcal{T}(s_1, s_2, \dots, s_n) := -\sum_{i=1}^{N} s_i$, with $s_i = \phi(z_i)$, and $\mathcal{F}_c = \left\{ z \in \mathbb{R}^N_{\geq 0} \,:\, \sum_{i=1}^{N} z_i = R \right\}$, where $R > 0$ is the available traffic demand to be allocated. Therefore, distributed optimization problems with coupled constraints can also be modeled as (34). However, the existence of a communication graph limiting the flow of information between nodes precludes the implementation of the centralized dynamics (26).

Since it is unrealistic to assume that *every* agent of the MAS has enough information-rich data, as well as access to the states of *all* other agents of the system, the DES dynamics considered in Section 3 are not applicable anymore. Instead, we now consider a class of cooperative data-enabled extremum seeking (CODES) dynamics that will rely on data that is only *cooperatively sufficiently rich*.

## 4.1 | Individual Approximation of Response Maps and Richness of the Network Data

In order to implement the CODES dynamics, each agent runs an individual estimate of the homogenous cost function $\phi$, given by

$$\hat{\phi}_i(z_i) = b_i(z_i)^\top \hat{w}_i, \tag{35}$$

where $\hat{w}_i \in \mathbb{R}^p$ is an auxiliary individual state, and $b_i : \mathbb{R}^n \to \mathbb{R}^p$ is a vector of basis functions. Since $\mathcal{F}_c$ is compact, there exists $M > 0$ such that $\mathcal{F}_c \subset M\mathbb{B}$. Thus, since $\phi$ is smooth and $\mathcal{F}_c$ is compact, by the Weierstrass high-order approximation theorem[53] we know that for any $\delta > 0$ there always exist basis functions $b_i$ and weights $w^\star \in \mathbb{R}^p$ such that

$$\sup_{z_i \in M\mathbb{B}} \left( |\phi(z_i) - b_i(z_i)^\top w^\star| + |\nabla \phi(z_i) - \nabla b(z_i)^\top w^\star| \right) \leq \delta, \tag{36}$$

for all $i \in \{1, 2, \dots, N\}$. In many cases, the simplest way to satisfy this bound is by endowing each agent of the network with the same basis function $b_i$, i.e., $b_i = b_j$ for all $i, j \in \{1, 2, \dots, N\}$. However, this is not a necessary condition since the error approximation bounded by $\delta$ gives room to consider different basis functions that may generate errors with similar bounds using the same ideal weights $w^\star$. Thus, each agent can approximate the cost function $\phi$ as

$$\phi(z_i) = b_i(z_i)^\top w^\star + \epsilon_i(z_i), \tag{37}$$

where $\sup_{z_i \in M\mathbb{B}} |\epsilon_i(z_i)| \leq \delta$. By defining $\tilde{w}_i := \hat{w}_i - w^\star$, the individual response map's estimation error and its gradient can be computed as

$$\begin{aligned} e_{ss,i}(z_i) &:= \hat{\phi}_i(z_i) - \phi(z_i) \\ &= b_i(z_i)^\top \tilde{w}_i - \epsilon(z_i), \end{aligned}$$

and

$$\begin{aligned} \nabla e_{ss,i}(z_i) &= \nabla \hat{\phi}_i(z_i) - \nabla \phi(z_i) \\ &= \nabla b_i(z_i)^\top \tilde{w}_i - \nabla \epsilon_i(z_i). \end{aligned}$$

In order to guarantee *uniform* convergence of $\hat{w}_i$ to $w^\star$ by minimizing the square of the error $e_i$, traditional approaches require restrictive individual PE conditions on the basis functions. On the other hand, if *each* agent of the MAS has data $\{b_i(z_{i,k})\}_{k=1}^{J}$ that satisfies the $(\gamma, J)$-SR condition, the DES dynamics (18) could be individually implemented by the agents in order to solve problem (34). However, this is a restrictive assumption for large-scale MAS since it requires the satisfaction of the full rank condition for N different matrices of data $D_i := \left[\tilde{b}_i(z_{i,1}), \tilde{b}_i(z_{i,2}), \dots, \tilde{b}_i(z_{ik}), \dots, b_i(z_{iJ})\right]$, where $i \in \{1, 2, \dots, N\}$. Therefore, in order to dispense with the standard PE assumption as well as the individual $(\gamma, J)$-SR condition, we consider the following "cooperative" richness condition.

**Definition 2.** A collection of $N$ sequences of data points $\left\{ \{x_{i,k}\}_{k=1}^{J} \,:\, x_{i,k} \in \mathbb{R}^n, i \in \mathcal{V} \right\}$ satisfying the inequality

$$\sum_{k=1}^{J} \sum_{i=1}^{N} x_{i,k} x_{i,k}^\top \geq \gamma I, \tag{38}$$

is said to be $(\gamma, J, N)$-*Cooperative Sufficiently Rich* (CSR).

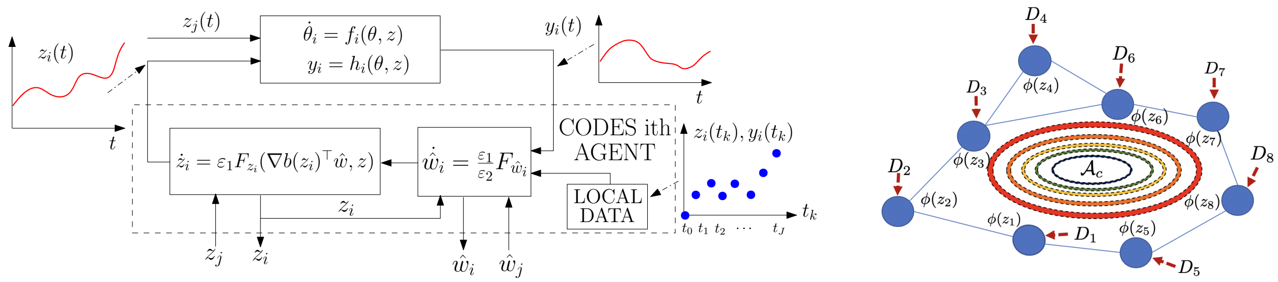


**FIGURE 2** Left. Schematic representation of the cooperative data-enabled extremum seeking dynamics for the $i^{th}$ agent of the MAS. Each agent implements a feedback-based optimization mechanism that uses real-time and past recorded data concurrently. Right. A network MAS where agents have access to individual data $D_i$ and measurements $\phi(z_i)$ of a homogenous cost function $\phi$ whose level sets are shown. Information is shared among the agents via a connected undirected graph.

Condition (38) will guarantee that the data in the *overall* MAS contains "sufficiently" rich information. Since the summation is taken over all agents of the network and over a finite number of times, condition (38) relaxes the $(\gamma, J)$-richness condition of Definition 1, as well as the cooperative PE conditions considered in the literature of adaptive control[47], see Remark 9 below. As a matter of fact, condition (38) can be satisfied even if some agents of the network have no recorded data at all, provided other agents of the network compensate with sufficiently rich data.

**Example 5.** Consider a system with 2 agents using individual basis functions $b_1(z_1(t)) = [\sin(t), 0]^\top$ and $b_2(z_2(t)) = [0, \cos(t)]^\top$. Note that none of these signals satisfy the classic PE condition (13). Consider now a sequence of measurement times $\{t_k\}_k^J$ satisfying $t_k = (k-1)\pi + \pi/4$, for all $k$. Then, the individual sequence of measurements $\{b_i(z_{i,k})\}$ satisfy

$$\sum_{k=1}^{J} b_1(z_1(t_k))b_1^\top(z_1(t_k)) = \begin{bmatrix} \frac{J}{2} & 0 \\ 0 & 0 \end{bmatrix} \text{ for agent 1,} \quad \text{and} \quad \sum_{k=1}^{J} b_2(z_2(t_k))b_2^\top(z_2(t_k)) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{J}{2} \end{bmatrix} \text{ for agent 2,}$$

which are not $(\gamma, J)$-SR for any $\gamma > 0$ and any $J > 0$. However, the overall multi-agent system satisfies condition (38), since

$$\sum_{k=1}^{J}\sum_{i=1}^{N} b_i(z_i(t_k))b_i^\top(z_i(t_k)) = \begin{bmatrix} \frac{J}{2} & 0 \\ 0 & \frac{J}{2} \end{bmatrix} \geq \frac{J}{2} I,$$

for any $J \in \mathbb{Z}_{>1}$.

It is important to note that even if the amount of memory in the nodes of a MAS is unbounded, condition (38) may not necessarily hold if the sampling times or the basis functions are not carefully selected.

**Example 6.** For the same 2-agent system of Example 5, consider now two sequences of data $\{b_i(z_i(t_k))\}_{k=1}^J$, $i \in \{1, 2\}$, which satisfy

$$\sum_{k=1}^{J}\sum_{i=1}^{2} b_i(z_{i,k})b_i^\top(z_{i,k}) = \begin{bmatrix} \sum_{k=1}^{J} \sin(t_k)^2 & 0 \\ 0 & \sum_{k=1}^{J} \cos(t_k)^2 \end{bmatrix}.$$

Then, for any sequence of measurement times $\{t_k\}_{k=1}^J$ satisfying $t_k = (k-1)\pi$ or $t_k = \frac{\pi(2k+1)}{2}$ for all $k$, there is no $\gamma > 0$ and $J \in \mathbb{Z}_{>0}$ such that condition (38) holds.

As shown in Example 6, satisfying the $(\gamma, J, N)$-CSR condition is not trivial even if the amount of data is unbounded. However, suitable training experiments can be designed a priori in order to gather a finite amount of data that satisfies the richness condition (38). These approaches are common in the context of reinforcement learning[71], transfer learning[72], and concurrent learning[18].

## 4.2 | Cooperative Data-Enabled Optimization Dynamics

Let $\tilde{\mathcal{T}}(z) := \mathcal{T}(\phi(z_1), \phi(z_2), \ldots, \phi(z_N))$ and let $\mathcal{A}_c \subset \mathcal{F}_c$ denote the set of solutions of problem (34), that is

$$\mathcal{A}_c := \{z^\star \in \mathcal{F}_c : (z - z^\star)^\top \nabla\tilde{\mathcal{T}}(z^\star) \geq 0, \ \forall\, z \in \mathcal{F}_c\}. \tag{39}$$

In order to guarantee that the overall state of the MAS converges to a neighborhood of $\mathcal{A}_c$ by using only individual data and online measurements of the outputs $y_i$, each agent implements the following cooperative data-enabled extremum seeking (CODES) dynamics:

$$\dot{\hat{\omega}}_i = \frac{\varepsilon_1}{\varepsilon_2} F_{\hat{\omega}_i}(\hat{\omega}_i, \hat{\omega}_j, b_i(z_i), y_i), \tag{40a}$$

$$\dot{z}_i = \varepsilon_1 F_{z_i}(g_i, g_j, z_i, z_j), \tag{40b}$$

where $z_i \in \mathbb{R}^n$ is the input of the $i^{th}$ agent (31), and $g_i$ is a place holder for $g_i = \nabla b_i(z_i)^\top \hat{\omega}_i$. The function $F_{z_i}$ already incorporates the communication graph of the MAS, and the function $F_{\hat{\omega}_i}$ is now defined as follows:

$$F_{\hat{\omega}_i} := -\alpha_1 \Psi_i(z_i)\left(\hat{\phi}_i(z_i) - y_i\right) - \alpha_2 \sum_{k=1}^{J} \Psi_i(z_{i,k})\left(\hat{\phi}_i(z_{i,k}) - y_{i,k}\right) - \alpha_3 \sum_{j \in \mathcal{N}_i} a_{ij}(\hat{\omega}_i - \hat{\omega}_j), \tag{41}$$

where $\alpha_1, \alpha_2, \alpha_3 > 0$ are tunable parameters and where the mapping $y \mapsto \Psi(y)$ is defined as in (20). The dynamics (40a) allow each agent to share their state $\hat{\omega}_i$ with neighboring agents via the last term of equation (41). However, it is important to note that agents do not share their individual data. Thus, the CODES dynamics are suitable for applications where privacy of data is relevant. Figure 2 shows a scheme representing the CODES dynamics and a cooperative ES problem with homogenous cost functions.

The class of functions $F_{z_i}$ characterizing the dynamics of equation (40b) are again application dependent and characterized by the following assumption.

**Assumption 9.** Suppose that the overall MAS implements the following ideal optimizing dynamics:

$$z \in \mathcal{F}_c, \quad \dot{z} = F_z(\nabla\phi, z), \tag{42}$$

where $z = [z_1^\top, z_2^\top, \ldots, z_N^\top]^\top$ and $F_z := F_{z_1} \times F_{z_2} \times \ldots \times F_{z_N}$. Then, system (42) satisfies the conditions of Assumption 5 with respect to the sets $\mathcal{A}_c$ and $\mathcal{F}_c$.

As in Assumption 5, the conditions of Assumption 9 ask for Lipschitz continuity of the mapping $F_z$, UGAS of the compact set $\mathcal{A}_c$, and forward invariance of the set $\mathcal{F}_c$ under bounded disturbances on the gradient.

**Example 7.** For the source seeking problem with bounded navigation sets considered in Example 3, consider a quadratic potential field $\phi$ and simple stable linear vehicle dynamics of the form $\dot{\theta}_i = -A_i\theta_i + B_i u_i$, with $A_i = B_i > 0$. Then, under Assumptions 6, 7, and 8, the projected dynamics (24) can be used to steer the position of the vehicles towards the point that maximizes the intensity $\phi(z_i)$ subject to bounded navigation sets $\mathcal{F}_i$. Thus, satisfying Assumption 9.

**Example 8.** In order to solve in a distributed way the resource allocation problem described in Example 4, we can now consider the distributed dynamics given by

$$\dot{z}_i = \sum_{j \in \mathcal{N}_i} z_j \max\left\{0, \frac{\partial\phi(z_j)}{\partial z_j} - \frac{\partial\phi(z_i)}{\partial z_i}\right\} - z_i \sum_{j \in \mathcal{N}_i} \max\left\{0, \frac{\partial\phi(z_i)}{\partial z_i} - \frac{\partial\phi(z_j)}{\partial z_j}\right\}. \tag{43}$$

As shown in[73, Thm. 3], whenever the function $\phi$ is smooth and bounded, the dynamics (43) render forward invariant the simplex $\mathcal{F}_c = \left\{z_i \in \mathbb{R}^N : \sum_{i=1}^N z_i = R\right\}$. Moreover, they render UGAS the optimal set $\mathcal{A}_c$ whenever $\phi$ is strictly convex.

The following theorem, corresponding to the second main result of this paper, establishes the convergence properties of the CODES dynamics (40) applied to the MAS (31).

**Theorem 2.** Suppose that:

(a) The plant dynamics (31) and problem (34) satisfy Assumptions 6, 7, and 8, and the approximation error $\epsilon_i$ in (37) satisfies the conditions of Assumption 3 for all $i \in \mathcal{V}$.

(b) For each $\tilde{\rho} > 0$ each agent has access to input-output data $\{z_{i,k}, y_{i,k}\}_{k=1}^J$ that satisfies condition (21) for all $k \in \{1, 2, \ldots, J\}$, and the collection of data $\left\{\{\tilde{b}_i(z_k)\}_{k=1}^J : i \in \mathcal{V}\right\}$, with $k^{th}$ entry of the $i^{th}$ sequence given by

$$\tilde{b}_i(z_{i,k}) := \frac{b_i(z_{i,k})}{1 + b_i(z_{i,k})^\top b(z_{i,k})}, \tag{44}$$

is $(\gamma, J, N)$-CSR.

(c) The optimizing dynamics (40b) satisfy Assumption 9.

Then, for each pair $\Delta > \nu > 0$ there exists $p \in \mathbb{Z}_{>0}$ and $\varepsilon_2^\star \in \mathbb{R}_{>0}$ such that for each $\varepsilon_2 \in (0, \varepsilon_2^\star)$ there exists $\varepsilon_1^\star > 0$ such that for all $\varepsilon_1 \in (0, \varepsilon_1^\star)$ there exists a $T \in \mathbb{R}_{>0}$ such that the trajectories of the closed-loop system given by equations (31) and (40) with $|\hat{w}_i(0) - w^\star| \leq \Delta$ for all $i \in \mathcal{V}$, satisfy

$$|z(t)|_{\mathcal{A}_c} \leq \nu, \quad |\hat{w}_i(t) - w_c^\star| \leq \nu, \quad |\theta(t)|_{\ell(\mathcal{A}_c)} \leq \nu, \quad |y_i(t) - y_i^\star| \leq \nu, \tag{45}$$

for all $t \geq T$ and all $i \in \mathcal{V}$.

Theorem 2 says that by selecting a sufficiently large vector of basis functions for each agent, by inducing enough time scale separation in the closed-loop system, and by using data that is *cooperative sufficiently rich* and consistent with input-output behaviors at steady state, the CODES dynamics converge in finite time to an arbitrarily small $\nu$-neighborhood of the optimal set $\mathcal{A}_c$. Moreover, by continuity and stability, the CODES dynamics also satisfy the robustness result of Corollary 1, i.e., there exists a $\rho^\star > 0$ such that any measurable additive perturbation $\rho$ satisfying $\sup_{t\geq 0} |\rho(t)| \leq \rho^\star$, and acting on the states and dynamics of the system, does not dramatically modify the convergence properties of the algorithm.

*Remark 9.* In the context of classic adaptive parameter estimation and stabilization, the work[47] introduced a *cooperative* persistence of excitation condition of the form

$$\int_t^{t+T} \sum_{i=1}^N \tilde{b}_i(\tau)\tilde{b}_i^\top(\tau)\mathrm{d}\tau \geq \gamma I, \tag{46}$$

which has also been used in[74] to study neuro-adaptive learning controllers for multi-agent systems. While the $(\gamma, J, N)$-CSR condition is similar to this excitation condition, inequality (46) needs to be verified for all past and future times. Therefore, the $(\gamma, J, N)$-CSR condition (38) can be seen as a data-driven relaxation of (46) that can be verified a priory.

We finish this section by pointing out that when the number of agents in the MAS is $N = 1$, the CODES dynamics reduce to the DES dynamics and the $(\gamma, J, N)$-CSR condition (38) reduces to the $(\gamma, J)$-SR condition (16). However, when $N > 1$ the requirements on the data and the optimizing dynamics for the CODES dynamics and the DES dynamics are in general different.

## 5 | ANALYSIS

In this section we present the convergence analysis of the DES dynamics and the CODES dynamics. Since some steps are identical, we present the repeated steps only once.

### 5.1 | Analysis of the DES Dynamics

The analysis of the DES dynamics is based on two main parts. First, we will establish suitable convergence properties for the data-driven dynamics (18) under the assumption that the plant (2) is at steady state. After this, we will show that the closed-loop system with plant dynamics is stable provided the parameters $(\varepsilon_1, \varepsilon_2)$ are orderly chosen sufficiently small with respect to the transient behavior of the agents.

The following four lemmas will be instrumental for our results:

**Lemma 1.** Let $b : \mathbb{R}^n \to \mathbb{R}^p$ be a continuous function. Then,

$$|\Psi(z)| = \left| \frac{b(z)}{\left(1 + b(z)^\top b(z)\right)^2} \right| \leq 1 \tag{47}$$

for all $z \in \mathbb{R}^n$.

*Proof.* We have that $|b(z)| \leq 1 + |b(z)|^2$ for all $z \in \mathbb{R}^n$. Adding $|b(z)|^2$ and $|b(z)|^4$ to the right hand side of this inequality we obtain:

$$|b(z)| \leq 1 + 2|b(z)|^2 + |b(z)|^4 = \left(1 + |b(z)|^2\right)^2, \tag{48}$$

for all $z \in \mathbb{R}^n$. Thus,

$$\left| \frac{b(z)}{\left(1 + b(z)^\top b(z)\right)^2} \right| = \frac{1}{\left(1 + |b(z)|^2\right)^2} |b(z)| \leq 1, \tag{49}$$

for all $z \in \mathbb{R}^n$. $\square$

**Lemma 2.** Suppose that Assumption 3 holds and that the recorded data is $(\gamma, J)$-SR. Then, for each pair $(\bar{v}, c) \in \mathbb{R}^2_{>0}$ with $\bar{v} < \sqrt{2c}$ there exists a sufficiently large $p \in \mathbb{Z}_{>0}$ and $\tilde{\rho}^\star > 0$ such that for each $\tilde{\rho} \in (0, \tilde{\rho}^\star)$ the perturbed dynamical system

$$(\hat{w}, x) \in \left(\{w^\star\} + \sqrt{2c}\mathbb{B}\right) \times \mathcal{F}, \quad \begin{cases} \dot{\hat{w}} = -\alpha_1 \Psi(z) \left(\hat{\phi}(z) - \phi(z)\right) - \alpha_2 \sum_{k=1}^{J} \Psi(z_k) \left(\hat{\phi}(z_k) - \phi(z_k) + \tilde{\rho}\right) \\ \dot{z} = 0 \end{cases} \tag{50}$$

renders UGAS a compact set $\mathcal{M} \subset \left(\{w^\star\} + \bar{v}\mathbb{B}\right) \times \mathcal{F}$.

*Proof.* The proof follows similar ideas as the proofs in [13] and [55]. Fix the pair $\bar{v}, c > 0$ and the constants $\alpha_1, \alpha_2 > 0$. Let the pair $(\gamma, J)$ be generated by the $(\gamma, J)$-SR assumption, and define

$$\bar{\alpha} := \max\{\alpha_1, \alpha_2\}, \quad \underline{\alpha} := \min\{\alpha_1, \alpha_2\}. \tag{51}$$

Define the constants $\tilde{\rho}^\star > 0$ and $\delta > 0$ as

$$\delta = \tilde{\rho}^\star = \frac{\bar{v}\gamma\underline{\alpha}}{4(1+J)\bar{\alpha}}, \tag{52}$$

Let the Weierstrass high-order approximation theorem generate sufficiently many basis functions $b_i : \mathbb{R}^n \to \mathbb{R}$, $i \in \{1, 2, \ldots, p\}$, such that (10) holds with $\delta$ given by (52). Define a function $\rho : \mathbb{R}^n \to \mathbb{R}^n$ as:

$$\rho(z) := \alpha_1 \frac{\tilde{b}(z)}{1 + b(z)^\top b(z)} \epsilon(z) + \alpha_2 \sum_{k=1}^{J} \frac{\tilde{b}(z_k)}{1 + b(z_k)^\top b(z_k)} \epsilon(z_k) + \alpha_2 \sum_{k=1}^{J} \Psi(z_k)\tilde{\rho}, \tag{53}$$

where $\tilde{b}$ is defined as in (44). By Lemma 1 and the triangle inequality it follows that

$$\begin{aligned} |\rho(z)| &\leq (1+J)\bar{\alpha}(\delta + \tilde{\rho}) \\ &= (1+J)\bar{\alpha}(2\delta), \end{aligned}$$

for all $z \in \mathcal{F}$. Define the following matrix-valued function:

$$P(z) := \alpha_1 \tilde{b}(z)\tilde{b}(z)^\top + \alpha_2 \sum_{k=1}^{J} \tilde{b}(z_k)\tilde{b}(z_k)^\top. \tag{54}$$

Since the data is $(\gamma, J)$-SR, we have that

$$\sum_{k=1}^{J} \tilde{b}(z_k)\tilde{b}(z_k)^\top \geq \gamma I_p, \tag{55}$$

and since the matrix $\alpha_1 \tilde{b}(z)\tilde{b}(z)^\top$ in (54) is symmetric and positive semidefinite for all $z \in \mathcal{F}$, the matrix-valued function $P$ satisfies

$$P(z(t)) \succ \underline{\alpha}\gamma I_p, \tag{56}$$

along any trajectory $z$ generated by system (50). Let $\tilde{w} = \hat{w} - w^\star$, and consider the error dynamics

$$\begin{aligned} \dot{\tilde{w}} &= -\alpha_1 \tilde{b}(z)\tilde{b}(z)^\top \tilde{w} - \alpha_2 \sum_{k=1}^{J} \tilde{b}(z_k)\tilde{b}(z_k)^\top \tilde{w} + \alpha_1 \frac{\tilde{b}(z)}{1 + b(z)^\top b(z)} \epsilon(z) + \alpha_2 \sum_{k=1}^{J} \frac{\tilde{b}(z_k)}{1 + b(z_k)^\top b(z_k)} \epsilon(z_k) + \alpha_2 \sum_{k=1}^{J} \Psi(z_k)\tilde{\rho}, \\ &= -P(z)\tilde{w} + \rho(z). \end{aligned}$$

Using the following quadratic Lyapunov function

$$V(\tilde{w}) = \frac{1}{2}\tilde{w}^\top \tilde{w}, \tag{57}$$

we obtain that the derivative of $V$ along the solutions of (50) satisfies:

$$\begin{aligned} \dot{V} &= -\tilde{w}^\top P(z)\tilde{w} + \tilde{w}^\top \rho(z), \\ &\leq -\underline{\alpha}\gamma|\tilde{w}|^2 + |\tilde{w}|(1+J)\bar{\alpha}(2\delta) \\ &\leq -\frac{1}{2}\underline{\alpha}\gamma|\tilde{w}|^2, \quad \forall\, |\tilde{w}| \geq \frac{4(1+J)\delta\bar{\alpha}}{\underline{\alpha}\gamma} = \bar{v}. \end{aligned} \tag{58}$$

Since $|\tilde{w}| \leq \bar{v}$ implies that $V(\tilde{w}) \leq 0.5\bar{v}^2$, it follows that for any $c \geq 0.5\bar{v}^2$ we have $\bar{v}\mathbb{B} \subset L_c$ where $L_c := \{\tilde{w} \in \mathbb{R}^p : V(\tilde{w}) \leq c\}$. Thus, for any $c \geq 0.5\bar{v}^2$ the sets $L_c$ are forward invariant. Therefore, every solution of (50) is complete, and by the inequality (58) there exists $T > 0$ such that $[\tilde{w}(t), z(t)]^\top \in \bar{v}\mathbb{B} \times \mathcal{F}$ for all $t \geq T$, i.e., the trajectories of $\tilde{w}$ are uniformly ultimately bounded.

By using[23, Corollary 7.7] we can conclude the existence of a uniformly globally asymptotically stable set $\mathcal{M} \subset \left(\{w^\star\} + \bar{v}\mathbb{B}\right) \times \mathcal{F}$ for the constrained dynamics (50). This establishes the result. $\square$

**Lemma 3.** Suppose that Assumption 5 holds and consider the perturbed optimization dynamics

$$z \in K, \quad \dot{z} = F_z\left(\nabla\phi(z) + \mathcal{O}(\bar{\delta}), z\right), \tag{59}$$

where $\bar{\delta} > 0$. Then, for each $v > 0$ there exists $\bar{\delta}^\star > 0$ such that for all $\bar{\delta} \in (0, \bar{\delta}^\star)$ every solution is complete and there exists a UGAS compact set $\Omega \subset \mathcal{A} + v\mathbb{B}$.

*Proof.* Let $\hat{F}_z(z) := F_z(\nabla\phi(z), z)$. By items (a) and (b) in Assumption 5, and[23, Lemma 7.20], the inflated system

$$\dot{z} \in F_\rho(z) := \overline{\mathrm{co}}\,\hat{F}_z\left((z + \rho\mathbb{B}) \cap \mathcal{F}\right) + \rho\mathbb{B}, \tag{60}$$

renders the set $\mathcal{A}$ SGPAS as $\rho \to 0^+$. Moreover, by[22, Lemma 3], for each $\rho > 0$ there exist $\bar{\delta}^\star > 0$ sufficiently small such that

$$F_z\left(\nabla\phi(z) + \bar{\delta}\mathbb{B}, z\right) \subset F_\rho(z), \quad \forall\ z \in \mathcal{F}. \tag{61}$$

Thus, for any $\bar{\delta} \in (0, \bar{\delta}^\star)$ every solution of the perturbed system (59) is also a solution of the perturbed differential inclusion (60), which implies that for any $v > 0$ there exists $\bar{\delta}^{\star\star} > 0$ such that for any $\bar{\delta} \in (0, \bar{\delta}^{\star\star})$ every solution of (59) satisfies the following bound

$$|z(t)|_{\mathcal{A}} \leq \beta(|z(0)|_{\mathcal{A}}, t) + v, \tag{62}$$

for all $t \in \mathrm{dom}(z)$ and for some $\beta \in \mathcal{KL}$. By item (c) in Assumption 5 it follows that $\mathrm{dom}(z) = [0, \infty)$. By[23, Corollary 7.7], we obtain the existence of a compact set $\Omega \subset \mathcal{A} + v\mathbb{B}$ that is UGAS for system (59). This establishes the result. $\square$

**Lemma 4.** Suppose that Assumptions 2, 3, 4 and 5 hold, and that the sequence of normalized data $\{\tilde{b}(z_k)\}_{k=1}^J$ with $k^{th}$ entry given by (44) is $(\gamma, J)$-SR. Then, for each pair $(\bar{v}, c) \in \mathbb{R}^2_{>0}$ with $\bar{v} < \sqrt{2c}$ there exists $p \in \mathbb{Z}_{>0}$ and $\tilde{\rho}^\star > 0$ such that for each $\tilde{\rho} \in (0, \tilde{\rho}^\star)$ there exists $\varepsilon_2^\star \in \mathbb{R}_{>0}$ such that for each $\varepsilon_2 \in (0, \varepsilon_2^\star)$ the dynamics

$$\frac{d\hat{w}}{d\tau} = -\frac{1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - \phi(z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - \phi(z_k) + \tilde{\rho}\right)\right). \tag{63a}$$

$$\frac{dz}{d\tau} = F_z\left(\nabla b(z)^\top\hat{w}, z\right), \tag{63b}$$

constrained to the set $(\{w^\star\} + c\mathbb{B}) \times \mathcal{F}$ render UGAS a compact set $\Omega'_{\varepsilon_2,\tilde{\rho}} \subset (\{w^\star\} + c\mathbb{B}) \times (\mathcal{A} + v\mathbb{B})$.

*Proof.* System (63) is a singularly perturbed system in normal form[75]. Its *boundary layer dynamics* are given by

$$\dot{\hat{w}} = -\frac{1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z_k) - \phi(z_k)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - \phi(z_k) + \tilde{\rho}\right)\right), \tag{64a}$$

$$\dot{z} = 0. \tag{64b}$$

By Lemma 2, the dynamics (64) constrained to the sets $(\{w^\star\} + c\mathbb{B}) \times \mathcal{F}$ render UGAS a set $\mathcal{M} \subset \left(\{w^\star\} + \bar{v}\mathbb{B}\right) \times \mathcal{F}$. Therefore, the quasi-steady state value of $\hat{w}$ satisfies $|\hat{w}_{ss} - w^\star| \leq \bar{v}$. The *reduced dynamics* of (63) are obtained by substituting $\hat{w}$ in (63b) by its quasi steady-state value $\hat{w}_{ss} = w^\star + \mathcal{O}(\bar{v})$, i.e.,

$$\begin{aligned} z \in \mathcal{F}, \quad \dot{z} &= F_z\left(\nabla b(z)^\top w^\star + \mathcal{O}(\bar{v}), z\right), \\ &= F_z\left(\nabla\phi(z) + \mathcal{O}(\bar{v} + \delta), z\right). \\ &= F_z\left(\nabla\phi(z) + \mathcal{O}\left(\bar{v} + \frac{\bar{v}\gamma\underline{\alpha}}{4(1+J)\bar{\alpha}}\right), z\right). \end{aligned}$$

where the last equality follows by (52). Using the definition of $\bar{v}$ in (58) we obtain:

$$\bar{v} + \frac{\bar{v}\gamma\underline{\alpha}}{4(1+J)\bar{\alpha}} = \bar{v}\left(\frac{4(1+J)\bar{\alpha} + \gamma\underline{\alpha}}{4(1+J)\bar{\alpha}}\right) \tag{65}$$

$$= \frac{4(1+J)\bar{\alpha}}{4(1+J)\bar{\alpha} + \gamma\underline{\alpha}}\left(\frac{4(1+J)\bar{\alpha} + \gamma\underline{\alpha}}{4(1+J)\bar{\alpha}}\right)\bar{\delta} \tag{66}$$

$$= \bar{\delta} \tag{67}$$

Therefore, the reduced dynamics are given by

$$z \in \mathcal{F}, \quad \dot{z} = F_z\left(\nabla\phi(z) + \mathcal{O}(\bar{\delta}), z\right).$$

Since $\bar{\delta}$ was generated by Lemma 3, it follows that the reduced dynamics render UGAS a compact set $\Omega \subset \mathcal{A} + \nu\mathbb{B}$. By singular perturbation theory[75, Thm. 2] it follows that the original dynamics (63) render SGPAS as $\varepsilon \to 0^+$ the set $(\{w^\star\} + c\mathbb{B}) \times \Omega \subset (\{w^\star\} + c\mathbb{B}) \times (\mathcal{A} + \nu\mathbb{B})$. Therefore, every complete solution of (63) generates trajectories $z$ that converge to $\mathcal{A} + \nu\mathbb{B}$ before some finite time $T > 0$. Completness of solutions follows by the forward invariance properties of the set $(\{w^\star\} + c\mathbb{B}) \times \mathcal{F}$. By[23, Corollary 7.7] we obtain the existence of a UGAS compact set $\Omega'_{\varepsilon_2,\tilde{\rho}} \subset (\{w^\star\} + c\mathbb{B}) \times (\mathcal{A} + \nu\mathbb{B})$. □

## Proof of Theorem 1 and Corollary 1

Fix $\Delta > \nu > 0$, and let $\alpha_1, \alpha_2 > 0$. Let $c = \Delta$ and let Lemma 3 generate $\bar{\delta}^\star > 0$. Let $\bar{\delta} \in (0, \bar{\delta}^\star)$ be sufficiently small such that $\frac{4(1+J)\bar{\alpha}}{4(1+J)+\gamma\underline{\alpha}}\bar{\delta} < \sqrt{2c}$, where the pair $(\gamma, J)$ comes from the $(\gamma, J)$-SR assumption. Define the constant $\bar{\nu} := \frac{4(1+J)\bar{\alpha}}{4(1+J)+\gamma\underline{\alpha}}\bar{\delta}$ and consider the closed-loop system dynamics

$$\dot{\theta} = f(\theta, z) \tag{68a}$$

$$\dot{\hat{w}} = -\frac{\varepsilon_1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - h(\theta, z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - y_k\right)\right) \tag{68b}$$

$$\dot{z} = \varepsilon_1 F_z\left(\nabla b(z)^\top \hat{w}, z\right), \tag{68c}$$

constrained to the set $\Theta \times (\{w^\star\} + c\mathbb{B}) \times \mathcal{F}$. Consider the change of time scale induced by the change of variable $\tau = t\varepsilon_1$, which generates the following dynamics in the $\tau$-time scale

$$\frac{d\theta}{d\tau} = \frac{1}{\varepsilon_1}f(\theta, z) \tag{69a}$$

$$\frac{d\hat{w}}{d\tau} = -\frac{1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - h(\theta, z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - y_k\right)\right) \tag{69b}$$

$$\frac{dz}{d\tau} = F_z\left(\nabla b(z)^\top \hat{w}, z\right), \tag{69c}$$

For values of $\varepsilon_1 > 0$ sufficiently small, this system is a singularly perturbed system with fast dynamics given by equation (69a), and slow dynamics given by equation (69b)-(69c). The *boundary layer dynamics* of this system are obtained by setting $\varepsilon_1 = 0$ in (68), i.e.,

$$\dot{\theta} = f(\theta, z), \quad \dot{\hat{w}} = 0, \quad \dot{z} = 0. \tag{70}$$

By Assumption 1, for each fixed pair $(\hat{w}, z)$, the plant dynamics in (70) render globally asymptotically stable the quasi-steady state manifold $\theta^\star = \ell(z)$. Therefore, the dynamics (69) have a well-defined *reduced system*, obtained by substituting $\theta = \ell(z)$ in equation (69b):

$$\frac{d\hat{w}}{d\tau} = -\frac{1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - h(\ell(z), z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - y_k\right)\right) \tag{71a}$$

$$\frac{dz}{d\tau} = F_z\left(\nabla b(z)^\top \hat{w}, z\right). \tag{71b}$$

Using equation (4):

$$\frac{d\hat{w}}{d\tau} = -\frac{1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - \phi(z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - \phi(z_k)\right) + \tilde{\rho}\right) \tag{72a}$$

$$\frac{dz}{d\tau} = F_z\left(\nabla b(z)^\top \hat{w}, z\right), \tag{72b}$$

where $\tilde{\rho} := \sum_{k=1}^{J}\Psi(z_k)(\phi(z_k) + y_k)$, which, by Lemma 1 and Assumption 4 satisfies $|\tilde{\rho}| \leq \varepsilon$. Therefore, system (72) corresponds to the same dynamics (63) studied in Lemma 4, and the convergence result of the Theorem follows directly by combining the stability result of Lemma 4 with the stability properties of the open-loop plant dynamics (70), and singular perturbation theory[75, Thm. 2]. Corollary 1 follows now directly by the continuity and stability properties of the DES dynamics, and by[23, Lemma 7.20].

## 5.2 | Analysis of the CODES Dynamics

The analysis of the CODES dynamics (40) follows similar ideas to the analysis of the DES dynamics (18): First, we establish suitable convergence properties for the learning dynamics (2) under condition (b) of Theorem 2. After this, the dynamics (40) are analyzed under the assumption that the outputs $\phi_i$ are at steady state. Finally, we study the stability of the closed-loop system by using singular perturbation theory. Since some of the steps are identical to the proof of Theorem 1, we present only the new technical lemmas needed to establish the result.

Let $\tilde{w}_i = \hat{w}_i - w^\star$ be the individual parameter estimation error and let $y_{ik} = \phi(z_{i,k}) + \mathcal{O}(\tilde{\rho})$. Using (35)-(37) the parameter estimation error dynamics of each agent are given by

$$\dot{\tilde{w}}_i = -\alpha_1 \tilde{b}_i(z_i)\tilde{b}_i(z_i)^\top \tilde{w}_i - \alpha_2 \sum_{k=1}^{J} \tilde{b}_i(z_{i,k})\tilde{b}_i(z_{i,k})^\top \tilde{w}_i - \alpha_3 \sum_{j\in\mathcal{N}_i} a_{ij}(\tilde{w}_i - \tilde{w}_j) + \rho_i(z_i). \tag{73}$$

where $z_i : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is generated by the dynamics (40b), and $\rho(z_i)$ is given by

$$\rho_i(z_i) := \alpha_1 \frac{\tilde{b}_i(z_i)}{b_i(z_i)^\top b_i(z_i) + 1}\epsilon(z_i) + \alpha_2 \sum_{k=1}^{J} \frac{\tilde{b}_i(x_{ik})}{b_i(z_{ik})^\top b_i(z_{ik}) + +1}\epsilon(z_{i,k}) + \alpha_2 \sum_{k=1}^{J} \Psi(z_k)\tilde{\rho}. \tag{74}$$

Define $\tilde{w} := [\tilde{w}_1^\top, \tilde{w}_2^\top, \dots, \tilde{w}_N^\top]^\top$, $\tilde{B}(z) := \text{diag}\left\{\tilde{b}_1(z_1), \tilde{b}_2(z_2), \dots, \tilde{b}_N(z_N)\right\}$, $\tilde{B}_k := \text{diag}\{\tilde{b}_1(z_{1k}), \tilde{b}_2(z_{2,k}), \dots, \tilde{b}_N(z_{N,k})\}$, and $\rho(z) := [\rho_1(z_1)^\top, \rho_2(z_2)^\top, \dots, \rho_N(z_N)^\top]^\top$, which leads to the error estimation dynamics in vectorial form:

$$\begin{aligned} \dot{\tilde{w}} &= -\alpha_1 \tilde{B}(z)\tilde{B}(z)^\top \tilde{w} - \alpha_2 \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k \tilde{w} - (\alpha_3 \mathcal{L} \otimes I_p)\tilde{w} + \rho(x), \\ &= -\left[\alpha_1 \tilde{B}(z)\tilde{B}(z)^\top + \alpha_2 \sum_{k=1}^{\bar{k}} \tilde{B}_k \tilde{B}_k^\top + (\alpha_3 \mathcal{L} \otimes I_p)\right]\tilde{w}(t) + \rho(x) \\ &= -\Omega(z)\tilde{w} + \rho(z), \end{aligned} \tag{75}$$

where

$$\Omega(z) := \alpha_1 \tilde{B}(z)\tilde{B}(z)^\top + \alpha_2 \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k^\top + \alpha_3(\mathcal{L} \otimes I_p). \tag{76}$$

The following two lemmas characterizes the convergence properties of the error dynamics (75).

**Lemma 5.** Let $z : \mathbb{R}_{\geq 0} \to \mathbb{R}^{Nn}$ be a continuous function, and suppose that condition (b) of Theorem 2 holds. Then, there exists $\delta_1, \delta_2 \in \mathbb{R}_{>0}$ such that

$$\delta_1 \bar{\alpha} I_{Np} \geq \Omega(z(t)) \geq \delta_2 \underline{\alpha} I_{Np} \tag{77}$$

for all $t \geq 0$, where $\bar{\alpha} := \max\{\alpha_1, \alpha_2, \alpha_3\}$, and $\underline{\alpha} := \min\{\alpha_1, \alpha_2, \alpha_3\}$.

*Proof.* The proof combines ideas from the proofs of[47, Theorem 1] and [13]. Since the graph is connected and undirected, the Laplacian matrix $\mathcal{L}$ is symmetric and positive semidefinite. Similarly, the matrices $\alpha_1\Phi(z)\Phi(z)^\top$ and $(\alpha_3\mathcal{L}\otimes I_p)$ are symmetric and positive semidefinite for any pair $(\alpha_1, \alpha_3) \in \mathbb{R}^2_{>0}$. Moreover, the matrix $\mathcal{L} \otimes I_p$ has only $p$ zero eigenvalues, whose orthogonal unit eigenvectors are $v_i = \frac{1}{\sqrt{N}}\mathbf{1} \otimes e_i$, for all $i \in \{1, 2, \dots, p\}$, with $e_i \in \mathbb{R}^N$ being the unitary vector with nonzero element at the $i^{th}$ entry. The orthogonal unit eigenvectors associated to the positive eigenvalues $\{\lambda_{p+1}, \dots, \lambda_{Np}\}$ are denoted as $v_{p+1}, \dots, v_{Np}$. Since the eigenvectors of the symmetric matrix $\alpha_3\mathcal{L} \otimes I_p$ form an orthogonal basis in $\mathbb{R}^{Np}$, any vector $y \in \mathbb{R}^{Np}$ can be written as $y = \sum_{i=1}^{p} c_i v_i + \sum_{i=p+1}^{Np} c_i v_i$, with $c_i \in \mathbb{R}$ for all $i$, where without loss of generality we consider unitary vectors $y \in \mathbb{R}^{Np}$. We then have two possible cases:

(a) Suppose that $\sum_{i=p+1}^{Np} c_i^2 \neq 0$, then

$$y^\top \Omega(z) y = y^\top \left[\alpha_1 \tilde{B}(z)\tilde{B}(z)^\top + \alpha_2 \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k^\top\right] y + y^\top \left(\alpha_3 \mathcal{L} \otimes I_p\right) y. \tag{78}$$

Using spectral decomposition of $(\alpha_3 \mathcal{L} \otimes I_p)$, the second term of (78) reduces to $\sum_{i=p+1}^{Np} \alpha_3 c_i^2 \lambda_i$. Let $\lambda_2$ be the smallest positive eigenvalue of the Laplacian matrix $\mathcal{L}$. It follows that $\sum_{i=p+1}^{Np} \alpha_3 c_i^2 \lambda_i \geq \alpha_3 \lambda_2 \sum_{i=p+1}^{Np} c_i^2$, and (78) satisfies

$$y^\top \Omega(z) y \geq \alpha_3 \lambda_2 \sum_{i=p+1}^{Np} c_i^2 > 0, \tag{79}$$

for all $z \in \mathbb{R}^{nN}$ and all $y \in \mathbb{R}^{Np}$, where the last inequality follows by the assumption that $\sum_{i=p+1}^{Np} c_i^2 \neq 0$. By the contradiction argument of[47, Eq. (29)-(30)] it follows that $\Omega(z(t))$ is uniformly positive definite, i.e., the eigenvalues of $\Omega(z(t))$ have a uniform positive lower bound that holds for all $t \geq t_0$ and $t_0 \in \mathbb{R}_{\geq 0}$.

(b) Suppose that $\sum_{i=p+1}^{Np} c_i^2 = 0$, which implies that $\sum_{i=1}^{Np} c_i^2 = 1$. Then, it must be the case that $\sum_{i=1}^{p} c_i^2 \neq 0$. Using $y = \sum_{i=1}^{p} c_i v_i$, we obtain

$$y^\top \Omega(z) y = \alpha_1 y^\top \tilde{B}(z)\tilde{B}(z)^\top y + \alpha_2 y^\top \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k y. \tag{80}$$

Expanding $z$ in the second term of (80) we obtain

$$\alpha_2 \left( \sum_{i=1}^{p} c_i v_i \right)^\top \left[ \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k^\top \right] \left( \sum_{i=1}^{p} c_i v_i \right), \tag{81}$$

which can be written as $\alpha_2 C^\top V^\top \left[ \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k^\top \right] V C$, where $C := [c_1, c_2, \ldots, c_m]^\top$ and $V := [v_1, v_2, \ldots, v_m]$. Since $v_i = \frac{1}{\sqrt{N}} \mathbf{1} \otimes e_i$, we obtain

$$V^\top \left[ \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k^\top \right] V = \sum_{k=1}^{J} V^\top [\tilde{B}_k \tilde{B}_k^\top] V = \sum_{k=1}^{J} \sum_{i=1}^{N} \tilde{b}_{i,k} \tilde{b}_{i,k}^\top \geq \underline{\varepsilon} I_p, \tag{82}$$

for some $\underline{\varepsilon} \in \mathbb{R}_{>0}$, where the inequality follows by the $(\gamma, J, N)$-CSR condition. Using (81) and (82) we obtain

$$\alpha_2 C^\top V^\top \left[ \sum_{k=1}^{J} \tilde{B}_k \tilde{B}_k^\top \right] V C \geq \alpha_2 \underline{\varepsilon} \sum_{i=1}^{p} c_i^2 > 0 \tag{83}$$

which implies that $y^\top \Omega(z) y$ given by (80) is uniformly positive definite for all $y \in \mathbb{R}^{Np}$ and $z \in \mathbb{R}^{Nn}$.

The conclusion of the two cases establishes the existence of the pair $\delta_1, \delta_2$ in (77). □

**Lemma 6.** Suppose that condition (b) of Theorem 2 holds. Then, for each pair $(\bar{v}, c) \in \mathbb{R}^2_{>0}$ with $\bar{v} < \sqrt{2c}$ there exists a sufficiently large $p \in \mathbb{Z}_{>0}$ such that the constrained dynamical system

$$(\tilde{w}, x) \in \sqrt{2c}\mathbb{B} \times \mathcal{F}_c, \quad \begin{cases} \dot{\tilde{w}} = -\Omega(z)\tilde{w} + \rho(z), \\ \dot{z} = 0 \end{cases} \tag{84}$$

renders UGAS a compact set $\mathcal{M}_c \subset \bar{v}\mathbb{B} \times \mathcal{F}_c$.

*Proof.* Fix the pair $\bar{v}, c > 0$ and the constants $\alpha_1, \alpha_2, \alpha_3 > 0$. Let the Weierstrass high-order approximation theorem generate sufficiently many basis functions $b_{i\ell} : \mathbb{R}^n \to \mathbb{R}$, $\ell \in \{1, 2, \ldots, p\}$, for each agent $i \in \mathcal{V}$, such that (36) holds for all agents with

$$\delta = \frac{\bar{v}\varepsilon_2 \underline{\alpha}}{4\sqrt{N}(1+J)\bar{\alpha}}, \tag{85}$$

where the constants $\gamma, J$ are generated by the $(\gamma, J, N)$-CSR condition on the data. Let $\tilde{\rho}^\star = \delta$ and consider the quadratic Lyapunov function $V(\tilde{w}) = 0.5\tilde{w}^\top \tilde{w}$, which is positive definite and radially unbounded. The derivative $V$ along the solutions of (84) satisfies

$$\begin{aligned} \dot{V} &\leq -\varepsilon_2 \underline{\alpha} |\tilde{w}|^2 + \tilde{w}^\top \rho(x), \\ &\leq -\varepsilon_2 \underline{\alpha} |\tilde{w}|^2 + |\tilde{w}||\rho(x)|, \end{aligned}$$

where we used the lower bound of (77). Since $|\rho_i(z_i)| \leq (1+J)\bar{\alpha}\delta$ for all $i \in \mathcal{V}$, it follows that $|\rho(z)| \leq \sqrt{N}(1+J)\bar{\alpha}\delta$. Thus, $\dot{V}$ satisfies

$$\begin{aligned} \dot{V} &\leq -\varepsilon_2\underline{\alpha}|\tilde{w}|^2 + |\tilde{w}|\sqrt{N}(1+J)\bar{\alpha}2\delta, \\ &\leq -\frac{1}{2}\varepsilon_2\underline{\alpha}|\tilde{w}|^2, \quad \forall\, |\tilde{w}| \geq \frac{2\sqrt{N}(1+J)\bar{\alpha}2\delta}{\varepsilon_2\underline{\alpha}} = \bar{v}. \end{aligned}$$

where the last equality follows by the definition of $\delta$. This establishes forward invariance of the level sets $L_c := \{\tilde{w} \in \mathbb{R}^{Np} : V(\tilde{w}) \leq c\}$ for $c > 0.5\bar{v}^2$, and uniform ultimate boundedness of the trajectories $\tilde{w}$, with a uniform ultimate bound satisfying $|\tilde{w}| \leq \bar{v}$. Since every solution of (84) with $(\tilde{w}(0), x(0)) \in L_c \times \mathcal{F}_c$ is complete, by[23, Corollary 7.7] there exists a UGAS compact set $\mathcal{M}_c \subset \bar{v}\mathbb{B} \times \mathcal{F}_c$ for the dynamics (84). □

### Proof of Theorem 2

With Lemma 6 in hand we can now follow the exact same steps of the proof of Theorem 1 in order to analyze the closed-loop system, given by

$$\dot{\theta} = f(\theta, z) \tag{86a}$$

$$\dot{\hat{w}} = -\frac{\varepsilon_1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - h(\theta, z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - y_k\right) + \alpha_3\sum_{j\in\mathcal{N}_i} a_{ij}(\hat{w}_i - \hat{w}_j)\right) \tag{86b}$$

$$\dot{z} = \varepsilon_1 F_z\left(\nabla b(z)^\top \hat{w}, z\right). \tag{86c}$$

In particular, in the $\tau$-time scale system (86) is a singularly perturbed system with fast dynamics corresponding to the plant dynamics (31), which have a well-defined quasi-steady state manifold. The slow dynamics correspond to the system

$$\dot{\hat{w}} = -\frac{1}{\varepsilon_2}\left(\alpha_1\Psi(z)\left(\hat{\phi}(z) - h(\theta, z)\right) + \alpha_2\sum_{k=1}^{J}\Psi(z_k)\left(\hat{\phi}(z_k) - y_k\right) + \alpha_3\sum_{j\in\mathcal{N}_i} a_{ij}(\hat{w}_i - \hat{w}_j)\right) \tag{87a}$$

$$\dot{z} = F_z\left(\nabla b(z)^\top \hat{w}, z\right). \tag{87b}$$

which is also a singularly perturbed system, with fast dynamics corresponding to the dynamics (73), and slow dynamics corresponding to the optimizing dynamics (42). By Lemma 6 and Assumption 9, both dynamics have suitable stability properties. Therefore, as in the proof of Theorem 1, singular perturbation theory[75, Thm. 2] establishes the semi-global practical convergence result for the closed-loop system.

## 6 | NUMERICAL EXAMPLE: COOPERATIVE SOURCE SEEKING IN MULTI-VEHICLE SYSTEMS

In this section, we present a numerical example that illustrates the main properties of the algorithms considered in this paper. In particular, we consider a multi-vehicle localization problem characterized by four vehicles aiming to locate the source of a potential field by using only intensity measurements. We also illustrate the importance of the data-driven and the cooperative terms in the function (41).

### Model of the System

We consider a MAS with four vehicles, where each vehicle can only sense the intensity of the potential field with respect to its current position. The vehicles share information via an undirected connected graph with edges $(1,2), (2,3), (4,1)$. For simplicity, each vehicle is modeled as a simple linear system with quadratic output, of the form

$$\dot{\theta}_i = A_i\theta_i + B_i z_i, \tag{88}$$

$$y_i = \theta_i^\top Q_i \theta_i + c_i^\top \theta_i + d_i, \tag{89}$$

for all $i \in \{1,2,3,4\}$, where $\theta_i \in \mathbb{R}^2$ is the position on the plane of the $i^{th}$ vehicle, $z_i \in \mathbb{R}^n$ is the input, and $y_i \in \mathbb{R}$ is the output. We assume that the matrices $(A_i, B_i)$ have already been designed to guarantee steady-state regulation, i.e., $0 = A_i\theta_i^\star + B_i z_i \Rightarrow$

$\theta_i^\star = \ell_i(z_i) = z_i$. In particular, we consider the following matrices $(A_i, B_i)$ that satisfy this property:

$$A_1 = -10I_2,\ \ B_1 = 10I_2,\ \ A_2 = -20I_2,\ \ B_2 = 20I_2,\ \ A_3 = -30I_2,\ \ B_3 = 30I_2,\ \ A_4 = -10I_2,\ \ B_4 = 10I_2,$$

and $Q_i = -I_2$, $c_i = [4,8]^\top$, $d_i = -20$ for all $i \in \{1,2,3,4\}$. The response map of the agents is then given by

$$\phi(z_i) = -z_i^\top I_2 z_i + z_i^\top [4,8]^\top - 20, \quad \forall\ i \in \{1,2,3,4\}, \tag{90}$$

In order to locate the maximizer using individual real-time measurements $y_i$ and input-output data $\{z_{i,k}, y_{i,k}\}$, the vehicles implement the CODES dynamics (40). The input of each vehicle is restricted to the individual navigation set $\mathcal{F}_i = p_i + 2\mathbb{B}$, where $p_1 = [-2,8]^\top$, $p_2 = [6,8]^\top$, $p_3 = [2,0]^\top$, and $p_4 = [2,4]^\top$. The feasible set of the overall MAS is $\mathcal{F}_c = \mathcal{F}_1 \times \mathcal{F}_2 \times \ldots \times \mathcal{F}_N$. The individual optimizing algorithm used by each vehicle is given by equation (24), which exploits the fact that the Euclidean projection on a ball of radius $r > 0$ is given by $P_{r\mathbb{B}}(z) = \min\left\{\frac{r}{|z|}, 1\right\} z$. To approximate the response map (90), each agent implements polynomial basis functions that satisfy

$$b_i(z_i) = \left[z_{i,1}^2, z_{i,1}, z_{i,2}^2, z_{i,2}, z_{i1}z_{i2}, 1\right]^\top \in \mathbb{R}^6, \quad \nabla b_i(z_i) = \begin{bmatrix} 2z_{i,1} & 1 & 0 & 0 & z_{i,2} & 0 \\ 0 & 0 & 2z_{i,2} & 1 & z_{i,1} & 0 \end{bmatrix}^\top \in \mathbb{R}^{6\times 2}. \tag{91}$$

Based on the transient performance of the plant dynamics (88), the parameters of the CODES are selected as $\varepsilon_1 = 7$, $\varepsilon_2 = 0.58$.

### Generating the Individual Data

We assume that each vehicle has access to only 6 points of data $\{z_{i,k}, y_{i,k}\}$ generated as follows:

$$z_{1,k} = \begin{bmatrix} \sin(40(k-1)T_s) + 1 \\ 1 \end{bmatrix},\ z_{2,k} = \begin{bmatrix} 1 \\ \cos(40(k-1)T_s) \end{bmatrix},\ z_{3,k} = \begin{bmatrix} \cos(15(k-1)T_s) \\ 1 \end{bmatrix},\ z_{4,k} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \forall\ k \in \{1,2,\ldots,6\}, \tag{92}$$

where $T_s = 0.1$. The output data $y_{i,k}$ is obtained by sampling the output of each vehicle (88) after 5 seconds of applying the input $z_{i,k}$ with initial conditions given by $\theta_i(0) = p_i$ for all $i \in \{1,2,3,4\}$ and all $k \in \{1,2,\ldots,6\}$, which guarantees that condition (21) holds with $\tilde{\rho} = 0.01$. It can be verified that these data generate matrices $D_i$ with normalized entries (44) with the following ranks:

$$\mathrm{rank}(D_1) = 3,\ \ \mathrm{rank}(D_2) = 3,\ \ \mathrm{rank}(D_3) = 3,\ \ \mathrm{rank}(D_4) = 1. \tag{93}$$

Therefore, none of the vehicles has enough information to satisfy the $(\gamma, J)$-SR condition (16). However, it can also be verified that $\mathrm{rank}(D_1 + D_2 + D_3 + D_4) = 6$, which implies that their joint data satisfies the $(\gamma, J, N)$-CSR condition (38).

### Simulations

We simulate the CODES dynamics with $\alpha_1 = \alpha_2 = \alpha_3 = 1$. Vehicles $\{1,2,3\}$ are initialized at the center of their navigation set $\mathcal{F}_i$, and vehicle 4 is initialized at the point $\theta_4(0) = [1,3]^\top$. Figure 3 shows the resulting trajectories of the vehicles on the plane. As it can be observed, all the vehicles converge to the maximizer of the response map $\phi$ subject to the local navigation constraints. Figure 4 shows the evolution in time of the position $\theta_i$ and input $z_i$ of the vehicles, which converge to the optimal value $\theta^\star = z^\star$. As it can be observed, and in contrast to the traditional extremum seeking approaches, the trajectories of the vehicles do not exhibit oscillatory behaviors since no persistent exploratory signal is added to the input of the vehicles. Figure 5 shows the evolution in time of the parameter estimation errors $\tilde{w}_i$, which converge to zero as expected.

### No Cooperation between the Vehicles

In order to illustrate the importance of the cooperative term in equation (41), we now set $\alpha_3 = 0$, which implies that agents are not allowed to share their parameter estimation with their neighbors. By setting $\alpha_3 = 0$ the CODES dynamics reduce to the DES dynamics, which require that each agent satisfies the $(\gamma, J)$-SR condition to guarantee convergence. Figure 6 shows the trajectories of the vehicles on the plane. As it can be observed, in this case the vehicles do not converge to the optimal joint set $\mathcal{A}_c$, which is expected given that the individual matrices of data $D_i$ are not full column rank. Figures 7 and 8 show the evolution in time of the positions and inputs of the agents, as well as the parameter estimation error. As it can be observed, the estimation error does not converge to zero for all vectors $\tilde{w}_i$.

### No Recorded Data used by the Vehicles

We finish this section by considering the situation where agents are not allowed to use recorded data in the learning dynamics, i.e., $\alpha_2 = 0$ in (41). In this case, the trajectories of the vehicles are shown in Figure 9. Since the trajectories of the vehicles do not satisfy the PE condition, the vehicles are not able to achieve parameter estimation of the optimal weights $w^\star$, and they

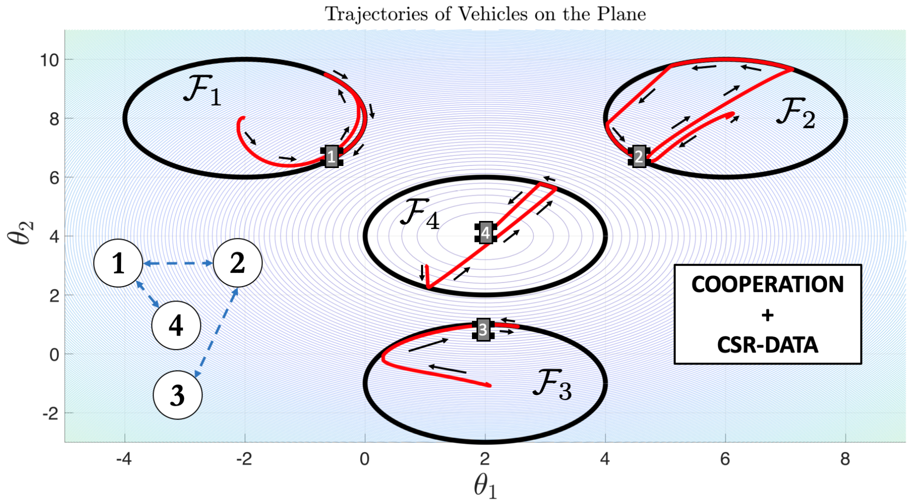


**FIGURE 3** Trajectories of a MAS comprised of four vehicles with internal dynamics (88), implementing the CODES dynamics to maximize their response map subject to local navigation constraints. The graph describes the communication link between the vehicles. The vehicle figure's indicate the final position of each agent.

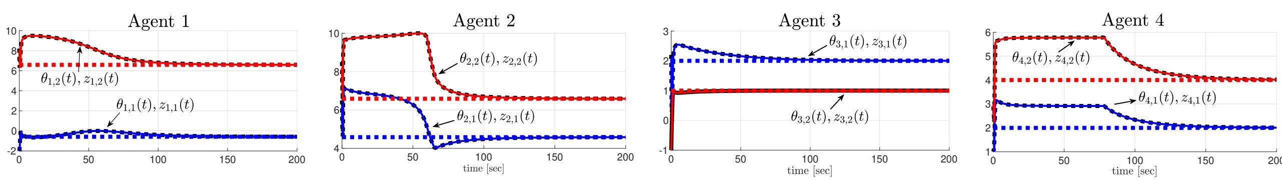


**FIGURE 4** Evolution in time of the states and inputs of the vehicles. The dotted colored lines indicate the optimal points. The black dotted line corresponds to the inputs of the agents, while the solid colored line describes their position.

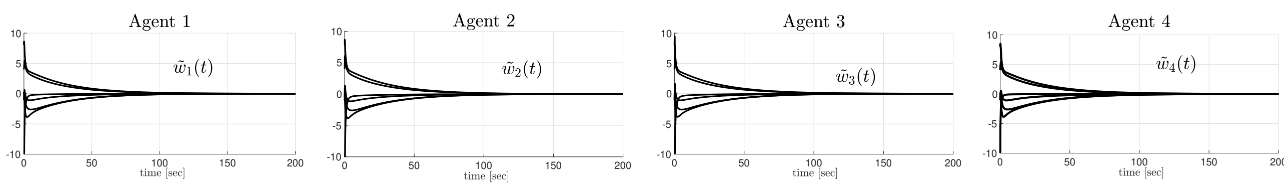


**FIGURE 5** Evolution in time of the vector of parameter estimation errors associated to each vehicle.

converge to a non-optimal location insider their navigation set. Figures 10 an 11 show the evolution in time of the states, inputs, and parameter estimation error. As expected, the parameter estimation error does not converge to zero.

## 7 | CONCLUSIONS AND OUTLOOK

In this paper we presented a new class of data-enabled extremum seeking dynamics that rely on information-rich data sets instead of external time-varying dither signals. The algorithms are suitable for single-agent and multi-agent optimization problems subject to constraints characterized by compact sets. Sufficient conditions on the optimization dynamics and the richness of the data were presented for both single-agent and multi-agent systems. In the latter case, it was shown that cooperation between agents can be harnessed to compensate for the absence of individual information-rich data sets. Different examples of suitable optimization dynamics were also presented, and connections and differences with respect to existing results in concurrent learning and cooperative adaptive control were also discussed.

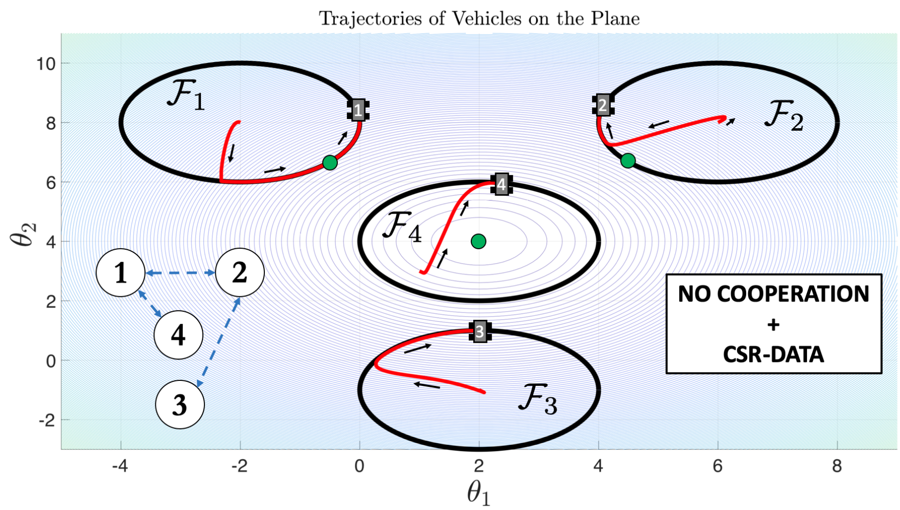


**FIGURE 6** Trajectories of a MAS comprised of four vehicles with internal dynamics (88), implementing the CODES dynamics with no cooperation, i.e., $\alpha_3 = 0$. The graph describes the communication link between the vehicles. The vehicle figure's indicate the final position of each agent. The green circle indicates the theoretical optimal location of each vehicle.

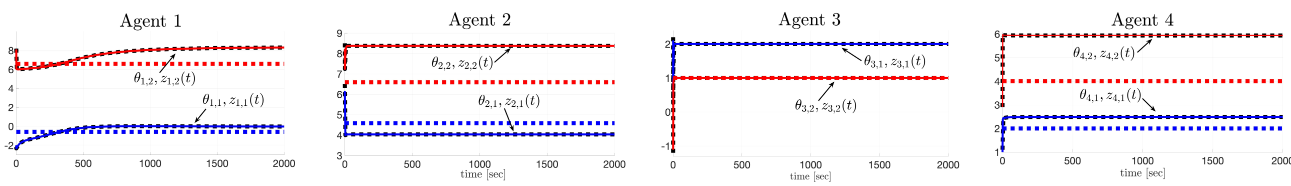


**FIGURE 7** Evolution in time of the states and inputs of the vehicles. The dotted colored lines indicate the optimal points. The black dotted line corresponds to the inputs of the agents, while the solid colored line describes their position. Agents do not converge to their optimal value due to lack of coordination and sufficiently rich data.

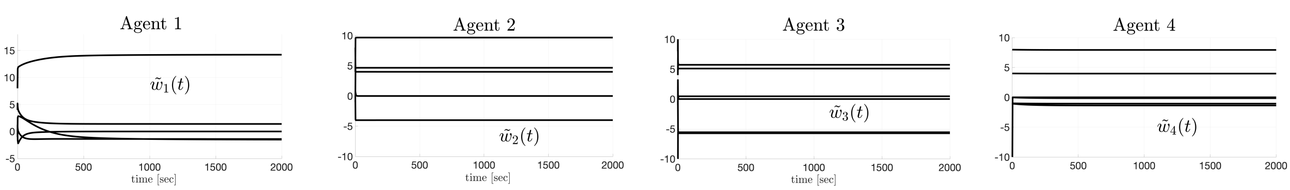


**FIGURE 8** Evolution in time of the vector of parameter estimation errors associated to each vehicle when no cooperation is used by the vehicles. As expected, all the errors do not converge to zero.

There exist several potential future extensions to the results presented in this paper. First, it is of interest to consider optimization dynamics that are characterized by hybrid dynamical systems rather than ODEs. This setting is relevant because hybrid optimization dynamics can be exploited to achieve global convergence results in some non-convex optimization problems, and can also be exploited to induce robust acceleration via resetting mechanisms[9]. Second, it is desirable to relax the homogeneity assumption on the cost function of the agents considered in Section 4. This would probably require either strong richness condition on the data of each agent, or stronger conditions on the basis that parameterize the functions. Third, while all the dynamics considered in this paper were based on one-layer neural network approximations, it is of interest to design multi-layer approximations in the spirit of deep learning. Since multi-layer approximations usually lead to quadratic estimation errors $e^2$ that are not convex with respect to the estimation error $\tilde{w}$, this may provide further motivations for the development of hybrid data-driven ES dynamics able to escape local minima.

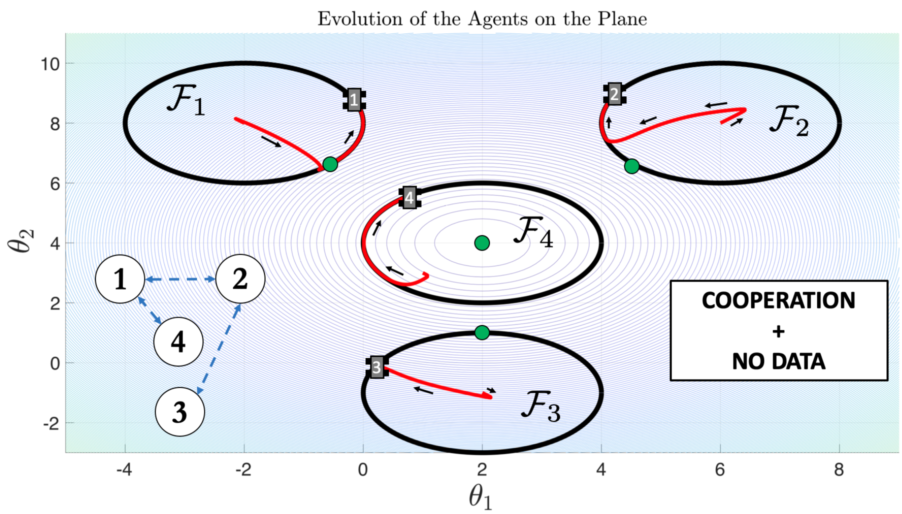


**FIGURE 9** Trajectories of a MAS comprised of four vehicles with internal dynamics (88), implementing the CODES dynamics with no data, i.e., $\alpha_2 = 0$. The graph describes the communication link between the vehicles. The vehicle figure's indicate the final position of each agent. The green circle indicates the theoretical optimal location of each vehicle.

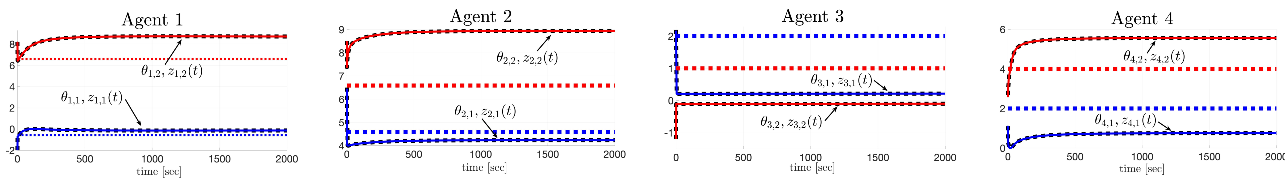


**FIGURE 10** Evolution in time of the states and inputs of the vehicles. The dotted colored lines indicate the optimal points. The black dotted line corresponds to the inputs of the agents, while the solid colored line describes their position. Agents do not converge to their optimal value due to lack of representative data in the learning dynamics.

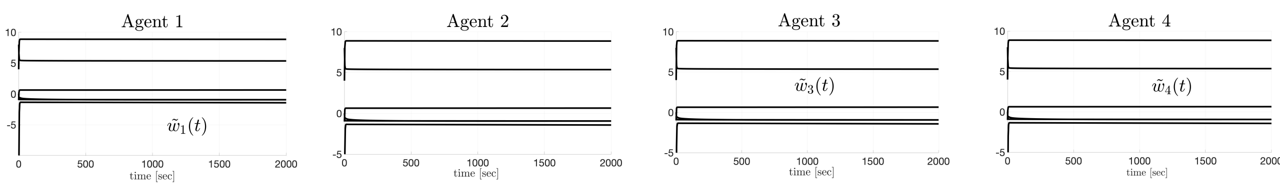


**FIGURE 11** Evolution in time of the vectors of parameter estimation errors associated to each vehicle when no data is used in the dynamics. As expected, the errors do not converge to zero.

## ACKNOWLEDGMENTS

The authors would like to acknowledge support in part from CU Boulder ASIRT Seed Grant No. 11005946, by ONR Minerva under grant No. N00014−18−1−2160, by NSF under grant Nos. CAREER CPS-1851588, S&AS 1849198 and SATC-1801611.

## Conflict of interest

The authors declare no potential conflict of interests.

## References


1. Tanaskovic M, Fagiano L, Novara C, Morari M. Data-driven control of nonlinear systems: An on-line direct approach. *Automatica* 2017; 75: 1-10.

2. Coulson J, Lygeros J, Dörfler F. Data-Enabled Predictive Control: In the Shallows of the DeePC. *arXiv:1811.05890*.

3. Vamvoudakis K, Antsaklis P, Dixon W, Hespanha J, Lewis F. Autonomy and machine intelligence in complex systems: A tutorial. *In Proc. of American Contr. Conf.* 2015: 5062-5079.

4. Poveda JI, Benosman M, Teel AR. Hybrid online learning control in networked multiagent systems: A Survey. *International Journal of Adaptive Control and Signal Processing* 2019; 33(2): 228-261.

5. Benosman M. Model-based vs data-driven adaptive control: An overview. *International Journal of Adaptive Control and Signal Processing* 2018: 753-776.

6. Colombino M, Dall'Anese E, Bernstein A. Online Optimization as a Feedback Controller: Stability and Tracking. *IEEE Trans. on Control of Network Systems, to appear.*

7. Ariyur KB, Krstić M. *Real-Time Optimization by Extremum-Seeking Control*. Wiley . 2003.

8. Sastry S, Bodson M. *Adaptive Control: Stability, Convergence, and Robustness*. Englewood Cliffs, NJ: Prentice-Hall . 1989.

9. Poveda JI, Li N. Inducing Uniform Asymptotic Stability in Non-Autonomous Accelerated Optimization Dynamics via Hybrid Regularization. *58th IEEE Conference on Decision and Control* 2019: 3000-3005.

10. Efimov D, Barabanov N, Ortega R. Robust Stability Under Relaxed Persistent Excitation Conditions. *IEEE Conference on Decision and Control* 2018: 7243-7248.

11. Barabanov N, Ortega R, Wang J. On Global Asymptotic Stability of $\dot{x} = -\phi(t)\phi(t)^\top x$ with $\phi(t)$ bounded and not persistently exciting. *Systems & Control Letters* 2017; 109: 24-27.

12. Wang J, Efimov D, Aranovskiy S, Bobtsov A. Fixed-time estimation of parameters for non-persistent excitation. *https://doi.org/10.1016/j.ejcon.2019.07.005* 2019.

13. Chowdhary G, Johnson E. Concurrent Learning for Convergence in Adaptive Control without Persistence of Excitation. *49th IEEE Conference on Decision and Control* 2010: 3674-3679.

14. Chowdhary G, Yucelen T, Muhlegg M, Johnson EN. Concurrent Learning Adaptive Control of Linear Systems with Exponentially Convergent Bounds. *International Journal of Adaptive Control and Signal Processing* 2012; 27(4): 280-301.

15. Chowdhary G, Wu T, Cutler M, Ure NK, How JP. Experimental Results of Concurrent Learning Adaptive Controllers. *AIAA Guidance, Navigation and Control Conference,* 2012: 1-14.

16. Modares R, Lewis F, Yucelen T, Chowdhary G. Adaptive optimal control of partially-unknown constrained-input systems using policy iteration with experience replay. *AIAA Guidance, Navigation and Control Conference, Boston, MA, doi:10.2514/6.2013-4519* 2013.

17. Owens DH. *Iterative Learning Control: An Optimization Paradigm*. London: Springer . 2015.

18. Kamalapurkar R, Klotz JR, Dixon WE. Concurrent Learning-based Approximate Feedback-Nash Equilibrium Solution of N-player Nonsero-sum Differential Games. *IEEE/CAA Journal of Automatica Sinica* 2014; 1: 239-247.

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

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

21. Nešić D, Tan Y, Moase W, Manzie C. A Unifying Approach to Extremum Seeking: Adaptive Schemes Based on Estimation of Derivatives. *Proc. of 49th IEEE Conference on Decision and Control* 2010.

22. Poveda JI, Teel AR. A framework for a class of hybrid extremum seeking controllers with dynamic inclusions. *Automatica* 2017; 76: 113-126.

23. Goebel R, Sanfelice RG, Teel AR. *Hybrid Dynamical Systems: Modeling, Stability, and Robustness*. Princeton University Press . 2012.

24. Dürr H, Stanković M, Ebenbauer C, Johansson KH. Lie bracket approximation of extremum seeking systems. *Automatica* 2013; 49: 1538-1552.

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

26. Feiling J, Koga S, Krstić M, Oliveira TR. Gradient extremum seeking for static maps with actuation dynamics governed by diffusion PDEs. *Automatica* 2018; 95: 197-206.

27. Scheinker A, Scheinker D. Bounded extremum seeking with discontinuous dithers. *Automatica* 2016; 69: 250-257.

28. Teel AR, Popović D. Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming. *Proc. of the American Control Conference* 2001: 2394–2399.

29. Popovic D, Jankovic M, Magner S, Teel A. Extremum Seeking Methods for Optimization of Variable Cam Timing Engine Operation. *IEEE Transactions on control systems technology* 2006; 14: 398–407.

30. Khong SZ, Nešić D, Tan Y, Manzie C. Unified framework for sampled-data extremum seeking control: Global optimisation and multi-unit systems.. *Automatica* 2013; 49: 2720-2733.

31. Poveda JI, Teel AR. A Robust Event-Triggered Approach for Fast Sampled-Data Extremization and Learning. *IEEE Transactions on Automatic Control* 2017; 62(10): 4949-4964.

32. Poveda JI, Benosman M, Sanfelice RG, Teel AR. A Hybrid Adaptive Feedback Law for Robust Obstacle Avoidance and Coordination in Multiple Vehicle Systems. *In proc. of American Control Conferece* 2018: 616-621.

33. Vandemeulen I, Guay M, McLellan PJ. Discrete-time distributed extremum seeking control over networks with unstable dynamics. *IEEE Transactions on Control of Network Systems* 2017.

34. Kvaternik K, Pavel L. An Analytic Framework for Decentralized Extremum Seeking Control. *In Proc. IEEE Ame. Con. Conf* 2012: 3371–3376.

35. Khong SZ, Tan Y, Manzie C, Nesic D. Multi-agent source seeking via discrete-time extremum seeking control. *Automatica* 2014; 50(9): 2312-2320.

36. Suttner R, Dashkovskiy S. Exponential Stability for Extremum Control Systems. *20th IFAC World Congress* 2017: 15464-15470.

37. Poveda JI, Benosman M, Teel AR. Distributed Extremum Seeking in Multi-Agent Systems with Arbitrary Switching Graphs. *IFAC World Congress* 2016; 50(1): 735-740.

38. Guay M, Zhang T. Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. *Automatica* 2003; 39: 1283–1293.

39. Guay M, Dochain D. A time-varying extremum-seeking control approach. *Automatica* 2015; 51: 356-363.

40. Atta KT, Johansson A, Guay M. On the Generalization and Stability Analysis of Pareto Seeking Control. *IEEE Control Systems Letters* 2018; 2(1): 145-150.

41. Atta KT, Guay M. Adaptive amplitude fast proportional integral phasor extremum seeking control for a class of nonlinear system. *Journal of Process Control* 2019; 83: 147-154.

42. Atta KT, Johansson A, Gustafsson T. Accuracy Improvement of Extremum Seeking Control. *IEEE Transactions on Automatic and Control* 2017; 62(4): 1952-1958.

43. Poveda JI, Vamvoudakis KG, Benosman M. A Neuro-Adaptive Architecture for Extremum Seeking Control Using Hybrid Learning Dynamics. *In proc. of American Control Conferece. Boston, MA*. 2017: 542-547.

44. Guay M, Vandemeulen I, Dougherty S, Mclellan PJ. Distributed extremum-seeking control over networks of dynamic agents. *In proc. of American Control Conferencel* 2015: 159-164.

45. Dougherty S, Guay M. An extremum-seeking controller for distributed optimization over sensor networks. *IEEE Transactions on Automatic Control* 2017; 62(2): 930-933.

46. Adetola V, Guay M. Parameter Convergence in adaptive extremum-seeking control. *Automatica* 2007; 43: 105-110.

47. Chen W, Hua S, Ge S. Consensus-based distributed cooperative learning control for a group of discrete-time nonlinear multi-agent systems using neural networks. *Automatica* 2014; 50: 2254-2268.

48. Poveda JI, Vamoudakis K, Benosman M. DEES: A Class of Data-Enabled Robust Feedback Algorithms for Real-Time Optimization. *IFAC Symposium on Nonlinear Control Systems* 2019: 670-675.

49. Poveda JI, Vamvoudakis K, Benosman M. CODES: Cooperative Data-Enabled Extremum Seeking for Multi-Agent Systems. *IEEE 58th Conference on Decision and Control* 2019: 2988-2993.

50. Poveda JI. Robust Hybrid Systems for Control, Learning, and Optimization in Networked Systems. *PhD Thesis, University of California, Santa Barbara* 2018.

51. Rockafellar RT, Wets RJ. *Variational Analysis*. Springer . 1998.

52. Facchinei F, Pang J. *Finite-Dimensional Variational Inequalities and Complementary Problems*. Springer . 2003.

53. Hornik K, Stinchcombe S, White H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks* 1990; 3: 551-560.

54. Hornik K. Multilayer feedforward networks are universal approximators. *Neural Networks* 1989; 2(5): 359-366.

55. Vamvoudakis KG, Lewis FL. Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* 2010; 46(5): 878-888.

56. Dudley RM. *Real Analysis and Probability*. Cambridge Uniersity Press . 2002.

57. Gao W, Jiang ZP. Adaptive Dynamic Programming and Adaptive Optimal Output Regulation of Linear Systems. *IEEE, Transactions on Automatic Control* 2016; 61(12): 4164-4169.

58. Lewis FL, Vrabie D, Vamvoudakis KG. Reinforcement learning and feedback control: using natural decision methods to design optimal adaptive controllers. *IEEE Control Systems Magazine* 2012; 32: 76-105.

59. Vrabie D, Vamvoudakis KG, Lewis FL. *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*. IET Press . 2012.

60. Johnson M, Kamalapurkar R, Bhasin S, Dixon W. Approximate N-Player Nonzero-Sum Game Solution for an Uncertain Continuous Nonlinear System. *IEEE Transactions on Neural Networks* 1645; 26(8): 1645-1658.

61. Ioannou PA, Sun J. *Robust Adaptive Control*. Mineola, NY.: Dover Publications Inc. . 2012.

62. Loria A, Panteley E, Popović D, Teel AR. A Nested Matrosov Theorem and Persistency of Excitation for Uniform Convergence in Stable Nonautonomous Systems. *IEEE Transactions on Automatica Control* 2005; 50(2): 183-198.

63. Bruggemann S, Possieri C, Poveda JI, Teel AR. Robust constrained model predictive control with persistent model adaptation. *IEEE 55th Conference on Decision and Control* 2016: 2364-2369.

64. Xia YS, Wang J. On the Stability of Globally Projected Dynamical Systems. *Journal of Optimization Theory and Applications* 2000; 106(1): 129-150.

65. Gao X. Exponential Stability of Globally Projected Dynamic Systems. *IEEE Transactions on Neural Networks* 2003; 14(2): 426-431.

66. Sandholm W. *Population Games and Evolutionary Dynamics*. The MIT Press . 2010.

67. Cochran J, Krstić M. Nonholonomic Source Seeking With Tuning of Angular Velocity. *IEEE Trans. on Autom. Control* 2009; 54: 717-731.

68. Acharya DS, Nath N. Applications of Multi Agent Systems in Control Engineering: A State of the Art Survey. *International Journal of Innovative Research in Advanced Engineering* 2015; 2(4): 113-128.

69. Poveda JI, Quijano N. Shahshahani gradient-like extremum seeking. *Automatica* 2015; 58: 51-59.

70. Brinon-Arranz L, Schenato L, Seuret A. Distributed Source Seeking Via a Circular Formation of Agents Under Communication Constraints. *IEEE Trans. on Control of Network Systems* 2016: 104-115.

71. Lewis FL, Liu D. *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. Hoboken, New Jersey: John Wiley/IEEE Press, Computational Intelligence Series . 2012.

72. Christiano P, Shah Z, Mordatch I, et al. Transfer from simulation to real world through learning deep inverse dynamics model. *arXiv preprint arXiv:1610.03518* 2016.

73. Barreiro-Gomez J, Obando G, Quijano N. Distributed Population Dynamics: Optimization and Control Applications. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 2017; 47(2): 304-314.

74. Chen W, Hua S, Zhang H. Consensus-Based Distributed Cooperative Learning From Closed-Loop Neural Control Systems. *IEEE Trans. on Neural Net. and Learning Syst.* 2015; 26(2): 331-345.

75. Wang W, Teel AR, Nešić D. Analysis for a class of singularly perturbed hybrid system via averaging. *Automatica* 2012; 48.

## AUTHOR BIOGRAPHY

**Jorge I. Poveda** received the M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the University of California at Santa Barbara in 2016 and 2018, respectively. He was a Postdoctoral Fellow at Harvard University in 2018 and a research intern at the Mitsubishi Electric Research Laboratories (MERL) during the summers of 2016 and 2017. He was a Best Student Paper Award Finalist in the 2017 IEEE Conference on Decision and Control, and he received the CCDC Outstanding Scholar Fellowship at UCSB. His research interests lie in the design and analysis of high-performance optimization, learning, and control algorithms for cyber-physical systems using tools from nonlinear control theory, hybrid dynamical systems, and game theory. Since January 2019 he is affiliated with the Department of Electrical, Computer, and Energy Engineering at the University of Colorado, Boulder, where he is currently an Assistant Professor.

**Mouhacine Benosman.** Mouhacine Benosman is a Senior Research Scientist at Mitsubishi Electric Research Labs (MERL) in Cambdrige, USA. Before joining MERL he worked at Reims University, France, Strathclyde University, Scotland, and National University of Singapore. His research interests include nonlinear control of flexible systems, multi-agent distributed control with applications to robotics and smart-grid systems, and learning and adaptive control. He has published more than 70 peer-reviewed journal articles and conference papers, and more than 20 patents in the field of mechatronics systems control. He is a senior member of the IEEE, Associate Editor of the Control System Society Conference Editorial Board, Associate Editor of the Journal of Optimization Theory and Applications, and Senior Editor of the International Journal of Adaptive Control and Signal Processing.

**Kyriakos G. Vamvoudakis** was born in Athens, Greece. He received the Diploma (a 5 year degree, equivalent to a Master of Science) in Electronic and Computer Engineering from Technical University of Crete, Greece in 2006 with highest honors. After moving to the United States of America, he studied at The University of Texas and he received his M.S. and Ph.D. in Electrical Engineering in 2008 and 2011 respectively. During the period from 2012 to 2016 he was a project research scientist at the Center for Control, Dynamical Systems and Computation at the University of California, Santa Barbara. He was an assistant professor at the Kevin T. Crofton Department of Aerospace and Ocean Engineering at Virginia Tech until 2018. He is now an assistant professor at The Daniel Guggenheim School of Aerospace Engineering at Georgia Institute of Technology. His research interests include optimal control, reinforcement learning, and game theory. Recently, his research has focused on cyber-physical security, and safe autonomy. Prof. Vamvoudakis is the recipient of a 2018 National Science Foundation CAREER Award, the 2016 International Neural Network Society Young Investigator (INNS) Award, the Best Paper Award for Autonomous/Unmanned Vehicles at the 27th Army Science Conference in 2010, and the Best Researcher Award from the Automation and Robotics Research Institute in 2011. He currently is an associate editor of Automatica, an associate editor of the IEEE Computational Intelligent Magazine, an associate editor of IEEE Control Systems Letters, an associate editor of Journal of Optimization Theory and Applications, an associate editor of Control Theory and Technology, a registered electrical/computer engineer (PE), and a member of the IEEE Control Systems Society Conference Editorial Board.

**How to cite this article:** J. I. Poveda, M. Benosman, K. Vamvoudakis (2020), Data-Enabled Extremum Seeking: A Cooperative Concurrent Learning-Based Approach, *under review*.