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

CODES: Cooperative Data-Enabled Extremum Seeking for Multi-Agent Systems

Poveda, J.; Vamvoudakis, K.; Benosman, M.

TR2019-148 December 11, 2019

Abstract

In this paper, we study the problem of modelfree cooperative real-time optimization in multi-agent network systems (MAS). Unlike existing adaptive extremum seeking approaches that presume the satisfaction of a persistence of excitation condition on the agents of the network, we propose a novel approach that leverages the presence of cooperation and information-rich data sets in the system. This approach is based on the idea that in MAS with sufficient communication and information resources, agents can efficiently learn a common cost function under mild individual excitation requirements by leveraging cooperation. Therefore, our main result can be seen as a spatiotemporal condition that guarantees model-free optimization in MAS with agents having homogeneous but unknown cost functions. To solve this model-free optimization problem, we characterize a class of robust dynamics that can be safely interconnected with the data-enabled learning mechanism in order to achieve a stable closed-loop system. A numerical result is presented to illustrate the approach.

IEEE Conference on Decision and Control (CDC)

CODES: Cooperative Data-Enabled Extremum Seeking for Multi-Agent Systems

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

Abstract— **In this paper, we study the problem of model-free cooperative real-time optimization in multi-agent network systems (MAS). Unlike existing adaptive extremum seeking approaches that presume the satisfaction of a persistence of excitation condition on the agents of the network, we propose a novel approach that leverages the presence of cooperation and information-rich data sets in the system. This approach is based on the idea that in MAS with sufficient communication and information resources, agents can efficiently learn a common cost function under mild individual excitation requirements by leveraging cooperation. Therefore, our main result can be seen as a spatiotemporal condition that guarantees model-free optimization in MAS with agents having homogeneous but unknown cost functions. To solve this model-free optimization problem, we characterize a class of robust dynamics that can be safely interconnected with the data-enabled learning mechanism in order to achieve a stable closed-loop system. A numerical result is presented to illustrate the approach.**

I. INTRODUCTION

Due to the inherent model uncertainties that emerge in practical applications, as well as the growing amount of available data sets in engineering systems, there is an urgent need to design robust and data-enabled algorithms with suitable adaptive and learning properties, see for instance [26], [27], [22], [2], and references therein. In this paper, we study a specific learning problem in multi-agent systems (MAS), namely, optimizing a global reward function using a cooperative data-enabled approach in the context of zero-order optimization and extremum seeking control (ES). ES control has emerged as a promising model-free feedback-based optimization strategy that provides certifiable stability and robustness properties in applications where the mathematical form of the cost function is unknown [18], [20], [10], [23], [8]. For multi-agent systems, different types of ES dynamics have been presented in [21], [28], and [19], to just name a few. ES dynamics based on persistence of excitation (PE) conditions have been studied in [12], [11], [7], [25]. A set point-based relaxed PE condition for ES was presented in [1], and a relaxed memory-based PE condition for ES in single-agent systems has been recently developed in [24]. However, while these approaches have provided valuable insight, guaranteeing *a priori* the satisfaction of a PE condition along the trajectories of a MAS remains a challenging problem in many practical settings, e.g., network and wireless control, control of autonomous robots, optimization of wind farms, etc. Since existing distributed ES approaches rely entirely on current (online) measurements of the cost function, available information-rich data sets that could be used to either improve the performance of the controller, or to relax the excitation conditions, are usually ignored. However, as shown in [5] and [6], in certain applications it is indeed possible to design learning dynamics that dispense with the classic PE assumption by leveraging the existence of information-rich data sets. Moreover, as shown in [3], when multiple agents with local learning dynamics cooperatively interact in a networked system, individual PE conditions imposed to all the agents of system may not be necessary to guarantee that the overall MAS achieves successful learning.

Motivated by this background, in this paper we extend our previous results of [24] by studying a novel class of cooperative data-enabled extremum seeking (CODES) algorithms that combine ideas from cooperative adaptive estimation and concurrent learning in order to achieve real-time optimization in multi-agent systems (MAS). In particular, we present a spatiotemporal condition that guarantees the solution of a common optimization problem in MAS with unknown mathematical forms of the cost, where agents are not required anymore to individually satisfy a classic persistence of excitation condition, but only to cooperatively have "sufficiently rich" stored data. This spatiotemporal condition merges together past data (temporal) information and share data (spatial) information from other agents. We characterize a class of robust optimization dynamics that can be safely interconnected with the learning mechanism, and we show that the solutions of the closed-loop system converge to a neighborhood of the minimizers of the global cost function. Connections with the classic settings of concurrent learning and cooperative adaptive control are also discussed. To the knowledge of the authors the results of this paper correspond to the first family of ES controllers for MAS that exploit cooperation and memory (data) simultaneously.

The rest of this paper is organized as follows. In Section II, we present the model-free optimization problem and we characterize the CODES dynamics as well as their convergence properties. Sections III and IV illustrate the similarities and differences with respect to other existing approaches that relaxed the excitation conditions in the temporal domain and in the spatial domain. An example is presented in Section

J. I. Poveda is with the Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO 80309 USA email: jorge.poveda@colorado.edu

K. G. Vamvoudakis is with the Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0150, USA email: kyriakos@gatech.edu

M. Benosman is with the Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, 02139 email: benosman@ieee.org

This work was supported in part by NATO under grant No. SPS G5176, in part by ONR Minerva under grant No. N00014-18-1-2160, by an NSF CAREER CPS-1851588. The work of M. Benosman was fully supported by MERL.

V, and Section VI ends with conclusions.

II. PROBLEM FORMULATION AND MAIN RESULTS

In this paper, we consider networked systems characterized by a time-invariant undirected connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \ldots, N\}$ is the set of vertices or nodes, and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges between the nodes, which can be seen as communication links. Each node $i \in \mathcal{V}$ of the network can be seen as an agent aiming to solve the following individual optimization problem

$$\min \quad f(x) \quad \text{subject to} \quad x \in K_i, \tag{1}$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ is a $\mathcal{C}^2$ and convex cost function, and $K_i \subset \mathbb{R}^n$ is an individual compact set that incorporates constraints into the optimization problem. Our main standing assumption is that the mathematical form of f is unknown to all agents. Indeed, we assume that agents have access only to current individual evaluations of the global function $f(x_i)$, which are used to update their individual state $x_i \in \mathbb{R}^n$. Additionally, in order to solve problem (1), agents are allowed to share information with their neighbors. Examples where this type of model-free optimization problem emerges include cooperative source seeking with local constraints, cooperative surveillance with individual bounded action spaces, and cooperative formation control under unknown potential fields, to just name a few. These problems can be solved by leveraging the well-known paradigm of *exploration* vs *exploitation*. However, our goal in this paper is to use minimal online *exploration* (i.e., excitation) in the network by exploiting the existence of recorded data that can be accessed by the control system.

A. Feedback Structure and Uniform Function Approximation

In order to solve problem (1), and similar to [24], we propose a class of algorithm where each agent $i \in \mathcal{V}$ implements the following dynamics

$$\dot{\hat{w}}_i = \frac{1}{\varepsilon} F_{w,i}(\hat{w}, \phi(x), e_i), \tag{2a}$$

$$\dot{z}_i = F_{z,i}(\hat{w}_i^\top \nabla \phi_i(x_i), z), \quad z_i \in C_{z,i} \tag{2b}$$

where $z_i := [x_i^\top, s_i^\top]^\top \in \mathbb{R}^{n+r}$, $s_i \in \mathbb{R}^r$ is an auxiliary state of dimension $r \in \mathbb{Z}_{\geq 0}$, and $\varepsilon > 0$. The mapping $F_{z,i}$ and the set $C_{z,i} \subset \mathbb{R}^{n+r}$ will be designed based on the structure of the optimization problem (1), and they will be characterized in Section II-C. The mapping $F_{w,i} : \mathbb{R}^{np} \times \mathbb{R}^{np} \times \mathbb{R}^p \to \mathbb{R}^p$ is assumed to be a continuous function that only depends on the information of agent i and its neighboring agents $j \in \mathcal{N}_i := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$. This function, which describes the "learning dynamics", will be designed based on the available temporal and spatial information to each agent of the network.

The individual error signal e_i in (2a) is given by

$$e_i := \hat{f}_i(x_i) - f(x_i), \tag{3}$$

where $\hat{f}_i(x_i)$ is agent's i approximation of the global common cost function $f(x)$. In this paper, we consider approximations of the form

$$\hat{f}_i(x_i) = \hat{w}_i^\top \phi_i(x_i), \tag{4}$$

where the auxiliary state $\hat{w}_i \in \mathbb{R}^p$ is an estimation of an ideal weight $w_i^\star \in \mathbb{R}^p$ that satisfies

$$f(x) = w_i^{\star\top} \phi_i(x_i) + \epsilon_i(x_i) \tag{5}$$

for some approximation error $\epsilon_i(x_i)$ that satisfies $\max\{|\epsilon_i(x_i)|, |\nabla \epsilon_i(x_i)|\} \leq \delta$ for all $x_i \in K_i$, with $\delta > 0$. Such ideal weight always exists due to the Stone-Weirstrass high-order approximation Theorem, see [14], [13]. The set of all such optimal weights $w_i^\star$ is defined as

$$\mathbb{W}_i^\delta := \left\{ w_i^\star \in \mathbb{R}^p : |f(x_i) - w_i^{\star\top} \phi_i(x_i)| \leq \delta, \forall\, x_i \in K_i \right\},$$

which we assume to be compact, i.e., $\|w_i^\star\| \leq \overline{w}$ for all $w^\star \in \mathbb{W}_i^\delta$, and all $i \in \mathcal{V}$, for some $\overline{w} \in \mathbb{R}_{>0}$. The individual vector valued regressor functions $\phi_i : \mathbb{R}^n \to \mathbb{R}^p$ are assumed to be known and $\mathcal{C}^2$. For each agent $i \in \mathcal{V}$, the regressor function $\phi_i := [\phi_{i,1}, \phi_{i,2}, \ldots, \phi_{i,p}]^\top$ should be selected such that the functions $\phi_{i,j}$, $j \in \{1, \ldots, p\}$, define a complete independent basis set for $f(x_i)$. Typical choices of basis functions include quadratic functions, radial basis functions, or sigmoid functions [27]. We will impose the following technical assumption on the approximation (5).

Assumption 1: For all $i \in \mathcal{V}$, the approximation error function $\epsilon_i(\cdot)$ in (5) is continuously differentiable. □

Since K_i is compact, by the Stone-Weierstrass high-order approximation Theorem, as $p \to \infty$ we have that $\epsilon(x_i) \to 0$ and $\nabla \epsilon_i(x_i) \to 0$ uniformly in K_i. Moreover, due to Assumption 1, the compactness of K_i and the fact that $\phi_i(x_i)$ is $\mathcal{C}^2$, the mappings $\phi_i(x_i)$, $\nabla \phi_i(x_i)$, $\epsilon_i(x_i)$, and $\nabla \epsilon_i(x_i)$ are all uniformly bounded in K_i. However, since the set $\mathbb{W}_i^\delta$ depends on the individual functions ϕ_i and the individual compact sets K_i, we need the following assumption regarding the existence of a common optimal weight $w^\star \in \mathbb{R}^p$ for the MAS.

Assumption 2: There exists a $\delta^* > 0$ such that for each $\delta \in (0, \delta^*)$ there exists a $p^* \in \mathbb{Z}_{>0}$ such that for each $p > p^*$ the set $\mathbb{W}^\delta := \bigcap_{i \in \mathcal{V}} \mathbb{W}_i^\delta$ is not empty. □

Let $w^\star \in \mathbb{W}^\delta$ and define $\tilde{w}_i := \hat{w}_i - w^\star$ as the estimation error of agent i. Using (3), (4), and (5), the approximation error of each agent satisfies

$$e_i = \hat{f}_i(x_i) - f(x_i) = \tilde{w}_i^\top \phi_i(x_i) - \epsilon_i(x_i). \tag{6}$$

It is well-known, e.g. [15, Ch. 4] that under several learning dynamics $F_{w,i}$, exponential convergence of $e_i(t)$ to zero (or to an ε-neighborhood of zero) is achieved if and only if the regressor vector ϕ_i is "sufficiently rich", a requirement that is usually expressed in terms of a PE condition of the form

$$\int_t^{t+T_i} \phi_i(\tau)\phi_i^\top(\tau) \mathrm{d}\tau \geq \gamma_i I, \tag{7}$$

which must hold for all $t \geq 0$ and for some constants $T_i, \gamma_i > 0$. However, in general, for many practical applications it may be difficult to verify a priori this excitation condition [5]. Moreover, as discussed in [3], even if some agents of the MAS individually satisfy the PE condition, it is unrealistic to assume that *all* agents of large-scale network will satisfy the PE condition. Motivated by this limitation we propose to synergistically combine the availability of communication networks in the MAS, and the existence of information-rich data sets, to design cooperative learning dynamics that guarantee a correct estimation of $w^\star$ without assuming a priori that the trajectories of all the regressors satisfy a PE condition.

B. Learning in Networks using Memory and Cooperation

In order to establish the main ideas behind the cooperative learning dynamics, let $k \in \{1, 2, \ldots, \mathbf{k}\}$ denote the index of a stored data point $x_{i,k}$ collected at some time t_k by agent $i \in \mathcal{V}$, i.e., $x_{i,k} = x_i(t_k)$, and let $\phi_i(x_{i,k})$ be the regressor vector of agent i evaluated at that point. For each agent i and each $k \in \{1, 2, \ldots, \mathbf{k}\}$, we also introduce an estimation error associated to the data previously collected at time t_k, given by

$$\begin{aligned} e_i(t_k, t) &= \hat{f}_i(x_i(t_k)) - f(x_i(t_k)), \\ &= \tilde{w}_i(t)^\top \phi_i(x_{i,k}) - \epsilon(x_{i,k}). \end{aligned} \tag{8}$$

Note that the estimation error $\tilde{w}_i$ of each agent i depends on the current time t. To streamline the presentation of the learning dynamics, and to be consistent with (8), we will use t_0 to represent the current time t, i.e.,

$$e_i(t_0, t) := e_i(t), \ t \geq 0. \tag{9}$$

Using this notation, we consider the following adaptive dynamics that incorporate individual memory and cooperation between the agents:

$$\dot{\hat{w}}_i = -\alpha \sum_{k=0}^{\mathbf{k}} \Psi_i(x_i(t_k)) e_i(t_k, t) - \gamma \sum_{j \in \mathcal{N}_i} a_{ij}(\hat{w}_i - \hat{w}_j), \tag{10}$$

where $\alpha, \gamma > 0$ are tunable parameters, $a_{i,j} \in \{0, 1\}$ corresponds to the entry (i, j) of the adjacency matrix of the graph $\mathcal{G}$, and where the term $\Psi_i(x_i(t_k))$ is defined as

$$\Psi_i(x_i(t_k)) := \frac{\phi_i(x_i(t_k))}{(\phi_i(x_i(t_k))^\top \phi_i(x_i(t_k)) + 1)^2}. \tag{11}$$

The dynamic mechanism (10) is comprised of two main terms. The first term exploits the available memory in the system by considering the current approximation error, as well as the approximation error associated to a sequence of $\mathbf{k}$ measurements $\{x_{i,k}\}_{k=1}^{\mathbf{k}}$. On the other hand, the second term exploits the flow of information in the system between neighboring agents.

In order to study the convergence properties of (10) we introduce the following definition.

Definition 1: A collection of N sequences of stored data points, denoted by $\Phi := \{\{y_{i,k}\}_{k=1}^{\mathbf{k}} : y_{i,k} \in \mathbb{R}^n, i \in \mathcal{V}\}$, satisfying inequality

$$\sum_{k=1}^{\mathbf{k}} \sum_{i=1}^{N} y_{i,k} y_{i,k}^\top \succ 0, \tag{12}$$

is said to be $\mathbf{k}$-*cooperatively sufficiently rich* ($\mathbf{k}$-CSR). □

Condition (12) in Definition 1 ensures that the stored data in the *entire* network contains "sufficiently" rich information. Since the summation is taken over all agents and over a finite number of times, it relaxes the cooperative PE condition and the classic memory condition used in the context of concurrent learning [5], [24]. Indeed, unlike the PE condition (7), which applies to the past and future behavior of $\phi_i(t)$, the condition given by (12) relies only on *past* recorded data associated to the nodes of the network. Therefore, it can be verified a priori. The data can be recorded during a finite amount of time where the overall networked system is sufficiently excited, e.g., during a training phase. Note, however, that memory and bandwidth limitations will determine the number of points $\mathbf{k}$ that can be stored in the system.

Remark 1: If the number of data points collected by each agent differs, i.e., $\mathbf{k}_i \neq \mathbf{k}_j$, condition (12) can be tested with $\mathbf{k} := \min\{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\}$. □

The following proposition, corresponding to the first result of this paper, characterizes the convergence properties of the dynamics (10) under a $\mathbf{k}$-CSR condition on the normalized regressor vector

$$\tilde{\phi}_{i,k} := \frac{\phi_i(x_i(t_k))}{\phi_i(x_i(t_k))^\top \phi_i(x_i(t_k)) + 1}. \tag{13}$$

The proof is omitted due to lack of space.

Proposition 1: Suppose that Assumptions 1 and 2 hold, and that the functions $e_i(\cdot)$ and $\Psi_i(x_i(\cdot))$ are measurable. Suppose also that the collection of data points

$$\Phi := \{\{\tilde{\phi}_{i,k}\}_{k=1}^{\mathbf{k}} : \tilde{\phi}_{i,k} \in \mathbb{R}^n, i \in \mathcal{V}\} \tag{14}$$

is $\mathbf{k}$-CSR, and let $\alpha, \gamma > 0$. Then, for each $\bar{\nu} > 0$ there exists a sufficiently small $\delta > 0$ and a sufficiently large $p^* \in \mathbb{Z}_{>1}$ such that for each $p \in \mathbb{Z}_{\geq p^*}$ and each compact set $K_w \subset \mathbb{R}^p$ every solution of (10) with $\hat{w}_i(0) \in K_w$ for all $i \in \mathcal{V}$, converges in finite time to $\mathbb{W}^\delta + \bar{\nu}\mathbb{B}$. Moreover, the rate of convergence is exponential outside the set $\mathbb{W}^\delta + \bar{\nu}\mathbb{B}$. □

The convergence result of Proposition 1 is established by assuming the richness condition (12) on the data $\tilde{\phi}_{i,k}$, which can be verified a priori for the MAS. Further connections and differences with respect to existing cooperative and data-based PE conditions are discussed in Sections III and IV.

C. Robust Optimization Dynamics

Once a data-enabled learning mechanism for the estimation of $w^\star$ has been designed, we proceed to characterize the optimization dynamics for the solution of problem (1). Since the result of Proposition 1 implies a residual estimation error, in order to obtain a stable interconnection in the closed-loop system, the optimization dynamics should be robust to small but persistent additive disturbances on the gradient. To characterize these optimization dynamics, we assume in (2b) that $\nabla f_i = \hat{w}_i^\top \nabla\phi_i(x_i)$, i.e., we consider the ideal gradient-based optimization dynamics

$$\dot{z}_i = F_{z,i}(\nabla f_i(x_i), z_i), \quad z_i \in C_{z,i}. \tag{15}$$

For this system we impose the following Assumption on the mappings $F_{z,i}$ and the sets $C_{z,i}$.

Assumption 3: For each $i \in \mathcal{V}$ the dynamics (15) satisfy the following conditions:

(a) The mapping $F_{z,i}$ is continuous with respect to both arguments.
(b) The set $C_{z,i}$ satisfies $C_{z,i} = K_i \times S_i$, where $S_i \subset \mathbb{R}^r$ is a compact set.
(c) There exists a nonempty compact set $\mathcal{S}_i \subset S_i$ such that the set $\mathcal{A}_i \times \mathcal{S}_i$ is globally asymptotically stable, where $\mathcal{A}_i$ is the optimal solution of (1) for agent i.
(d) There exists an $\delta_i > 0$ such that for each measurable function $e : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ satisfying $\sup_{t\geq 0} |e(t)| \leq \delta_i$, the perturbed system $\dot{z}_i = F_{z,i}(\nabla f_i(x_i) + e, z_i)$, $z_i \in C_{z,i}$, generates complete solutions from each $z_i(0) \in C_{z,i}$. □

Generally, in order to construct optimization dynamics that satisfy Assumption 3, it suffices to known the structure of K_i and the convexity properties of f. Moreover, in some cases, item (d) can be relaxed, and complete solutions are only required from compact subsets of $C_{z,i}$. A particular example will be presented in Section V.

D. Main Result

We now present the main result of this paper, which combines the cooperative data-enabled learning mechanism (10) and the optimization dynamics characterized in Assumption 3, which generate a class of cooperative data-enabled extremum seeking (CODES) dynamics. The proof is omitted due to lack of space.

Theorem 1: Suppose that Assumptions 1, 2 and 3 hold. For each agent $i \in \mathcal{V}$ consider the dynamics (2) with $F_{w,i}$ given by (10). Suppose that the collection of sequences (14) is $\mathbf{k}$-CSR. Then, for each $\nu > 0$ there exists a sufficiently large $p^* \in \mathbb{Z}_{>1}$ such that for each $p \in \mathbb{Z}_{\geq p^*}$ and each compact set $K \subset \mathbb{R}^{Np}$ there exists $\varepsilon^* \in \mathbb{R}_{>0}$ and $T \in \mathbb{R}_{>0}$ such that for each $\varepsilon \in (0, \varepsilon^*)$ every solution of the closed-loop system with $\hat{w}(0) \in K$ generates trajectories x_i satisfying

$$x_i(t) \in \mathcal{A}_i + \nu\mathbb{B},$$

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

In words, Theorem 1 establishes that by selecting a sufficiently large vector of basis functions, by inducing enough time scale separation in the closed-loop, and by using data satisfying (12), the CODES algorithm guarantees convergence in finite time to an arbitrarily small ν-neighborhood of the optimal solution $\mathcal{A}_i$ of each agent.

Remark 2: Even though so far we have assumed that the optimization problems are uncoupled and that the interaction between the agents occurs via the cooperative estimation dynamics (10), there is no loss of generality by considering coupled optimization problems with global constraints. In that case, the optimization dynamics (15) must be designed to stabilize the global optimal set assuming that the gradient is known, and using the information available by the communication graph $\mathcal{G}$.

III. SPATIAL RELAXATION: COOPERATIVE LEARNING WITH NO MEMORY

The result of the previous section provides a framework for the design of CODES for MAS with memory and communication networks. On the other hand, when there is no memory in the nodes, i.e., $\mathbf{k} = 0$, and condition (12) applied to the normalized regressor vectors (13) must hold uniformly on every compact time domain $[t, t+T]$ for all $t \geq 0$ and some $T > 0$, condition (12) becomes

$$\int_t^{t+T} \sum_{i=1}^N \tilde{\phi}_i(\tau)\tilde{\phi}_i^\top(\tau)\mathrm{d}\tau \geq \gamma I, \quad \gamma > 0, \tag{16}$$

which is the cooperative PE condition studied in [3] in the context of classic adaptive parameter estimation and stabilization, as well as in [4] in the context of neuro-adaptive learning. Instead of asking that every agent satisfies condition (7), the cooperative condition (16) must hold for the overall MAS. The following example, corresponding to [3, Remark 4] illustrates this idea.

Example 1: [3, Remark 4] For a 2-agent system consider the individual normalized regressor signals $\tilde{\phi}_1(t) = [\sin(t), 0]$ and $\tilde{\phi}_2(t) = [0, \cos(t)]$. Note that none of these signals individually satisfy the classic PE condition (7). However the cooperative signal $\sum_{i=1}^2 \tilde{\phi}_i(t)\tilde{\phi}_i^\top(t) = \begin{bmatrix} \sin(t)^2 & 0 \\ 0 & \cos(t)^2 \end{bmatrix} \succ 0$, satisfies (16). □

On the other hand, while the cooperative PE condition relaxes the individual PE assumption in MAS, the $\mathbf{k}$-CSR condition (12) can additionally be verified a priori when sufficiently rich data exists.

Example 2: For the same 2-agent system of Example 1, consider the $\mathbf{k}$-sequence of *recorded* data obtained from the individual regressor signals $\{\tilde{\phi}_i(t_k)\}_{k=1}^{\mathbf{k}}$. In this case, condition (12) gives

$$\sum_{k=1}^{\mathbf{k}}\sum_{i=1}^{2} \tilde{\phi}_i(t_k)\tilde{\phi}_i^\top(t_k) = \begin{bmatrix} \sum_{k=1}^{\mathbf{k}} \sin(t_k)^2 & 0 \\ 0 & \sum_{k=1}^{\mathbf{k}} \cos(t_k)^2 \end{bmatrix},$$

thus for any sequence of points $\{t_k\}_{k=1}^{\mathbf{k}}$ such that $t_k \neq (k-1)\pi$ for all k, or $t_k \neq \frac{(k+1)\pi}{2}$ for all k, condition (12) is satisfied. On the other hand, if either $t_k = (k-1)\pi$ or $t_k = \frac{\pi(k+1)}{2}$ for all k, there is no $\mathbf{k} \in \mathbb{Z}_{>0}$ such that (12) holds. Thus, condition (12) may not be trivially satisfied even if the amount of memory is unbounded in the system. This is also true for condition (16) if, for instance, one selects $\phi_i := [1, 0]$ for all agents $i \in \mathcal{V}$. □

When there is no memory in the nodes, i.e., $\mathbf{k} = 0$, we recover the estimation dynamics of [3], and the learning dynamics (10) reduce to

$$\dot{\hat{w}}_i = -\alpha\Psi_i(x_i(t))e_i(t) - \gamma \sum_{j\in\mathcal{N}_i} a_{ij}(\hat{w}_i - \hat{w}_j), \tag{17}$$

where $e_i(t)$ corresponds to the standard error (6), and $\Psi_i(x_i(t))$ corresponds to the term (11) evaluated at $x_i(t)$. Under the cooperative PE condition (16) on the normalized regressors

$$\tilde{\phi}_i(t) := \frac{\phi_i(x_i(t))}{\phi_i(x_i(t))^\top \phi_i(x_i(t)) + 1}, \tag{18}$$

the following Corollary can be obtained as in [3]. The proof is omitted due to lack of space. While related stability results have been obtained for neuro-adaptive controllers [4], to our knowledge the following result is also novel in the context of extremum seeking control.

Corollary 1: Suppose that Assumptions 1, 2 and 3 hold. Consider the dynamics (2) with F_{w_i} given by (17). Suppose that the normalized regressors (18) satisfy the cooperative PE condition (16) with (γ, T) independent of $x_i(0)$ and t_0. Then, for each $\nu > 0$ there exists a $p^* \in \mathbb{Z}_{>1}$ such that for each $p \in \mathbb{Z}_{\geq p^*}$ and each compact set $K \subset \mathbb{R}^{Np}$ there exists $\varepsilon^* \in \mathbb{R}_{>0}$ and $T > 0$ such that for each $\varepsilon \in (0, \varepsilon^*)$ every solution of the closed-loop system with $\hat{w}(0) \in K$ generates trajectories satisfying

$$x_i(t) \in \mathcal{A}_i + \nu \mathbb{B},$$

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

IV. TEMPORAL RELAXATION: LEARNING WITH MEMORY AND NO COOPERATION

In the previous section we studied distributed ES dynamics with cooperation and no memory. In this section, we now discuss the opposite case. Namely, agents have access to memory and data, but there is no cooperation to estimate the parameters $w^\star$. In this case, condition (12) reduces to

$$\sum_{k=1}^{\mathbf{k}} \tilde{\phi}(t_k) \tilde{\phi}(t_k)^\top \succ 0, \tag{19}$$

and we recover the standard "richness" condition used in the setting of concurrent learning, e.g., [5], [16]. For networked systems this condition has been studied in [17, Assumption 2] by assuming that all agents of the network individually satisfy condition (19). Note, however, that condition (19) is stronger than the $\mathbf{k}$-CSR condition (12).

Example 3: Consider again the network of 2 agents studied in Examples 1 and 2, and let the sequence of times $\{t_k\}_k^{\mathbf{k}}$ be such that $t_k = (k-1)\pi + \pi/4$, for all $k \in \mathbb{Z}_{>0}$. Then, the normalized regressor vectors of the agents satisfy $\sum_{k=1}^{\mathbf{k}} \tilde{\phi}_1(t_k)\tilde{\phi}_1^\top(t_k) = \begin{bmatrix} \frac{\mathbf{k}}{2} & 0 \\ 0 & 0 \end{bmatrix}$, and $\sum_{k=1}^{\mathbf{k}} \tilde{\phi}_2(t_k)\tilde{\phi}_2^\top(t_k) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\mathbf{k}}{2} \end{bmatrix}$, respectively, which do not satisfy condition (19) individually for any $\mathbf{k} \in \mathbb{Z}_{>1}$. However, the overall multi-agent system satisfies condition (12), since $\sum_{k=1}^{\mathbf{k}} \sum_{i=1}^{N} \tilde{\phi}_2(t_k)\tilde{\phi}_1^\top(t_k) = \begin{bmatrix} \frac{\mathbf{k}}{2} & 0 \\ 0 & \frac{\mathbf{k}}{2} \end{bmatrix} \succ 0$, for any $\mathbf{k} \in \mathbb{Z}_{>1}$. □

When there is no cooperation between the agents, we recover the estimation results of [5], and since $a_{ij} = 0$ for all $i \neq j$, the individual learning dynamics (10) reduce to

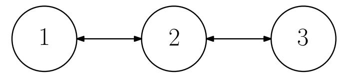

Fig. 1: Communication topology of the MAS.

$$\dot{\hat{w}}_i = -\alpha \sum_{k=0}^{\mathbf{k}} \Psi_i(x_i(t_k)) e_i(t_k, t), \tag{20}$$

where $\Psi_i(x_i(t_k))$ is given by (11). Note that under the stronger richness condition imposed by (19), each agent is now able to individually learn the optimal weights of a local cost function $f_i(x_i)$ that may be different from the cost functions of the other agents. Thus, under this stronger excitation assumption, we can now study distributed optimization problems with heterogenous cost functions. Using the data-enabled dynamics (20), a class of single-agent data-enabled extremum seeking (DEES) algorithms were recently proposed in [24].

V. NUMERICAL EXAMPLE

Consider a system of 3 agents, with a communication graph as the one shown in Figure 1. The agents aim to cooperatively solve local optimization problems of the form (1), with a homogenous cost function f. Each agent controls its own state $x_i \in \mathbb{R}$, and we assume that the common cost function has the quadratic form $f(x) = x^2 - 2x + 1$, and the individual sets of constraints are given by $K_1 = [1, 4]$, $K_2 = [0.9, 2.4]$ and $K_3 = [0.8, 2]$. We consider the following CODES:

$$\dot{\hat{w}}_i = -\alpha \sum_{k=0}^{\mathbf{k}} \Psi_i(x_i(t_k)) e_i(t_k, t) - \gamma \sum_{j \in \mathcal{N}_i} a_{ij}(\hat{w}_i - \hat{w}_j), \tag{21a}$$

$$\dot{x}_i = -x_i + \mathrm{P}_{K_i}\left(x_i - \hat{w}_i^\top \nabla \phi_i(x_i)\right), \tag{21b}$$

where the Lipschitz projection $\mathrm{P}_K(z) = \mathrm{argmin}_{y \in K} \|z - y\|$ is used to guarantee the satisfaction of the individual constraints for all time. Since the cost f is $\mathcal{C}^2$ and strongly convex, and the sets K_i are closed, convex, and bounded, by [9, Thm. 3], the optimization dynamics (21b) satisfy Assumption 3. Figure 2 shows the evolution in time of the solutions generated by the algorithm. As it can be observed, all trajectories converge to the global minimizer of $f(x)$ given by $x^* = 1$. The data was recorded by exciting each regressor vector during 5 seconds and collecting enough measurements to guarantee the satisfaction of the richness condition (12). As shown in Figure 3, the trajectories x_i also satisfy the individual constraints K_i during the initial transient. On the other hand, Figure 4 shows the behavior obtained when the richness condition (12) is not satisfied. In this case, the local estimations of the weights $\hat{w}_i$ do not converge to their true values, and the states x_i oscillate between the upper and lower bounds defined by the individual sets K_i.

VI. CONCLUSIONS

We presented a novel class of cooperative data-enabled extremum seeking (CODES) algorithms for static optimization

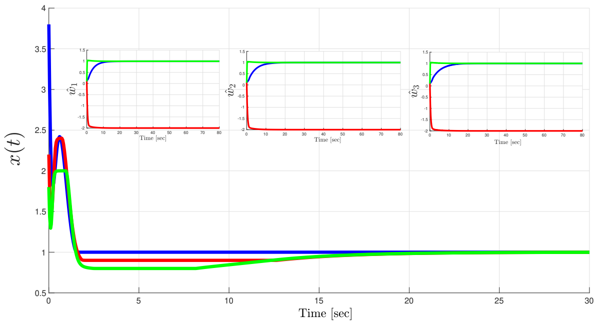

Fig. 2: Evolution of states $x_i(t)$. The inset shows the evolution of the weights $\hat{w}_i$.

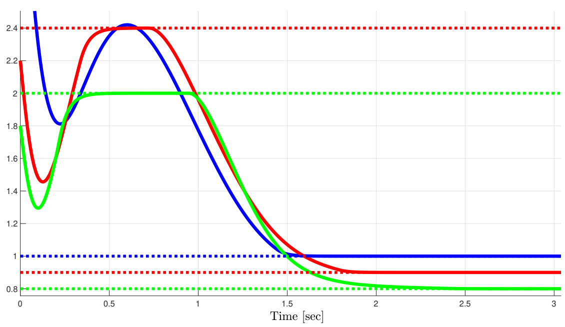

Fig. 3: Transient performance of the states x_i, satisfying the individual constraints K_i for all time.

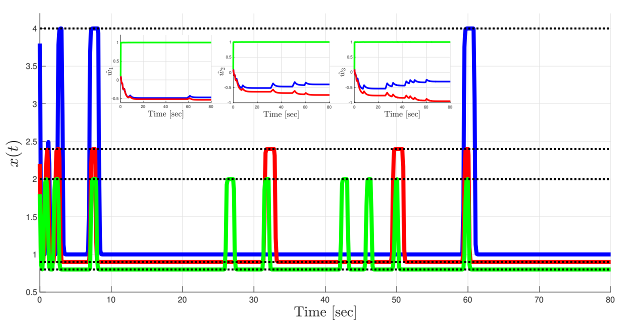

Fig. 4: Lack of convergence when the condition (12) is not satisfied. The inset shows the evolution of the states $\hat{w}_i$.

problems in network multi-agent systems with homogeneous cost functions and heterogenous constraints characterized by compact sets. Under a spatiotemporal richness condition on the recorded data of the MAS, convergence to a neighborhood of the set of optimizers can be guaranteed for a family of optimization dynamics with suitable regularity and stability properties. Future directions will study function approximations based on multi-layer neural networks, as well as time-varying communication graphs.

References

[1] V. Adetola and M. Guay. Parameter convergence in adaptive extremum-seeking control. *Automatica*, 43:105–110, 2007.

[2] M. Benosman. Model-based vs data-driven adaptive control: An overview. *International Journal of Adaptive Control and Signal Processing*, pages 753–776, 2018.

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

[4] W. Chen, S. Hua, and H. Zhang. Consensus-based distributed cooperative learning from closed-loop neural control systems. *IEEE Trans. on Neural Net. and Learning Syst.*, 26(2):331–345, 2015.

[5] G. Chowdhary and E. Johnson. Concurrent learning for convergence in adaptive control without persistence of excitation. *49th IEEE Conference on Decision and Control*, pages 3674–3679, 2010.

[6] G. Chowdhary, T. Wu, M. Cutler, N. K. Ure, and J. P. How. Experimental results of concurrent learning adaptive controllers. *AIAA Guidance, Navigation and Control Conference,*, pages 1–14, 2012.

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

[8] J. Feiling, S. Koga, M. Krstić, and T. R. Oliveira. Gradient extremum seeking for static maps with actuation dynamics governed by diffusion pdes. *Automatica*, 95:197–206, 2018.

[9] X. Gao. Exponential stability of globally projected dynamic systems. *IEEE Transactions on Neural Networks*, 14(2):426–431, 2003.

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

[11] M. Guay, I. Vandemeulen, S. Dougherty, and P. J. Mclellan. Distributed extremum-seeking control over networks of dynamic agents. *In proc. of American Control Conferencel*, pages 159–164, 2015.

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

[13] K. Hornik. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989.

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

[15] P. A. Ioannou and J. Sun. *Robust Adaptive Control*. Dover Publications Inc., Mineola, NY., 2012.

[16] R. Kamalapurkar, J. R. Klotz, and W. E. Dixon. Concurrent learning-based approximate feedback-nash equilibrium solution of n-player nonsero-sum differential games. *IEEE/CAA Journal of Automatica Sinica*, 1:239–247, 2014.

[17] J. Klotz, R. Kamalapurkar, and W. E. Dixon. Concurrent learning-based network synchronization. *American Cont. Conf.*, pages 796–801, 2014.

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

[19] K. Kvaternik and L. Pavel. An analytic framework for decentralized extremum seeking control. *In Proc. IEEE Ame. Con. Conf*, pages 3371–3376, 2012.

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

[21] J. I. Poveda, M. Benosman, and A. R. Teel. Distributed extremum seeking in multi-agent systems with arbitrary switching graphs. *IFAC World Congress*, 50(1):735–740, 2016.

[22] J. I. Poveda, M. Benosman, and A. R. Teel. Hybrid online learning control in networked multiagent systems: A survey. *International Journal of Adaptive Control and Signal Processing*, pages 1–34, 2018.

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

[24] J. I. Poveda, K. Vamoudakis, and M. Benosman. DEES: A class of data-enabled robust feedback algorithms for real-time optimization. *IFAC Symposium on Nonlinear Control Systems, to appear.*, 2019.

[25] J. I. Poveda, K. G. Vamvoudakis, and M. Benosman. A neuro-adaptive architecture for extremum seeking control using hybrid learning dynamics. *In proc. of American Control Conferece. Boston, MA.*, pages 542–547, 2017.

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

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

[28] I. Vandemeulen, M. Guay, and P. J. McLellan. Discrete-time distributed extremum seeking control over networks with unstable dynamics. *IEEE Transactions on Control of Network Systems*, 2017.