



HAL
open science

Distributed Hybrid Gradient Algorithm with Application to Cooperative Adaptive Estimation

Mohamed Maghenem, Antonio Loria, Adnane Saoud

► **To cite this version:**

Mohamed Maghenem, Antonio Loria, Adnane Saoud. Distributed Hybrid Gradient Algorithm with Application to Cooperative Adaptive Estimation. HSCC 2022 - 25th ACM International Conference on Hybrid Systems: Computation and Control, May 2022, Milan, Italy. 10.1145/3501710.3519528 . hal-03789071

HAL Id: hal-03789071

<https://hal.science/hal-03789071>

Submitted on 27 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed Hybrid Gradient Algorithm with Application to Cooperative Adaptive Estimation

Mohamed Maghenem
GIPSA-lab, CNRS
Grenoble, France
mohamed.maghenem@gipsa-lab.fr

Adnane Saoud
L2S-CentraleSupélec, Univ
Paris-Saclay
Gif sur Yvette, France
adnane.saoud@centralesupelec.fr

Antonio Loria
L2S, CNRS
Gif sur Yvette, France
antonio.loria@cnrs.fr

ABSTRACT

We address a classical identification problem that consists in estimating a vector of constant unknown parameters from a given linear input/output relationship. The proposed method relies on a network of gradient-descent-based estimators, each of which exploits only a portion of the input-output data. A key feature of the method is that the input-output signals are hybrid, so they may evolve in continuous time (*i.e.*, they may *flow*), or they may change at isolated time instances (*i.e.*, they may *jump*). The estimators are interconnected over a weakly-connected directed graph, so the alternation of flows and jumps combined with the distributed character of the algorithm introduce a rich behavior that is impossible to obtain using continuous- or discrete-time estimators. A condition of *persistence of excitation* in hybrid form ensures exponential convergence of the estimation errors. The proposed approach generalizes the existing centralized gradient-descent algorithms and yields relaxed sufficient conditions for (uniform-exponential) parameter estimation. In addition, we address the observation/identification problem for a class of hybrid systems with unknown parameters using a distributed network of adaptive observers/identifiers.

CCS CONCEPTS

• **Theory of computation** → **Logic and verification; Modal and temporal logics; Linear logic**; • **Computer systems organization** → Embedded and cyber-physical systems;

KEYWORDS

Gradient Descent Algorithm; Distributed Estimation; Hybrid Systems; Lyapunov Direct Method

ACM Reference Format:

Mohamed Maghenem, Adnane Saoud, and Antonio Loria. 2022. Distributed Hybrid Gradient Algorithm with Application to Cooperative Adaptive Estimation. In *25rd ACM International Conference on Hybrid Systems: Computation and Control (HSCC '22)*, April 04–06, 2022, Milan, Italy. ACM, New York, NY, USA, 10 pages. <https://doi.org/...>

The work of M. Maghenem and A. Loria was supported by the ANR (project "HANDY") under the contract ANR-18-CE40-0010.

1 INTRODUCTION

Estimation theory is key in many applications such as autonomous navigation [2, 28], weather predictions [8, 14], applied economics [4], etc. In many cases, the estimation problem consists in estimating the parameters $\theta \in \mathbb{R}^{m_\theta}$ of the linear regression model of a measurable output $y \in \mathbb{R}$ satisfying the relationship $y = \psi^\top \theta$ [27]. In such model, ψ , which is generally referred to as the *regressor*, is a map from its domain, denoted by 'dom ψ ', to an Euclidean space of dimension m_θ . An elementary, but efficient estimation algorithm is the so-called *gradient-based*, which updates the value of a parameter estimate $\hat{\theta}$ using the gradient of the cost function $J(e) := \frac{e^2}{2}$, where $e := \hat{y} - y$ and $\hat{y} := \psi^\top \hat{\theta}$ is the estimated output, function of the parameter estimate $\hat{\theta}$. Thus, the estimation problem consists in designing an update algorithm for $\hat{\theta}$ in function of the computable error e .

There are several ways to *update* the estimate $\hat{\theta}$. Most commonly, time is tracked by a real variable $t \in \mathbb{R}_{\geq 0}$ [11, 25] or by a sequence of discrete instances $j \in \mathbb{N}$ [27]. In the first case, dom $\psi = \mathbb{R}_{\geq 0}$, so, for each $t \in \mathbb{R}_{\geq 0}$, $e(t) := \hat{y}(t) - y(t) = \tilde{\theta}(t)^\top \psi(t)$, where $\tilde{\theta} := \hat{\theta} - \theta$. Then, the *gradient-descent* algorithm is defined by the ordinary differential equation

$$\dot{\hat{\theta}}(t) = -\gamma \nabla_{\hat{\theta}} J(e(t)) = -\gamma \psi(t) [\psi(t)^\top \hat{\theta}(t) - y(t)], \quad (1)$$

where $\gamma > 0$ is a design parameter and $\nabla_{\hat{\theta}} J$ denotes the gradient of $J(e)$ with respect to $\hat{\theta}$ [16], so the behavior of $\hat{\theta}(t)$ is governed by

$$\dot{\tilde{\theta}}(t) = -\gamma \psi(t) \psi(t)^\top \tilde{\theta}(t). \quad (2)$$

The convergence of the estimation errors $\tilde{\theta}(t)$ may be assessed by establishing the stronger property of uniform (in the initial time) exponential stability (UES) of the origin $\{\tilde{\theta} = 0\}$ for (2). In the case that ψ is bounded on its domain, UES is known (see, *e.g.*, [15]) to be equivalent to ψ possessing the continuous-time persistency-of-excitation (CT-PE) property:

ASSUMPTION 1 (CT-PE). *There exist $T, \mu > 0$ such that, for each $t_0 \geq 0$, $\int_{t_0}^{t_0+T} \psi(s) \psi(s)^\top ds \geq \mu I_{m_\theta}$.*

When y is measured at certain sampled instants of time, that is, dom $y = \text{dom } \psi = \mathbb{N}$, the gradient algorithm may be implemented in discrete time, using the difference equation

$$\hat{\theta}(j+1) = \hat{\theta}(j) - \sigma(j) \nabla_{\hat{\theta}} J(e(j)), \quad (3)$$

where $\sigma(j) := \frac{\gamma}{1 + \gamma |\psi(j)|^2}$ and $\gamma > 0$ is the adaptation rate [27]. As a result, the dynamics of the estimation error is given by

$$\tilde{\theta}(j+1) = \left(I - \frac{\gamma \psi(j) \psi(j)^\top}{1 + \gamma |\psi(j)|^2} \right) \tilde{\theta}(j). \quad (4)$$

It is known, at least since [1], that the following hypothesis is necessary and sufficient to ensure UES of the set $\{\tilde{\theta} = 0\}$ for (4).

ASSUMPTION 2 (DT-PE). *There exist $J, \mu > 0$, such that, for each $j_0 \geq 0$, $\sum_{j=j_0}^{j_0+J} \psi(j)\psi(j)^\top \geq \mu I_{m_\theta}$.*

While the necessity of persistency of excitation is well established in the literature, we can see that the form of such condition is determined by how ψ is defined mathematically. Clearly, the nature of $\text{dom } \psi$ leads to different mathematical conditions.

In [24], the regressor function ψ is viewed as a general ‘hybrid signal’ (in the sense of [9]), so it is allowed to exhibit both continuous and discrete behaviors. That is, the linear regression model becomes $y(t, j) = \psi(t, j)^\top \theta$, where $\text{dom } \psi$ (hence $\text{dom } y$) is a hybrid time domain in which t is the ordinary time capturing the continuous evolution of ψ and j captures the instantaneous jumps of ψ . Correspondingly, the estimate \hat{y} is defined as $\hat{y}(t, j) := \psi(t, j)^\top \hat{\theta}(t, j)$. As a result, for $((t, j), (t, j+1)) \in \text{dom } \psi \times \text{dom } \psi$, i.e., when ψ jumps, $\hat{\theta}$ is updated according to the discrete-time law (3), which in hybrid time becomes

$$\hat{\theta}(t, j+1) = \hat{\theta}(t, j) - \sigma(t, j)\psi(t, j)e(t, j), \quad \sigma(t, j) := \frac{\gamma}{1 + \gamma|\psi(t, j)|^2}.$$

On the other hand, when ψ flows, i.e., for all $(t, j) \in \text{dom } \psi$ with $(t, j+1) \notin \text{dom } \psi$, we update $\hat{\theta}$ according to the continuous-time adaptation law (1), that is,

$$\dot{\hat{\theta}}(t, j) = -\gamma\psi(t, j)\psi(t, j)^\top \tilde{\theta}(t, j).$$

The choice of using a hybrid gradient algorithm is not of pure academic interest. It is showed in [24] that parametric convergence, and more precisely UES of the set $\{\tilde{\theta} = 0\}$, is guaranteed under a relaxed form of persistence of excitation tailored for hybrid systems, which is less restrictive than the CT- or the DT-PE. In [24], the global behavior of $\tilde{\theta}$ is modeled and analyzed using the hybrid-systems framework in [9]. In particular, UES of the set $\{\tilde{\theta} = 0\}$ is guaranteed under the following hybrid persistence of excitation (HPE) condition that generalizes Assumptions 1 and 2.

ASSUMPTION 3 (HPE). *There exist $K, \mu > 0$ such that, for each $\bigcup_{j=j_0}^J ([t_j, t_{j+1}] \times \{j\}) \subset \text{dom } \psi$ with $(t_{j+1} - t_{j_0}) + (J - j_0) \geq K$, we have*

$$\sum_{j=j_0}^J \left[\int_{t_j}^{t_{j+1}} \psi(\tau, j)\psi(\tau, j)^\top d\tau + \frac{\gamma\psi(t_{j+1}, j)\psi(t_{j+1}, j)^\top}{2 + 2\gamma|\psi(t_{j+1}, j)|^2} \right] \geq \mu I_{m_\theta}.$$

Roughly speaking, the HPE condition in Assumption 3 is less conservative than its counter-parts CT- and DT-PE since it captures the fact that the richness of a signal may be enhanced by an appropriate mingling of exciting flows and jumps, which, otherwise, are insufficient to guarantee that either the CT- or the DT-PE hold—see [24] for an example.

In this paper, inspired by [5, 12, 19], we carry a step further the relaxation of the conditions for parameter estimation using the hybrid-gradient estimation algorithm. We propose an algorithm that relies on a network of estimators, each of which is designed using a hybrid gradient-based adaptation law and has access to only limited data, involving the parameters of the global process. However, to compensate for the lack of data, the different estimators are interconnected, so that they share their estimates, with

the goal of achieving consensus. Furthermore, we use our main result as a building block to address the observation/identification problem for a class of nonlinear hybrid systems with unknown parameters. After decomposing the hybrid system into (reduced-order) interconnected subsystems, a reduced-order hybrid adaptive observer/estimator is assigned to each subsystem. To that end, we borrow inspiration from [10, 30]. The different observers/estimators form a network and exchange their estimates of the parameters to reconstruct the state as well as the unknown parameters.

The considered distributed-estimation framework is useful in many contexts, including the following:

- (1) For large-scale systems, where local estimators are mounted on local sensors, so each estimator has access only to local data involving the parameters of the global system. Interconnecting the local estimators, so that they share their estimates, serves the estimation process. The idea *per se* is not new; distributed observers have been proposed in the literature of continuous- and discrete-time systems separately, e.g., in [13, 26, 29], but our hybrid algorithms are illustrated to perform better.
- (2) When reinforcing a single real-time estimator using old data. For example, when a single estimator processing real-time data has also access to old data, the old data may be exploited by designing several virtual estimators to interact with the one treating the real-time data; see the forthcoming Section 3.2 for more details. Such estimation-design approach fits within the paradigm of *exploration vs exploitation* as it is the case in concurrent-learning control strategies [7, 20].
- (3) In homogeneous networks of dynamical systems, where the agents need to estimate some of there (equal) parameters while achieving a cooperative task. Instead of each agent estimating the parameters on its own, a cooperative strategy can yield better convergence properties under relaxed conditions [5].

The considered distributed estimation framework has been studied in the literature for the continuous-time case only [5, 12, 19]. Indeed, under different graph conditions, distributed (continuous-time) persistence of excitation conditions are shown to be necessary and sufficient to guarantee UES of the set $\{\tilde{\theta} = 0\}$. However, to the best of our knowledge, estimating the convergence rate has not been addressed in existing literature. Furthermore, the whole problem is unsolved in the hybrid case.

The rest of the paper is organized as follows. In Section 2, we recall some basic concepts of hybrid systems and graph theory. In Section 3, we formulate the gradient-based estimation problem in a distributed setting and provide our main statement. In Section 4, we address the cooperative observation/identification problem for a class of hybrid systems with unknown parameters and, in Section 5, we present several illustrative examples.

2 PRELIMINARIES

Notations. $\mathbb{R}_{\geq 0} := [0, \infty)$, $\mathbb{N} := \{0, 1, \dots, \infty\}$. Given a nonempty set $K \subset \mathbb{R}^{m_x}$, $|x|_K := \inf_{y \in K} |x - y|$, $\text{int}(K)$ denotes the interior of K , and $\text{cl}(K)$ denotes its closure. For a nonempty set $O \subset \mathbb{R}^n$, $K \setminus O$ denotes the subset of elements of K that are not in O . For a map $\phi : \mathbb{R}_{\geq 0} \times \mathbb{N} \rightarrow \mathbb{R}^{m_x} \cup \{\emptyset\}$, $\text{dom } \phi \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$ denotes the domain of definition of ϕ . Furthermore, $|\phi|_\infty := \sup\{\phi(t, j) : (t, j) \in \text{dom } \phi\}$. For a differentiable map $(x) \mapsto V(x) \in \mathbb{R}$, ∇V denotes the gradient

of V with respect to x . For a symmetric semi-positive definite matrix $\Gamma \in \mathbb{R}^{m_x \times m_x}$, $\lambda_{\min}(\Gamma)$ and $|\Gamma| := \lambda_{\max}(\Gamma)$ denote the smallest and the largest eigenvalues of Γ , and the trace of Γ , respectively. By $\mathbf{1}_{m_x} \in \mathbb{R}^{m_x}$, we denote the vector whose entries are equal to $1/\sqrt{m_x}$. For a sequence $\{A_i\}_{i=1}^{m_y} \subset \mathbb{R}^{m_x \times m_x}$, $\text{blkdiag}\{A_i\}$ is the block-diagonal matrix whose i th diagonal block corresponds to the matrix A_i . Finally, \otimes denotes the Kronecker product.

2.1 On hybrid systems

Following [9], we view a hybrid dynamical system, denoted by $\mathcal{H} := (C, F, D, G)$, as a combination of a constrained differential and a constrained difference equations given by

$$\mathcal{H} : \begin{cases} \dot{x} = F(x) & x \in C \\ x^+ = G(x) & x \in D, \end{cases} \quad (5)$$

with the state variable $x \in \mathcal{X} \subset \mathbb{R}^{m_x}$, the flow set $C \subset \mathcal{X}$, the jump set $D \subset \mathcal{X}$, the flow and jump maps $F : C \rightarrow \mathbb{R}^{m_x}$ and $G : D \rightarrow \mathbb{R}^{m_x}$, respectively.

Solutions to hybrid systems as defined above, are defined using the following concept.

Definition 2.1 (Hybrid arc). A map $\phi : \text{dom } \phi \rightarrow \mathbb{R}^{m_x}$ with $\text{dom } \phi \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$ is a hybrid arc if ϕ is parameterized by an ordinary time variable $t \in \mathbb{R}_{\geq 0}$ and a discrete jump variable $j \in \mathbb{N}$ and $\text{dom } \phi$ is such that, for each $(T, J) \in \text{dom } \phi$, $\text{dom } \phi \cap ([0, T] \times \{0, 1, \dots, J\}) = \bigcup_{j=0}^J ([t_j, t_{j+1}] \times \{j\})$ for a sequence $\{t_j\}_{j=0}^{J+1}$, such that $t_{j+1} \geq t_j$, $t_0 = 0$, and $t_{J+1} = T$.

Definition 2.2 (Concept of solution to \mathcal{H}). A hybrid arc $\phi : \text{dom } \phi \rightarrow \mathbb{R}^{m_x}$ is a solution to \mathcal{H} if

- (S0) $\phi(0, 0) \in \text{cl}(C) \cup D$;
- (S1) for all $j \in \mathbb{N}$ such that $I^j := \{t : (t, j) \in \text{dom } \phi\}$ has nonempty interior, $t \mapsto \phi(t, j)$ is locally absolutely continuous and

$$\begin{aligned} \phi(t, j) &\in C && \text{for all } t \in \text{int}(I^j), \\ \dot{\phi}(t, j) &= F(\phi(t, j)) && \text{for almost all } t \in I^j; \end{aligned}$$
- (S2) for all $(t, j) \in \text{dom } \phi$ such that $(t, j+1) \in \text{dom } \phi$, we have $\phi(t, j) \in D$ and $\phi(t, j+1) = G(\phi(t, j))$.

Next, we recall from [23] the definition of uniform exponential stability (UES) of a general closed set for a hybrid system.

Definition 2.3 (UES). Consider the hybrid system $\mathcal{H} := (C, F, D, G)$ and let $\mathcal{S} \subset \mathcal{X}$ be a closed set. The set \mathcal{S} is said to be UES for \mathcal{H} on $\mathcal{X}_o \subset \mathcal{X}$ if there exist $\kappa, \lambda > 0$ such that, each solution ϕ to \mathcal{H} starting from $\phi(0, 0) \in \mathcal{X}_o$, we have

$$|\phi(t, j)|_{\mathcal{S}} \leq \kappa \exp^{-\lambda(t+j)} |\phi(0, 0)|_{\mathcal{S}} \quad \forall (t, j) \in \text{dom } \phi. \quad (6)$$

When $\mathcal{X}_o = \mathcal{X}$, we say \mathcal{S} is said to be UES for \mathcal{H} .

The constant λ in (6) is called a convergence rate of the solutions to \mathcal{H} towards \mathcal{S} . Finally, we recall a characterization of UES for hybrid systems using Lyapunov functions.

LEMMA 2.4. *A closed subset $\mathcal{S} \subset \mathcal{X}$ is UES for \mathcal{H} on \mathcal{X}_o if there exists a function $V : \mathbb{R}_{\geq 0} \times \mathbb{N} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ and positive constants $(\bar{\lambda}, \underline{\lambda}, \lambda_c, \lambda_d)$ such that*

$$\underline{\lambda}|x|_{\mathcal{S}}^2 \leq V(t, j, x) \leq \bar{\lambda}|x|_{\mathcal{S}}^2, \quad (7)$$

and, along each solution ϕ to \mathcal{H} starting from \mathcal{X}_o , we have

- For each $j \in \mathbb{N}$ such that $\text{int}(I^j) \neq \emptyset$, $t \mapsto V(t, j, \phi(t, j))$ is absolutely continuous on I^j and

$$\dot{V}(t, j, \phi(t, j)) \leq -\lambda_c |\phi(t, j)|_{\mathcal{S}} \quad \text{for almost all } t \in I^j.$$
- For each $((t, j), (t, j+1)) \in \text{dom } \phi \times \text{dom } \phi$,

$$V(t, j+1, \phi(t, j+1)) - V(t, j, \phi(t, j)) \leq -\lambda_d |\phi(t, j)|_{\mathcal{S}}.$$

Note that when the conditions in Theorem 4.1 are satisfied, a lower bound of the convergence rate is given by

$$\lambda := \min \left\{ \frac{\lambda_c}{\bar{\lambda}}, -\ln \left[1 - \frac{\lambda_d}{\bar{\lambda}} \right] \right\}.$$

2.2 On Graph theory

To formulate our distributed hybrid excitation condition, we lay on some graph-theoretical notions [22]. A directed graph or a digraph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is characterized by the set of nodes $\mathcal{V} = \{1, 2, \dots, m_y\}$, and the set of directed edges \mathcal{E} . The edge set \mathcal{E} consists of ordered pairs of the form (k, i) , which indicates a directed link from node k to node i . We assume that the digraphs are simple, i.e., there are no self-arcs. Note that when there exists a directed edge $(k, i) \in \mathcal{E}$, then node k is called an *in-neighbor* of node i . We assign a positive weight a_{ik} to each edge (k, i) . That is, $a_{ik} = 0$ if (k, i) is not an edge. A digraph is strongly connected if for any two distinct nodes i and j , there is a path from i to j . A digraph \mathcal{G} is weakly connected if the undirected graph obtained by ignoring the orientations of the edges is connected.

3 DISTRIBUTED HYBRID GRADIENT ALGORITHM

For simplicity, in the introduction, we discussed a classical parameter estimation problem based on *one* measurable output and a vectorial regressor function ψ . Nonetheless, the gradient algorithms previously presented also apply to multi-output systems, where $y \in \mathbb{R}^{m_y}$ for some integer $m_y > 1$. In this case, $\psi \in \mathbb{R}^{m_\theta \times m_y}$, so $\psi \psi^\top \in \mathbb{R}^{m_\theta \times m_\theta}$. Now let $y := [y_1 \cdots y_{m_y}]^\top$ and assume that each component y_i is available to a different estimator, but the different estimators exchange the values of their estimates over a network. With this in mind, we consider m_y estimators and, for each $i \in \{1, 2, \dots, m_y\}$, we introduce a hybrid regressor function, denoted by $\psi_i : \text{dom } \psi_i \rightarrow \mathbb{R}^{m_\theta}$, and an estimate of θ , denoted by $\hat{\theta}_i$, which is yet to be defined dynamically. Now, akin to the case of a single hybrid estimator described in the introduction, we define the output of the i th estimator as

$$\hat{y}_i(t, j) = \psi_i(t, j)^\top \hat{\theta}_i(t, j) \quad i \in \{1, 2, \dots, m_y\}. \quad (8)$$

Correspondingly, the error between the actual and the estimated output for the i -th estimator is given by

$$e_i(t, j) := \hat{y}_i(t, j) - y_i(t, j) = \psi_i(t, j)^\top \tilde{\theta}_i(t, j), \quad \tilde{\theta}_i := \hat{\theta}_i - \theta. \quad (9)$$

At this point, we introduce a standing assumption on the regularity of ψ_i for all i .

(SA1) For each $j \in \mathbb{N}$, the map $t \mapsto \psi(t, j)$ is continuous on $\text{int}(I^j)$, where $I^j := \{t : (t, j) \in \text{dom } \psi\}$. Furthermore, there exists $\bar{\psi} > 0$ such that $|\psi|_\infty \leq \bar{\psi}$.

REMARK 1. (SA1) is not restrictive, especially when the regressor ψ is excited by an external input that is a design parameter.

To achieve the identification of θ , the different estimators are designed to exchange their estimates, in real (hybrid) time, according to a communication graph denoted by \mathcal{G} . That is, similarly to the algorithm proposed in [12] for distributed estimation in continuous time, we design an update law for each $\hat{\theta}_i$ along the discrete-time evolution and along the continuous-time evolution of ψ as follows.

- For all $(t, j) \in \text{dom } \psi_i$ and $(t, j+1) \in \text{dom } \psi$ (i.e., when ψ jumps), inspired by [17], we set the update law for $\hat{\theta}_i$ to

$$\hat{\theta}_i^+ = \hat{\theta}_i - \sigma_i(t, j)\psi_i(t, j)e_i - \delta \sum_{k=1}^{m_y} a_{ik}(\hat{\theta}_i - \hat{\theta}_k), \quad (10)$$

where $\delta > 0$ is an interconnection gain, $\gamma > 0$, and $\sigma_i(t, j) := \frac{\gamma}{1+\gamma|\psi_i(t, j)|^2}$.

- On the other hand, when $(t, j) \in \text{dom } \psi$ and $(t, j+1) \notin \text{dom } \psi$ (i.e., ψ flows), we update $\hat{\theta}$ according to

$$\dot{\hat{\theta}}_i = -\gamma_1\psi_i(t, j)e_i^\top - \gamma_2 \sum_{k=1}^{m_y} a_{ik}(\hat{\theta}_i - \hat{\theta}_k). \quad (11)$$

The coefficients $a_{ik} \geq 0$ are positive if a reliable *directed* (that is, in general, $a_{ik} \neq a_{ki}$) interconnection exists between the i th and k th nodes and equal to zero otherwise. Thus, the estimators are assumed to be interconnected forming a network of multi-agent systems that can be modeled using a directed graph \mathcal{G} —see Section 2.2. For this graph, we introduce the following standing assumption: (SA2) The graph \mathcal{G} is unidirectional and weakly connected.

Recall that a weakly connected digraph \mathcal{G} can be decomposed into S strongly connected subgraphs $\{\mathcal{G}_s(\mathcal{V}_s, \mathcal{E}_s)\}_{s=1}^S$, for some $S \in \{1, 2, \dots, m_y\}$. Among the strongly connected subgraphs in \mathcal{G} , we identify the *leading* strongly connected subgraphs $\{\mathcal{G}_s\}_{s \in S_l}$, $S_l \subset \{1, 2, \dots, S\}$, with no in-neighbors from any other strongly connected subgraph.

REMARK 2. (SA2) is the least restrictive graph requirement, since the lack of it implies that some nodes are disconnected from the network.

Now we introduce the Laplacian matrix $\mathcal{L} := [l_{ik}] \in \mathbb{R}^{m_y \times m_y}$ with $l_{ik} := -a_{ik} \leq 0$ for $k \neq i$ and $\sum_{k=1}^{m_y} a_{ik}$ otherwise. Thus, we may express the right-hand sides of (10) and (11) as

$$\dot{\hat{\theta}} = -\gamma_1\Phi_c(t, j) - \gamma_2(\mathcal{L} \otimes I_{m_\theta})\hat{\theta}, \quad (12a)$$

$$\Phi_c(t, j) := \text{blkdiag}\{\Phi_{c1}(t, j), \Phi_{c2}(t, j)\}, \quad (12b)$$

$$\Phi_{c1}(t, j) := \gamma_1 \text{blkdiag} \left\{ \psi_i(t, j)\psi_i(t, j)^\top \right\}_{i \in \mathcal{V}_1}, \quad (12c)$$

$$\Phi_{c2}(t, j) := \gamma_1 \text{blkdiag} \left\{ \psi_i(t, j)\psi_i(t, j)^\top \right\}_{i \in \mathcal{V} \setminus \mathcal{V}_1}, \quad (12d)$$

for all times such that the solutions flow, and as

$$\hat{\theta}^+ = \hat{\theta} - \Phi_d(t, j)\hat{\theta} - \delta(\mathcal{L} \otimes I_{m_\theta})\hat{\theta}, \quad (13a)$$

$$\Phi_d(t, j) := \text{blkdiag}\{\Phi_{d1}(t, j), \Phi_{d2}(t, j)\}, \quad (13b)$$

$$\Phi_{d1}(t, j) := \text{blkdiag} \left\{ \frac{\gamma\psi_i(t, j)\psi_i(t, j)^\top}{1+\gamma|\psi_i(t, j)|^2} \right\}_{i \in \mathcal{V}_1}, \quad (13c)$$

$$\Phi_{d2}(t, j) := \text{blkdiag} \left\{ \frac{\gamma\psi_i(t, j)\psi_i(t, j)^\top}{1+\gamma|\psi_i(t, j)|^2} \right\}_{i \in \mathcal{V} \setminus \mathcal{V}_1}. \quad (13d)$$

for all times such that the solutions jump.

REMARK 3. Instead of the discrete-time distributed gradient algorithm in (10), one can use, for each $i \in \{1, 2, \dots, m_y\}$,

$$\hat{\theta}_i(t, j+1) = \frac{\hat{\theta}_i(t, j) + \sum_{k=1}^{m_y} a_{ik}\hat{\theta}_k(t, j)}{1 + \sum_{k=1}^{m_y} a_{ik}} - \sigma_i(t, j)\psi_i(t, j)e_i, \quad (14)$$

where $\sigma_i(t, j) := \frac{\gamma}{1+\gamma|\psi_i(t, j)|^2}$. Hence, in the error coordinates, (14) becomes

$$\tilde{\theta}^+ = -\Phi_d(t, j)\tilde{\theta} + \left((I_{m_y} + \mathcal{D})^{-1}(I_{m_y} + \mathcal{A}) \otimes I_{m_\theta} \right) \tilde{\theta},$$

where $\mathcal{A} := [a_{ik}] \in \mathbb{R}^{m_y \times m_y}$ is the adjacency matrix associated with the digraph by setting $a_{ii} := 0$ for all $i \in \{1, 2, \dots, m_y\}$ and $\mathcal{D} :=$

$\text{blkdiag} \left\{ \sum_{k=1}^{m_y} a_{ik} \right\}_{i \in \{1, 2, \dots, m_y\}}$. Note that solving the estimation problem under the discrete update in (14) is an open problem.

3.1 Main Result I

Our main statement is that the distributed hybrid estimation laws (12)-(13) guarantee the estimation of θ provided a property that we call distributed hybrid persistency of excitation (DHPE) holds.

ASSUMPTION 4 (DHPE). For each $s \in S_l$, there exist $K_s, \mu_s > 0$ such that, for each hybrid time domain $\bigcup_{j=j_0}^J ([t_j, t_{j+1}] \times \{j\}) \subset \text{dom } \psi$ with $(t_{j+1} - t_{j_0}) + (J - j_0) \geq K_s$, we have

$$\sum_{i \in \mathcal{V}_s} \sum_{j=j_0}^J \left[\int_{t_j}^{t_{j+1}} \psi_i(\tau, j)\psi_i(\tau, j)^\top d\tau + \frac{\gamma\psi_i(t_{j+1}, j)\psi_i(t_{j+1}, j)^\top}{2 + 2\gamma|\psi_i(t_{j+1}, j)|^2} \right] \geq m_s \mu_s I_{m_\theta}.$$

REMARK 4. Note that Assumption 4 is the hybrid counterpart of the distributed persistence of excitation condition used in [12], which is shown to be necessary and sufficient for convergence in the continuous-time setting.

Under the DHPE condition, we establish uniform exponential stability of $\{\tilde{\theta} = 0\}$ in the hybrid sense (see Def. 2.3) for the following hybrid dynamical system defined after the set of equations (12a) and (13a). That is,

$$\mathcal{H}_{dg} : \begin{cases} \begin{cases} \dot{\tilde{\theta}} &= -\gamma_1\Phi_c(t, j)\tilde{\theta} - \gamma_2(\mathcal{L} \otimes I_{m_\theta})\tilde{\theta} \\ \dot{t} &= 1 \\ \dot{j} &= 0 \end{cases} & (\tilde{\theta}, t, j) \in C_g \\ \begin{cases} \tilde{\theta}^+ &= \tilde{\theta} - \Phi_d(t, j)\tilde{\theta} - \delta(\mathcal{L} \otimes I_{m_\theta})\tilde{\theta} \\ t^+ &= t \\ j^+ &= j+1 \end{cases} & (\tilde{\theta}, t, j) \in D_g, \end{cases} \quad (15a)$$

where $D_g := \{(\tilde{\theta}, t, j) \in \mathbb{R}^{m_\theta} \times \text{dom } \psi : (t, j+1) \in \text{dom } \psi\}$ and $C_g := \text{cl}((\mathbb{R}^{m_\theta} \times \text{dom } \psi) \setminus D_g)$.

The formulation of our main statement relies on a specific representation of the Laplacian that holds under the standing assumption (SA2). More precisely, we recall that by an appropriate reordering of the nodes of \mathcal{G} , the Laplacian matrix \mathcal{L} of a weakly connected digraph \mathcal{G} can be written in the lower-block triangular form

$$\mathcal{L} = \begin{bmatrix} \text{blkdiag}\{\mathcal{L}_s\} & 0 \\ s \in S_l & \\ M_l & M_r \end{bmatrix},$$

where each $\mathcal{L}_s \in \mathbb{R}^{m_s \times m_s}$ is the Laplacian matrix of the strongly connected component \mathcal{G}_s , the lower-left block $M_l \in \mathbb{R}^{k \times m_y - k}$, $k := m_y - \sum_{s \in \mathcal{S}_l} m_s$, is a non-positive matrix, and the lower-right block $M_r \in \mathbb{R}^{k \times k}$ is a non-singular M-matrix; namely, $M_r := \gamma_r I_k - C_r$, where $C_r \in \mathbb{R}^{k \times k}$ is a non-negative matrix, $\gamma_r > \rho(C_r) > 0$, and $\rho(C_r)$ is the spectral radius of C_r ; see [6, 12] for more details.

For the sequel, we assume the following.

(SA3) The graph \mathcal{G} admits only one leading strongly connected component $\mathcal{G}_1(\mathcal{V}_1, \mathcal{E}_1)$; namely, $\mathcal{S}_l = \{1\}$.

Under (SA3), the Laplacian matrix \mathcal{L} enjoys the previous lower-triangular structure with a simple northeast block, that is,

$$\mathcal{L} = \begin{bmatrix} \mathcal{L}_1 & 0 \\ M_l & M_r \end{bmatrix}, \quad (16)$$

where $\mathcal{L}_1 \in \mathbb{R}^{m_1 \times m_1}$ is the Laplacian matrix of the strongly connected component \mathcal{G}_1 . It is to be noted that our approach extends trivially if we omit assume (SA3).

As a result, we decompose the state $\tilde{\theta}$ of \mathcal{H}_{dg} into $\tilde{\theta} := (\tilde{\theta}_s, \tilde{\theta}_f)$, where $\tilde{\theta}_s$ contains the state of the elements in the strongly connected graph \mathcal{G}_1 and $\tilde{\theta}_f$ contains the state of the elements of \mathcal{G} that are not in \mathcal{G}_1 . Using this decomposition, the hybrid system \mathcal{H}_{dg} can be expressed in the following cascaded form

$$\mathcal{H}_{dg} : \begin{cases} \begin{cases} \dot{\tilde{\theta}}_f = -\gamma_1 \Phi_{c2}(t, j) \tilde{\theta}_f - \gamma_2 (M_r \otimes I_{m_\theta}) \tilde{\theta}_f - \gamma_2 (M_l \otimes I_{m_\theta}) \tilde{\theta}_s \\ \dot{\tilde{\theta}}_s = -\gamma_1 \Phi_{c1}(t, j) \tilde{\theta}_s - \gamma_2 (\mathcal{L}_1 \otimes I_{m_\theta}) \tilde{\theta}_s \\ i = 1 \\ j = 0 \end{cases} & (\tilde{\theta}, t, j) \in C_g \\ \begin{cases} \tilde{\theta}_f^+ = \tilde{\theta}_f - \Phi_{d2}(t, j) \tilde{\theta}_f - \delta (M_r \otimes I_{m_\theta}) \tilde{\theta}_f - \delta (M_l \otimes I_{m_\theta}) \tilde{\theta}_s \\ \tilde{\theta}_s^+ = \tilde{\theta}_s - \Phi_{d1}(t, j) \tilde{\theta}_s - \delta (\mathcal{L}_1 \otimes I_{m_\theta}) \tilde{\theta}_s \\ t^+ = t \\ j^+ = j + 1 \end{cases} & (\tilde{\theta}, t, j) \in D_g. \end{cases}$$

THEOREM 3.1. Consider the hybrid system \mathcal{H}_{dg} under Assumption 4. Let SA2 generate \mathcal{L}_1 , M_l , and M_r such that (16) hold and let $\delta > 0$ be such that

$$Q_r := P_r M_r + M_r^T P_r \geq 4\delta M_r^T P_r M_r, \quad (17)$$

where $P_r := \left(\text{blkdiag}_{i \in \{1, 2, \dots, k\}} \{M_r^{-1} 1_k\} \right)^{-1}$,

$$P \mathcal{L}_1 + \mathcal{L}_1^T P - 4\delta \mathcal{L}_1^T P \mathcal{L}_1 \geq 0, \quad (18)$$

$P := \text{blkdiag}_{i \in \{1, 2, \dots, m_1\}} \{w_i\}$, and w is the left eigenvector of \mathcal{L}_1 associated to its null eigenvalue. Then, the set

$$\mathcal{S} := \{(\tilde{\theta}, t, j) \in \mathbb{R}^{m_y m_\theta} \times \text{dom } \psi : \tilde{\theta} = 0\}$$

is UES in the sense of Definition 2.3.

Moreover, $Z : \text{dom } \psi \times \mathbb{R}^{m_\theta m_y} \rightarrow \mathbb{R}_{\geq 0}$, given below, is a Lyapunov function for the closed-loop system:

$$\begin{aligned} Z(t, j, \tilde{\theta}) &:= \tilde{\theta}_f^T (P_r \otimes I_{m_\theta}) \tilde{\theta}_f + \zeta \beta \tilde{\theta}_s^T (P \otimes I_{m_\theta}) \tilde{\theta}_s \\ &\quad + \zeta \tilde{\theta}_s^T (1_{m_1} \otimes I_{m_\theta}) [\alpha I_{m_\theta} + Q(t, j)] (1_{m_1}^T \otimes I_{m_\theta}) \tilde{\theta}_s \end{aligned}$$

where $Q : \text{dom } \psi \rightarrow \mathbb{R}^{m_\theta \times m_\theta}$ is given by

$$\begin{aligned} Q(t, j) &:= \left(\frac{3\tilde{\psi}^2}{2 - 2\exp^{-1}} \right) I_{m_\theta} - Q_\Phi(t, j, +\infty), \\ Q_\Phi(t, j, \infty) &:= \sum_{i=j}^{m_\infty} \int_{t_i}^{t_{i+1}} \exp^{(t+j)-(s+i)} \Phi(s, i) ds \\ &\quad + \frac{1}{2} \sum_{i=j}^{m_\infty} \int_{t_{i+1}}^{t_{i+2}} \exp^{(t+j)-(s+i)} ds \Phi(t_{i+1}, i), \\ \Phi(t, j) &:= \begin{cases} \frac{1}{m_1} \sum_{i=1}^{m_1} \psi_i(t, j) \psi_i(t, j)^T & \text{if } t \in \text{int}(I_j), \\ \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{\gamma \psi_i(t, j) \psi_i(t, j)^T}{1 + \gamma |\psi_i(t, j)|^2} & \text{otherwise,} \end{cases} \end{aligned}$$

$m_\infty \in \mathbb{N} \cup \{+\infty\}$ is the maximum amount of jumps achieved on $\text{dom } \psi \cap [t, +\infty) \times \{j, j+1, \dots\} := [t, t_{j+1}] \times \{j\} \cup [t_{j+1}, t_{j+2}] \times \{j+1\} \cup \dots$, and (α, β, ζ) are positive constants appropriately chosen such that (7) holds and, along the solutions to \mathcal{H}_{dg} , we have

- For each $(t, j) \in \text{dom } \psi$ such that $(t, j+1) \notin \text{dom } \psi$,

$$\dot{Z}(t, j, \tilde{\theta}) \leq -\frac{\gamma_2 \lambda_{\min}(Q_r)}{2} |\tilde{\theta}|^2.$$

- For each $(t, j) \in \text{dom } \psi$ such that $(t, j+1) \in \text{dom } \psi$,

$$Z(t, j+1, \tilde{\theta}^+) - Z(t, j, \tilde{\theta}) \leq -\frac{\delta \lambda_{\min}(Q_r)}{4} |\tilde{\theta}|^2.$$

REMARK 5. Note that the first term in Z provides a strict Lyapunov function for the hybrid dynamics of $\tilde{\theta}_f$ when $\tilde{\theta}_s = 0$. Furthermore, the third term provides a non-strict Lyapunov function for the hybrid dynamics of $\tilde{\theta}_s$, which we strictify by adding the second term in Z . Indeed, the variations of the third term in Z along the solutions to \mathcal{H}_{dg} yield to a negative (quadratic) term in the errors $e_s := \left[\tilde{\theta}_s - (1_{m_1} \otimes I_{m_\theta}) (1_{m_1}^T \otimes I_{m_\theta}) \tilde{\theta}_s \right]$. Then, it is necessary to analyze the behavior of the average state $x_s := (1_{m_1}^T \otimes I_{m_\theta}) \tilde{\theta}_s$ when $e_s = 0$ —cf. [18]. This is the purpose of the second term in Z .

REMARK 6. Note that we can always find $\delta > 0$ such that conditions (17) and (18) hold. Indeed, the matrix Q_r is positive definite, which implies that we can always find $\delta > 0$ sufficiently small such that (17) holds. Moreover, we show in another step that $\text{Ker}(P \mathcal{L}_1 + \mathcal{L}_1^T P) = \text{Span}(1_{m_1})$; hence, $\text{Ker}(P \mathcal{L}_1 + \mathcal{L}_1^T P) \subset \text{Ker}(\mathcal{L}_1^T P \mathcal{L}_1)$. As a result, we can always design $\delta > 0$ sufficiently small such that (18) holds.

3.2 Connection to concurrent-learning-based estimation

Consider the scenario where a single estimator is used, let (ψ_1, y_1) be input-output data received in real time. Assume that the estimator has access to old data forming the sequence $\{(\psi_i, y_i)\}_{i=2}^{m_y}$. In this case, the hybrid update of $\hat{\theta}_1$ given in (10)-(11), while replacing i therein by 1, can viewed as a sum of two types of terms:

- (1) exploration terms, which are $-\gamma_1 \psi_1(t, j) e_1^T$ along flows and $\hat{\theta}_1 - \sigma_1(t, j) \psi_1(t, j) e_1$ along jumps and
- (2) exploitation terms, which are $-\gamma_2 \sum_{k=1}^{m_y} a_{1k} (\hat{\theta}_1 - \hat{\theta}_k)$ along flows and $-\delta \sum_{k=1}^{m_y} a_{1k} (\hat{\theta}_1 - \hat{\theta}_k)$ along jumps.

According to Theorem 3.1, the estimator can estimate θ exponentially fast, even if the real-time input ψ_1 is not hybrid persistently exciting. In contrast, estimating the parameters is impossible without exploiting the old data $\{(\psi_i, y_i)\}_{i=2}^{m_y}$. Furthermore, when the real-time input ψ_1 is hybrid persistently exciting, it is important to design well the graph topology so that adding virtual estimators improve, or at least do not degrade, the convergence properties. All these behaviors are illustrated, via simulations, in Section 5.

4 COOPERATIVE ADAPTIVE OBSERVATION/IDENTIFICATION

Based on the approach described in Section 3, we study the observation-plus-identification problem for hybrid systems of the form

$$\mathcal{H} : \begin{cases} \dot{x} = A_c x + \Psi_c(y, u)\theta & (x, u) \in C \\ x^+ = A_d x + \Psi_d(y, u)\theta & (x, u) \in D \\ y = E x, \end{cases} \quad (19)$$

where $x \in \mathcal{X} \subset \mathbb{R}^{m_x}$ is the state, $y \in \mathbb{R}^{m_y}$, with $m_y \leq m_x$, is the output, $\theta \in \mathbb{R}^{m_\theta}$ is a constant vector of unknown parameters, $E \in \mathbb{R}^{m_y \times m_x}$ is the output matrix, $(C, D) \subset \mathbb{R}^{m_x \times m_x} \times \mathbb{R}^{m_x \times m_x}$ are the flow and the jump sets, respectively, and $\Psi_c, \Psi_d : \mathbb{R}^{m_y} \times \mathbb{R}^{m_u} \rightarrow \mathbb{R}^{m_x \times m_\theta}$ are the continuous- and the discrete-time regressor functions, respectively, satisfying the following assumption.

(SA4) The regressors Ψ_c and Ψ_d are continuous.

Our approach is based on designing a network of reduced-order adaptive observers/identifiers that exchange their parameter estimates according to a communication graph. To do so, we start assuming that the hybrid system \mathcal{H} can be decomposed into m_y subsystems of the form

$$\mathcal{H}_i : \begin{cases} \dot{x}_i = A_{ci} x_i + \Psi_{ci}(y, u)\theta & (x, u) \in C \\ x_i^+ = A_{di} x_i + \Psi_{di}(y, u)\theta & (x, u) \in D \\ y_i = E_i x_i, \end{cases}$$

where, for each $i \in \{1, 2, \dots, m_y\}$, $x_i \in \mathbb{R}^{m_{xi}}$, $y_i \in \mathbb{R}$, and $\Psi_i : \mathbb{R}^{m_y} \times \mathbb{R}^{m_u} \rightarrow \mathbb{R}^{m_{xi} \times m_\theta}$, and $E_i \in \mathbb{R}^{m_{xi}}$. Furthermore, we let the following standing assumption holds.

(SA5) There exist scalars $a_c, a_d \leq 0$ and $\beta, M \geq 0$, matrices $K_{ci}, K_{di} \in \mathbb{R}^{m_{xi}}$, and a positive definite matrix $P_i = P_i^\top \in \mathbb{R}^{m_{xi} \times m_{xi}}$ such that

$$A_{ccli}^\top P_i + P_i A_{ccli} \leq a_c P_i, \quad (20a)$$

$$A_{dcli}^\top P_i A_{dcli} \leq e^{a_d} P_i, \quad (20b)$$

where $A_{ccli} := A_{ci} - K_{ci} E_i$ and $A_{dcli} := A_{di} - K_{di} E_i$, and

$$a_c t + a_d j \leq M - \beta(t + j) \quad \forall (t, j) \in \mathbb{R}_{\geq 0} \times \mathbb{N}_{\geq 0}.^1 \quad (21)$$

REMARK 7. According to (SA5), for each $i \in \{1, 2, \dots, m_y\}$, the pairs (A_{ci}, E_i) and (A_{di}, E_i) are observable and admit a common quadratic Lyapunov function characterized by the matrix P_i . Note that (SA5) can be relaxed under an extra knowledge on $\text{dom } \phi$; see [21]. For example, when ϕ is eventually discrete, eventually continuous, jumps periodically, has a persistent flow, or has a persistent jump.

The goal in this section is to design multiple reduced-order

(SA6) Each observer/identifier \mathcal{H}_{oi} knows the matrices (A_{ci}, A_{di}, E_i) , the maps (Ψ_{ci}, Ψ_{di}) , and the pair (y, u) .

¹One can easily check that for given $a_c, a_d \leq 0$, there always exist $M, \beta \geq 0$ such that (21) holds.

(SA7) The different observers/identifiers exchange their estimates of θ according to a graph \mathcal{G} satisfying (SA2) and (SA3).

(SA8) Knowing the pair (y, u) , we can detect instantaneously when the corresponding solution ϕ to \mathcal{H} jumps.

REMARK 8. Note that the decomposition of \mathcal{H} into $\{\mathcal{H}_i\}_{i=1}^{m_y}$ assumes a particular structure for \mathcal{H} . Extending our result using a more general decomposition based on the canonical form of observability could be a promising perspective to this work. Furthermore, one could envision relaxing (SA8) by assuming the jump times to be approximately known –cf. [3].

4.1 Networked Hybrid Adaptive Observers/Identifiers

As in the classical observer-design paradigm, we assume that each observer/identifier \mathcal{H}_{oi} contains a copy of \mathcal{H}_i only. Hence, it is reasonable to expect that, when the dimension \mathcal{H} is large (note that $m_x = \sum_{i=1}^{m_y} m_{xi}$), the proposed approach should impose rely on shorter computation time as opposed to using a single observer/identifier containing a copy of the entire system \mathcal{H} .

Consider a solution $\phi : \text{dom } \phi \rightarrow \mathbb{R}^{m_x}$ to \mathcal{H} and the corresponding input-output pair (u, y) . For each $i \in \{1, 2, \dots, m_y\}$, the state vector of the i th estimator \mathcal{H}_{oi} is

$$X_{oi} := (t, j, \hat{x}_i, \hat{\theta}_i, \Gamma_{ci}, \Gamma_{di}) \in \mathcal{X}_{oi},$$

where $\mathcal{X}_{oi} := \text{dom } \phi \times \mathbb{R}^{m_{xi} \times m_\theta} \times \mathbb{R}^{m_{xi} \times m_\theta} \times \mathbb{R}^{m_{xi}} \times \mathbb{R}^{m_\theta}$.

Furthermore, to simplify the notation, we introduce the hybrid arcs $\Psi_i, \Gamma_i : \text{dom } \phi \rightarrow \mathbb{R}^{m_{xi} \times m_\theta}$ and $\psi_i : \text{dom } \phi \rightarrow \mathbb{R}^{m_\theta}$ given by

$$\Psi_i(t, j) := \begin{cases} \Psi_{ci}(y(t, j), u(t, j)) & \text{if } t \in \text{int}(I^j) \\ \Psi_{di}(y(t, j), u(t, j)) & \text{otherwise,} \end{cases} \quad (22)$$

$$\Gamma_i(t, j) := \begin{cases} \Gamma_{ci}(t, j) & \text{if } t \in \text{int}(I^j) \\ \Gamma_{di}(t, j) & \text{otherwise,} \end{cases} \quad (23)$$

$$\psi_i(t, j) := \Gamma_i(t, j)^\top E_i^\top, \quad (24)$$

where $I^j := \{t : (t, j) \in \text{dom } \phi\}$.

Then, the hybrid dynamics of each \mathcal{H}_{oi} is given by

$$\mathcal{H}_{oi} : \begin{cases} \begin{cases} i = 1, & j = 0, \\ \dot{\hat{x}}_i = A_{ci} \hat{x}_i + \Psi_i \hat{\theta}_i + (K_{ci} + \gamma_1 \Gamma_i \psi_i) (y_i - E_i \hat{x}_i) \\ \quad \quad \quad - \gamma_2 \sum_{j=1}^{m_y} a_{ij} \Gamma_i (\hat{\theta}_i - \hat{\theta}_j) \\ \dot{\hat{\theta}}_i = \gamma_1 \psi_i (y_i - E_i \hat{x}_i) - \gamma_2 \sum_{j=1}^{m_y} a_{ij} (\hat{\theta}_i - \hat{\theta}_j) \\ \dot{\Gamma}_{ci} = A_{ccli} \Gamma_{ci} + \Psi_i & \dot{\Gamma}_{di} = 0, & X_{oi} \in C_{oi} \end{cases} \\ \begin{cases} t^+ = t, & j^+ = j + 1, \\ \hat{x}_i^+ = A_{di} \hat{x}_i + \Psi_i \hat{\theta}_i + (K_{di} + \sigma_i \Gamma_i^+ \psi_i) (y_i - E_i \hat{x}_i) \\ \quad \quad \quad - \delta \sum_{j=1}^{m_y} a_{ij} \Gamma_i^+ (\hat{\theta}_i - \hat{\theta}_j) \\ \hat{\theta}_i^+ = \hat{\theta}_i + \sigma_i \psi_i (y_i - E_i \hat{x}_i) - \delta \sum_{j=1}^{m_y} a_{ij} (\hat{\theta}_i - \hat{\theta}_j) \\ \Gamma_{ci}^+ = \Gamma_{ci}, & \Gamma_{di}^+ = A_{dcli} \Gamma_{di} + \Psi_i & X_{oi} \in D_{oi}, \end{cases} \end{cases} \quad (25a)$$

where

$$C_{oi} := \{X_{oi} \in \mathcal{X}_{oi} : (t, j) \in \text{dom } \phi \text{ and } (t, j+1) \notin \text{dom } \phi\},$$

$$D_{oi} := \{X_{oi} \in \mathcal{X}_{oi} : (t, j) \in \text{dom } \phi \text{ and } (t, j+1) \in \text{dom } \phi\},$$

the matrices K_{ci} and K_{di} are designed such that (SA5) hold, γ_1, γ_2 and $\delta > 0$ are constant adaptation rates, and $\sigma_i : \text{dom } \phi \rightarrow \mathbb{R}$ is given by $\sigma_i(t, j) := \frac{\gamma}{1 + \gamma |\psi_i(t, j)|^2}$, where $\gamma > 0$ is a constant design parameter.

4.2 Main Result II

Before presenting our main result, we introduce the error coordinates $e_i := x_i - \hat{x}_i$ and $\tilde{\theta}_i := \hat{\theta}_i - \theta_i$ for all $i \in \{1, 2, \dots, m_y\}$. Furthermore, for compact notation, we let $e := (e_1, e_2, \dots, e_{m_y})$, $\tilde{\theta} := (\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_{m_y})$, $X_o := (X_{o1}, X_{o2}, \dots, X_{om_y})$, $\Gamma_c := (\Gamma_{c1}, \Gamma_{c2}, \dots, \Gamma_{cm_y})$, $\Gamma_d := (\Gamma_{d1}, \Gamma_{d2}, \dots, \Gamma_{dm_y})$, and $X_o := (X_{o1} \times X_{o2} \times \dots \times X_{om_y})$. Furthermore, we introduce the set $S_{e\theta} := \{(X_o, x) \in X_o \times X : e = \tilde{\theta} = 0\}$ to render UES and, for a given $\bar{\Gamma}_o > 0$, we introduce the set of initial conditions

$$X_{in} := \{(X_o, x) \in X_o \times X : \max\{|\Gamma_c|, |\Gamma_d|\} \leq \bar{\Phi}_o\}.$$

Finally, we introduce what we refer to as distributed hybrid persistence of excitation (DHPE) condition along solutions.

ASSUMPTION 5 (DHPE ALONG SOLUTIONS). *There exist $K_1, \mu_1 > 0$ such that, for each $(X_o, x) \in X_{in}$ and for each hybrid time domain $\bigcup_{j=j_o}^J ([t_j, t_{j+1}] \times \{j\}) \subset \text{dom } u = \text{dom } y$, with $(t_{j+1} - t_{j_o}) + (J - j_o) \geq K_1$, we have*

$$\sum_{i \in \mathcal{V}_1} \sum_{j=j_o}^J \left[\int_{t_j}^{t_{j+1}} \psi_i(\tau, j) \psi_i(\tau, j)^\top d\tau + \frac{\gamma \psi_i(t_{j+1}, j) \psi_i(t_{j+1}, j)^\top}{2 + 2\gamma |\psi_i(t_{j+1}, j)|^2} \right] \geq m_1 \mu_1 I_{m_\theta},$$

where ψ_i is given in (24).

REMARK 9. *Note that each ψ_i in (24) is the output to the hybrid linear filter whose input is the arc Ψ_i . That is, if we consider the linear filter*

$$\mathcal{H}_{\psi_i} : \begin{cases} \begin{cases} i = 1, & j = 0, & \dot{\Gamma}_{di} = 0, \\ \dot{\Gamma}_{ci} = A_{ccli} \Gamma_{ci} + \Psi_i, & (t, j, \Gamma_{ci}, \Gamma_{di}) \in C_{\psi_i} \end{cases} \\ \begin{cases} t^+ = t, & j^+ = j + 1, & \Gamma_{ci}^+ = \Gamma_{ci}, \\ \Gamma_{di}^+ = A_{dcli} \Gamma_{di} + \Psi_i, & (t, j, \Gamma_{ci}, \Gamma_{di}) \in D_{\psi_i}. \end{cases} \end{cases}$$

the hybrid arc ψ_i can be seen as the output of \mathcal{H}_{ψ_i} given by

$$\psi_i(t, j) = \begin{cases} \Gamma_{ci}(t, j)^\top E^\top & \text{if } t \in \text{int}(I^j) \\ \Gamma_{di}(t, j)^\top E^\top & \text{otherwise,} \end{cases}$$

where

$$C_{\psi_i} := \{(t, j, \Gamma_{di}, \Gamma_{ci}) : (t, j) \in \text{dom } \phi \text{ and } (t, j+1) \notin \text{dom } \phi\}, \\ D_{\psi_i} := \{(t, j, \Gamma_{di}, \Gamma_{ci}) : (t, j) \in \text{dom } \phi \text{ and } (t, j+1) \in \text{dom } \phi\}.$$

Hence, it is interesting to investigate conditions on the inputs $\{\Psi_i\}_{i=1}^{m_y}$ and the data of $\{\mathcal{H}_{\psi_i}\}_{i=1}^{m_y}$ such that $\{\psi_i\}_{i=1}^{m_y}$ verifies Assumption 5.

THEOREM 4.1. *Given a solution $\phi : \text{dom } \phi \rightarrow \mathbb{R}^{m_x}$ to \mathcal{H} in (19), such that the corresponding input-output pair (u, y) is bounded, and the network of observers/identifiers $\mathcal{H}_o := \{\mathcal{H}_{oi}\}_{i=1}^{m_y}$. Then, the closed set $S_{e\theta}$ is UES on X_{in} for the hybrid system $(\mathcal{H}, \mathcal{H}_o)$ provided that Assumption 5 holds.*

Sketch of proof of Theorem 4.1. We start introducing the coordinate vector $X_{e\theta i} := (t, j, \Gamma_{ci}, \Gamma_{di}, e_i, \tilde{\theta}_i) \in X_{oi}$, whose dynamics is

given by

$$\mathcal{H}_{\eta\theta i} : \begin{cases} \begin{cases} i = 1, & j = 0, & \dot{\Gamma}_{di} = 0, & \dot{\Gamma}_{ci} = A_{ccli} \Gamma_{ci} + \Psi_i \\ \dot{e}_i = A_{ccli} e_i - \Psi_i \tilde{\theta}_i - \gamma_1 \Gamma_i \psi_i E_i e_i + \gamma_2 \sum_{j=1}^{m_y} a_{ij} \Gamma_i (\tilde{\theta}_i - \tilde{\theta}_j) \\ \dot{\tilde{\theta}}_i = \gamma_1 \psi_i E_i e_i - \gamma_2 \sum_{j=1}^{m_y} a_{ij} (\tilde{\theta}_i - \tilde{\theta}_j) & X_{e\theta i} \in C_{oi} \end{cases} \\ \begin{cases} t^+ = t, & j^+ = j + 1, & \Gamma_{ci}^+ = \Gamma_{ci}, & \Gamma_{di}^+ = A_{dcli} \Gamma_{di} + \Psi_i \\ e_i^+ = A_{dcli} e_i - \Psi_i \tilde{\theta}_i - \sigma_i \Gamma_i^+ \psi_i E_i e_i + \delta \sum_{j=1}^{m_y} a_{ij} \Gamma_i^+ (\tilde{\theta}_i - \tilde{\theta}_j) \\ \tilde{\theta}_i^+ = \tilde{\theta}_i + \sigma_i \psi_i E_i e_i - \delta \sum_{j=1}^{m_y} a_{ij} (\tilde{\theta}_i - \tilde{\theta}_j) & X_{e\theta i} \in D_{oi}. \end{cases} \end{cases}$$

Next, we introduce the change of variable $\eta_i := e_i + \Gamma_i(t, j) \tilde{\theta}_i$, where $\Gamma_i : \text{dom } \phi \rightarrow \mathbb{R}^{m_{x_i} \times m_\theta}$ is defined in (23), and we let the new coordinate vector $X_{\eta\theta i} := (t, j, \Gamma_{di}, \Gamma_{ci}, \eta_i, \tilde{\theta}_i) \in X_{oi}$, whose dynamics is governed by

$$\mathcal{H}_{\eta\theta i} : \begin{cases} \begin{cases} i = 1, & j = 0, & \dot{\eta}_i = A_{ccli} \eta_i \\ \dot{\tilde{\theta}}_i = -\gamma_1 \psi_i^\top (\psi_i \tilde{\theta}_i - E_i \eta_i) - \gamma_2 \sum_{j=1}^{m_y} a_{ij} (\tilde{\theta}_i - \tilde{\theta}_j) \\ \dot{\Gamma}_{ci} = A_{ccli} \Gamma_{ci} + \Psi_i, & \dot{\Gamma}_{di} = 0, & X_{\eta\theta i} \in C_{oi} \end{cases} \\ \begin{cases} t^+ = t, & j^+ = j + 1, & \eta_i^+ = A_{dcli} \eta_i \\ \tilde{\theta}_i^+ = \tilde{\theta}_i - \sigma_i \psi_i^\top (\psi_i \tilde{\theta}_i - E_i \eta_i) - \delta \sum_{j=1}^{m_y} a_{ij} (\tilde{\theta}_i - \tilde{\theta}_j) \\ \Gamma_{ci}^+ = \Gamma_{ci} & \Gamma_{di}^+ = A_{dcli} \Gamma_{di} + \Psi_i, & X_{\eta\theta i} \in D_{oi}. \end{cases} \end{cases}$$

Under (SA5), we use [9, Proposition 3.29] combined with the Lyapunov function $V(\eta_i) = \eta_i^\top P_i \eta_i$, to conclude that η_i converges to zero exponentially for all $i \in \{1, 2, \dots, m_y\}$.

As a result, it is enough to analyze the parameter-estimation system whose state variable is $X_{\theta i} := (t, j, \Gamma_{di}, \Gamma_{ci}, \tilde{\theta}_i) \in X_{\theta i}$, where $X_{\theta i} := \text{dom } \phi \times \mathbb{R}^{m_{x_i} \times m_\theta} \times \mathbb{R}^{m_{x_i} \times m_\theta} \times \mathbb{R}^{m_\theta}$, and whose dynamics is given by

$$\mathcal{H}_{\theta i} : \begin{cases} \begin{cases} i = 1, & j = 0, \\ \dot{\tilde{\theta}}_i = -\gamma_1 \psi_i \psi_i^\top \tilde{\theta}_i - \gamma_2 \sum_{j=1}^{m_y} a_{ij} (\tilde{\theta}_i - \tilde{\theta}_j) \\ \dot{\Gamma}_{ci} = A_{ccli} \Gamma_{ci} + \Psi_i, & \dot{\Gamma}_{di} = 0 & X_{\theta i} \in C_{\theta i} \end{cases} \\ \begin{cases} t^+ = t, & j^+ = j + 1, \\ \tilde{\theta}_i^+ = \left[I_{m_\theta} - \frac{\gamma \psi_i \psi_i^\top}{1 + \gamma \psi_i^\top \psi_i} \right] \tilde{\theta}_i - \delta \sum_{j=1}^{m_y} a_{ij} (\tilde{\theta}_i - \tilde{\theta}_j) \\ \Gamma_{ci}^+ = \Gamma_{ci}, & \Gamma_{di}^+ = A_{dcli} \Gamma_{di} + \Psi_i, & X_{\theta i} \in D_{\theta i}, \end{cases} \end{cases}$$

where

$$C_{\theta i} := \{X_{\theta i} \in X_{\theta i} : (t, j) \in \text{dom } \phi \text{ and } (t, j+1) \notin \text{dom } \phi\},$$

$$D_{\theta i} := \{X_{\theta i} \in X_{\theta i} : (t, j) \in \text{dom } \phi \text{ and } (t, j+1) \in \text{dom } \phi\}.$$

Note that, for each $i \in \{1, 2, \dots, m_y\}$, the matrices A_{ccli} and A_{dcli} are exponentially stable in the sense of continuous-time and discrete-time systems, respectively, (SA4) holds, and the pair (u, y) are bounded. As a result, $\max\{|\Gamma_c(0, 0)|, |\Gamma_d(0, 0)|\} \leq \bar{\Gamma}_o$, then there exists $\bar{\Gamma} > 0$ such that $\sup\{|\Gamma_c(t, j)|_\infty, |\Gamma_d(t, j)|_\infty\} \leq \bar{\Phi}$. Thus, $\psi := [\psi_1, \psi_2, \dots, \psi_{m_y}]$ satisfies (SA1). Now, using Theorem 3.1, under Assumption 5, we conclude that the parameter estimation error $\tilde{\theta}$ converges to zero exponentially. Finally, once both η and $\tilde{\theta}$ converge to zero exponentially, we conclude that the same property holds for the observation error e .

4.3 Observation/Identification Via Adaptive Concurrent Learning

Following the discussion in Section 3.2, we use the method described in Section 4.1 to reinforce the design of a single adaptive observer/identifier using old data. That is, we consider a hybrid system in the form of \mathcal{H} in (19) with $m_y = 1$ (we are not going to decompose \mathcal{H} into subsystems). Let the pair (u_1, y_1) represent the input-output data available in real time and let $\{(u_i, y_i)\}_{i=2}^N$, for some $N > 0$, be a sequence of old data. For simplicity, we assume that all the input-output pairs have the same hybrid time domain.

ASSUMPTION 6. $\text{dom } y_1 = \text{dom } u_1 = \dots = \text{dom } y_N = \text{dom } u_N$.

At this point, we introduce the sequence of hybrid systems $\{\mathcal{H}_i\}_{i=1}^N$, where each \mathcal{H}_i corresponds to \mathcal{H} subject the input u_i and thus generating the output y_i . Under (SA8), each \mathcal{H}_i can be expressed as

$$\mathcal{H}_{\psi_i} : \begin{cases} \begin{cases} \dot{t} = 1, & j = 0, \\ \dot{x}_i = A_c x_i + \Psi_i(t, j)\theta & (t, j, x_i) \in C_i \end{cases} \\ \begin{cases} t^+ = t, & j^+ = j + 1, \\ x_i^+ = A_d x_i + \Psi_i(t, j)\theta & (t, j, x_i) \in D_i \\ y_i = E x_i, \end{cases} \end{cases}$$

where $\Psi_i : \text{dom } u_i \rightarrow \mathbb{R}^{m_{x_i} \times m_\theta}$, in this case, is defined as

$$\Psi_i(t, j) := \begin{cases} \Psi_c(y_i(t, j), u_i(t, j)) & \text{if } t \in \text{int}(I^j) \\ \Psi_d(y_i(t, j), u_i(t, j)) & \text{otherwise,} \end{cases}$$

and

$$C_i := \{(t, j) \in \text{dom } \Psi_i : (t, j + 1) \notin \text{dom } \Psi_i\} \times \mathbb{R}^{m_x},$$

$$D_i := \{(t, j) \in \text{dom } \Psi_i : (t, j + 1) \in \text{dom } \Psi_i\} \times \mathbb{R}^{m_x}.$$

The proposed concurrent-learning-based observer/identifier uses a network $\{\mathcal{H}_{oi}\}_{i=1}^N$, where each \mathcal{H}_{oi} is designed as in (25) while replacing (A_{ci}, A_{di}) therein by (A_c, A_d) . Theorem 4.1 applies in this case, so we conclude that each \mathcal{H}_{oi} tracks the state x_i of \mathcal{H}_i and estimates the parameter θ . More precisely, we have the following.

COROLLARY 4.2. *Given the hybrid system \mathcal{H} in (19) with $m_y = 1$. Assume that the real-time input-output pair (u_1, y_1) and the sequence of old input-output data $\{(u_i, y_i)\}_{i=2}^N$ are bounded and satisfy Assumption 6. Consider the concurrent-learning-based observer/identifier $\mathcal{H}_0 := \{\mathcal{H}_{oi}\}_{i=1}^N$, where each \mathcal{H}_{oi} is designed as in (25) while replacing (A_{ci}, A_{di}) therein by (A_c, A_d) . Then, the closed set \mathcal{S}_{e_θ} is UES on \mathcal{X}_{in} for the hybrid system $(\mathcal{H}, \mathcal{H}_0)$ provided that Assumption 5 holds.*

REMARK 10. *According to Corollary 4.2, it is possible that the real-time data (u_1, y_1) fails to generate a signal ψ_1 that is hybrid persistently exciting. Yet, the concurrent-learning-based observer/identifier \mathcal{H}_0 can estimate the vector of parameters θ exponentially. In contrast, estimating the parameters is impossible using only the real-time data (y_1, u_1) .*

REMARK 11. *Assumption 6 is hard to satisfy in the general context of hybrid systems. However, we can always unify the domain of our data by creating fictitious jumps (static jumps) for all the data whenever one of them jumps. This procedure will affect the matrix A_d in each of the systems $\{\mathcal{H}_i\}_{i=1}^N$; thus, the analysis will require other machinery that is omitted in this version due to space limitation.*

5 ILLUSTRATIVE NUMERICAL EXAMPLES

5.1 Comparison between the continuous-, the discrete-time, and the hybrid gradient algorithms

Consider the linear input-output relationship

$$y(t, j) = \psi(t, j)^\top \theta,$$

where the hybrid regressor $\psi : \text{dom } \psi \rightarrow \mathbb{R}^2$ is given by

$$\psi_1(t, j) := \begin{cases} [\cos(2t) \ 0]^\top & \forall t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ [0.3 \ 0.9]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N}. \end{cases}$$

First, when viewing the regressor ψ as a discrete-time function defined on $\{0, 1, \dots\}$, we can see that the regressor

$$\psi(j)\psi(j)^\top := \begin{bmatrix} 0.09 & 0.27 \\ 0.27 & 0.81 \end{bmatrix} \quad j \in \{0, 1, \dots\}$$

is constant and not full rank. Hence, the DT-PE condition in Assumption 2 is not satisfied. Similarly, when viewing the regressor ψ as a continuous-time function defined on $\mathbb{R}_{\geq 0}$, we can see that

$$\psi(t)\psi(t)^\top := \begin{cases} \begin{bmatrix} \cos(2t)^2 & 0 \\ 0 & 0 \end{bmatrix} & t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ \begin{bmatrix} 0.09 & 0.27 \\ 0.27 & 0.81 \end{bmatrix} & \text{otherwise.} \end{cases}$$

Hence, for each $t_0 > 0$ and $T > 0$, we have

$$\int_{t_0}^{t_0+T} \psi(s)\psi(s)^\top ds = \begin{bmatrix} \int_{t_0}^{t_0+T} \cos(2s)^2 ds & 0 \\ 0 & 0 \end{bmatrix},$$

so the CT-PE condition in Assumption 1 is not satisfied. Yet, the HPE condition in Assumption 3 holds with $K = 2\pi + 1$ and $\mu = 0.21$. Figure 1 provides a comparison between the classical continuous-, discrete-time, and the hybrid gradient-descent algorithms in terms of convergence of the parameter estimation errors.

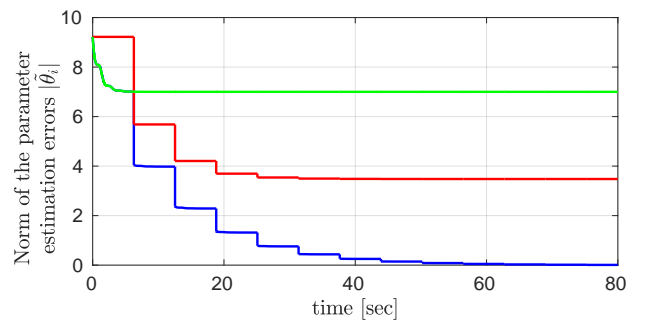


Figure 1: Evolution of $\|\tilde{\theta}\|$ using the continuous- (in green), the discrete-time (in red), and the hybrid (in blue) gradient-descent algorithms.

The initial condition for the simulation is $\tilde{\theta}(0) := [6 \ 7]^\top$. It can be seen that while the discrete and continuous-time algorithms fail to make the errors $\tilde{\theta}$ converge to zero, the hybrid gradient-descent algorithm succeeds.

5.2 The distributed gradient algorithm guarantees convergence when the centralized one fails

Consider the network of three linear estimators in the form of (8). For the three estimators, the corresponding hybrid regressors $\{\psi_1, \psi_2, \psi_3\}$ are given by

$$\begin{aligned} \psi_1(t, j) &:= \begin{cases} [\sin(t) \ 0]^\top & \forall t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ [0 \ 0]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N}, \end{cases} \\ \psi_2(t, j) &:= \begin{cases} [\cos(t) \ 0]^\top & \forall t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ [0 \ 0]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N}, \end{cases} \\ \psi_3(t, j) &:= \begin{cases} [0 \ 0]^\top & \forall t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ [0.5 \ 1]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N}. \end{cases} \end{aligned}$$

Using a similar reasoning as the one used in the previous example, one can easily check that none of the regressors $\{\psi_1, \psi_2, \psi_3\}$ is HPE according to Assumption 3. Now, we apply the distributed scenario presented in this paper, where the different estimators exchange their estimates according to a communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V} = \{1, 2, 3\}$ and $\mathcal{E} = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$, which corresponds to a fully connected graph with an adjacency matrix $\mathcal{A} := [a_{ik}] \in \mathbb{R}^{3 \times 3}$ defined by $a_{ik} = 1$ if $i \neq k$ and $a_{ii} = 0$. For this distributed algorithm, the DHPE condition in Assumption 4 is satisfied with $K_1 = 2\pi + 1$, $m_1 = 3$ and $\mu_1 = 0.11$.

Figures 2 and 3 provide a comparison between the centralized and the distributed hybrid gradient algorithms.

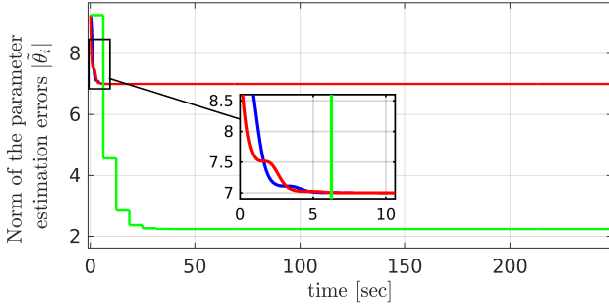


Figure 2: Evolution of $|\tilde{\theta}_1|$ (in blue), $|\tilde{\theta}_2|$ (in green), and $|\tilde{\theta}_3|$ (in red) using the centralized hybrid gradient algorithm.

The initial estimation errors are $\tilde{\theta}_1(0) = \tilde{\theta}_2(0) = \tilde{\theta}_3(0) := [6 \ 7]^\top$. It can be seen that while the centralized hybrid gradient-descent algorithm does not allow the convergence of $\tilde{\theta}_i$, for each $i = \{1, 2, 3\}$, to zero, the proposed distributed hybrid gradient descent algorithm ensures this property. This can be explained by the fact that none of the regressors ψ_1, ψ_2 and ψ_3 is hybrid persistently exciting, however, when considering the distributed scenario, where the three estimators exchange their estimates according to the communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ defined above, the distributed hybrid persistence of excitation condition is satisfied.

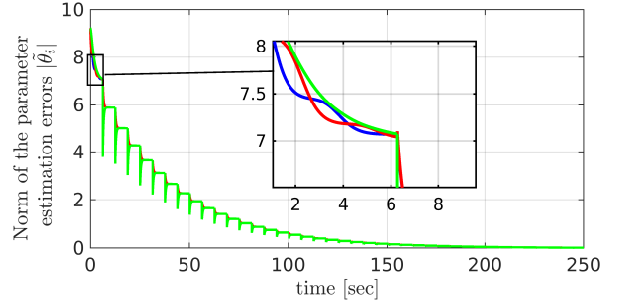


Figure 3: Evolution of $|\tilde{\theta}_1|$ (in blue), $|\tilde{\theta}_2|$ (in green) and $|\tilde{\theta}_3|$ (in red) using the distributed hybrid gradient algorithm.

5.3 The distributed hybrid gradient algorithm improves the convergence speed of the centralized one

Consider a network composed of three linear estimators in the form (8). The corresponding hybrid regressors are given by

$$\begin{aligned} \psi_1(t, j) &:= \begin{cases} \begin{bmatrix} \sin(t) \\ 0 \end{bmatrix} & \forall t \in (2j\pi, 2(j+1)\pi) \cap [0, 300], \\ & j \in \mathbb{N} \\ \begin{bmatrix} 0.05 \sin(t) \\ 0 \end{bmatrix} & \forall t \in (2j\pi, 2(j+1)\pi) \cap (300, +\infty), \\ & j \in \mathbb{N} \\ [0 \ 0]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N} \cap [0, 50] \\ [1 \ -1]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N} \cap (50, +\infty) \end{cases} \\ \psi_2(t, j) &:= \begin{cases} \begin{bmatrix} 0.05 \cos(t) \\ 0 \end{bmatrix} & \forall t \in (2j\pi, 2(j+1)\pi) \cap [0, 300], \\ & j \in \mathbb{N} \\ \begin{bmatrix} \cos(t) \\ 0 \end{bmatrix} & \forall t \in (2j\pi, 2(j+1)\pi) \cap (300, +\infty), \\ & j \in \mathbb{N} \\ [1 \ -1]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N}. \end{cases} \\ \psi_3(t, j) &:= \begin{cases} \begin{bmatrix} 0.05 \sin(t) & 0 \end{bmatrix}^\top & \forall t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ [1 \ -1]^\top & \forall t = 2j\pi, \quad j \in \mathbb{N}. \end{cases} \end{aligned}$$

We can easily check that the three regressors $\{\psi_1, \psi_2, \psi_3\}$ satisfy the HPE condition in Assumption 3. Now, we consider a distributed scenario, where the different estimators exchange their estimates according to the communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, 2, 3\}$ and $\mathcal{E} = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$, which corresponds to a fully connected graph with $a_{ik} = 1$ if $i \neq k$ and $a_{ii} = 0$. For such distributed scenario, one can also check that the DHPE condition in Assumption 4 is satisfied.

The initial estimation errors are $\tilde{\theta}_1(0) = \tilde{\theta}_2(0) = \tilde{\theta}_3(0) := [6 \ 7]^\top$. Figures 4 and 5, below, provide a comparison between the centralized hybrid gradient-descent algorithm and its distributed version in terms of the convergence speed. It can be seen that while the centralized hybrid gradient descent algorithm allows the convergence of $\tilde{\theta}_i$, $i = \{1, 2, 3\}$ to zero, in about 450 seconds, the distributed hybrid gradient-descent algorithm ensures a faster convergence, in less than 200 seconds, and which can be explained by the fact that in the distributed scenario, the three estimators collaborate, through the exchange of their estimates according to the communication graph \mathcal{G} .

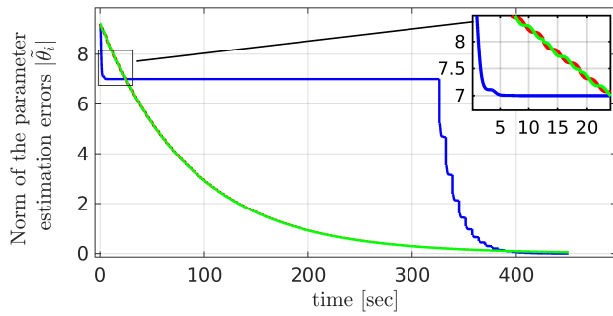


Figure 4: Evolution of $|\hat{\theta}_1|$ (in blue), $|\hat{\theta}_2|$ (in green) and $|\hat{\theta}_3|$ (in red) using the centralized hybrid gradient algorithm.

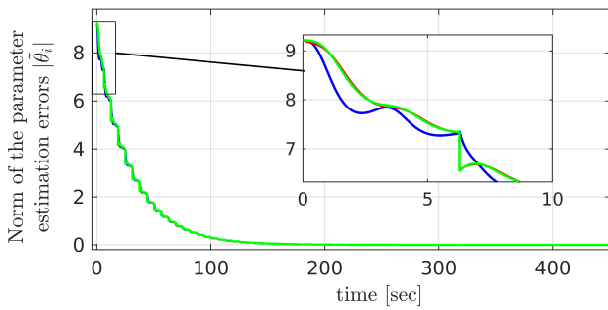


Figure 5: Evolution of $|\hat{\theta}_1|$ (in blue), $|\hat{\theta}_2|$ (in green) and $|\hat{\theta}_3|$ (in red) using the distributed hybrid gradient algorithm.

6 CONCLUSION

We proposed an estimation method involving a network of gradient-descent-based estimators, each of which exploits only a portion of the input-output data, which are viewed as hybrid signals. The different estimators exchange their estimates according to a weakly-connected directed graph. The alternation of flows and jumps combined with the distributed character of the proposed algorithm allow us to introduce a hybrid form of the well-known *persistence of excitation* condition. Under such a condition, we establish exponential convergence of the estimation errors. The proposed approach generalizes the existing ones, yields to relaxed sufficient conditions for (uniform-exponential) parameter estimation. Furthermore, we applied it to address the observation/identification problem for a class of hybrid systems with unknown parameters using a distributed network of adaptive observers/identifiers.

REFERENCES

- [1] K. J. Åström and Bohn. 1965. Numerical identification of linear dynamic systems from normal operating records. In Proc. of the 2nd IFAC Symp. on Theory of Self-adaptive Control Systems, P. H. Hammond (Ed.). Nat. Phys. Lab., Teddington, England, 96–111.
- [2] J. A. Bagnell, D. Bradley, D. Silver, B. Sofman, and A. Stentz. 2010. Learning for autonomous navigation. *IEEE Robotics & Automation Magazine* 17, 2 (2010), 74–84.
- [3] P. Bernard and R. Sanfelice. 2019. Robust Observer Design for Hybrid Dynamical Systems with Linear Maps and Approximately Known Jump Times. (Oct. 2019). <https://hal.archives-ouvertes.fr/hal-02187411> working paper or preprint.
- [4] A. Brown and A. Deaton. 1972. Surveys in applied economics: models of consumer behaviour. *The Economic Journal* 82, 328 (1972), 1145–1236.
- [5] W. Chen, C. Wen, S. Hua, and C. Sun. 2014. Distributed Cooperative Adaptive Identification and Control for a Group of Continuous-Time Systems With a Cooperative PE Condition via Consensus. *IEEE Trans. Automat. Control* 59, 1 (2014), 91–106.
- [6] X. Chen, B. Xudong, M-A. Belabbas, and T. Basar. 2017. Controllability of Formations Over Directed Time-Varying Graphs. *IEEE Transactions on Control of Network Systems* 4, 3 (2017), 407–416. <https://doi.org/10.1109/TCNS.2015.2504034>
- [7] G. Chowdhary, T. Yucelen, M. Mühlegg, and E. N. Johnson. 2013. Concurrent learning adaptive control of linear systems with exponentially convergent bounds. *International Journal of Adaptive Control and Signal Processing* 27, 4 (2013), 280–301.
- [8] M. Ghil, S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson. 1981. Applications of estimation theory to numerical weather prediction. In *Dynamic meteorology: Data assimilation methods*. Springer, New York, Heidelberg, Berlin, 139–224.
- [9] R. Goebel, R. G. Sanfelice, and A. R. Teel. 2012. *Hybrid Dynamical Systems: Modeling, stability, and robustness*. Princeton University Press, Princeton, USA.
- [10] A. Guyader and Q. Zhang. 2003. Adaptive observer for discrete time linear time varying systems. *IFAC Proceedings Volumes* 36, 16 (2003), 1705–1710.
- [11] P. Ioannou and J. Sun. 1996. *Robust adaptive control*. Prentice Hall, New Jersey, USA.
- [12] M. U. Javed, J. I. Poveda, and X. Chen. 2021. Excitation Conditions for Uniform Exponential Stability of the Cooperative Gradient Algorithm over Weakly Connected Digraphs. *IEEE Control Systems Letters* 6 (2021), 67–72. <https://doi.org/10.1109/LCSYS.2021.3049153>
- [13] T. Kim, C. Lee, and H. Shim. 2020. Completely Decentralized Design of Distributed Observer for Linear Systems. *IEEE Trans. Automat. Control* 65, 11 (2020), 4664–4678.
- [14] A. C. Lorenc. 1986. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society* 112, 474 (1986), 1177–1194.
- [15] K. S. Narendra and A. M. Annaswamy. 1987. Persistent excitation in adaptive systems. *Int. J. of Contr.* 45, 1 (1987), 127–160.
- [16] K. S. Narendra and A. M. Annaswamy. 2012. *Stable Adaptive Systems*. Prentice Hall, Englewood Cliffs, NJ, USA.
- [17] R. Olfati-Saber, J. A. Fax, and R. M. Murray. 2007. Consensus and cooperation in networked multi-agent systems. *Proc. IEEE* 95, 1 (2007), 215–233.
- [18] E. Panteley and A. Loria. 2017. Synchronization and Dynamic Consensus of Heterogeneous Networked Systems. *IEEE Trans. on Automatic Control* 62, 8 (2017), 3758–3773.
- [19] I. Papusha, E. Lavretsky, and R. M. Murray. 2014. Collaborative system identification via parameter consensus. In *2014 American Control Conference*. Portland, OR, USA, 13–19.
- [20] J. I. Poveda, M. Benosman, and K. G. Vamvoudakis. 2021. Data-enabled extremum seeking: a cooperative concurrent learning-based approach. *International Journal of Adaptive Control and Signal Processing* 35, 7 (2021), 1256–1284.
- [21] C. Prieur, A. R. Teel, and L. Zaccarian. 2014. Relaxed persistent flow/jump conditions for uniform global asymptotic stability. *IEEE Trans. Automat. Control* 59, 10 (2014), 2766–2771.
- [22] Z. Qu. 2009. *Cooperative control of dynamical systems: applications to autonomous vehicles*. Springer Verlag, London, UK.
- [23] R. G. Sanfelice. 2021. *Hybrid Feedback Control*. Princeton University Press, Princeton, NJ, USA.
- [24] A. Saoud, M. Maghenem, and R. G. Sanfelice. 2021. A Hybrid Gradient Algorithm for Linear Regression with Hybrid Signals. In *2021 American Control Conference (ACC)*. Austin, TX, USA, 4997–5002.
- [25] S. Sastry and S. Bodson. 2011. *Adaptive control: stability, convergence and robustness*. Courier Dover Publications, USA, 31–32 pages.
- [26] M. S. Stanković, S. S. Stanković, and D. M. Stipanović. 2015. Consensus-based decentralized real-time identification of large-scale systems. *Automatica* 60 (2015), 219–226.
- [27] G. Tao. 2003. *Adaptive Control Design and Analysis*. Vol. 37. John Wiley & Sons, Hooker, NJ, USA.
- [28] D. Wooden, M. Malchano, K. Blankespoor, A. Howardy, A. A. Rizzi, and M. Raibert. 2010. Autonomous navigation for BigDog. In *IEEE international conference on robotics and automation*. Anchorage, AL, USA, 4736–4741.
- [29] Y. Wu, A. Isidori, and R. Lu. 2021. On the Design of Distributed Observers for Nonlinear Systems. *IEEE Trans. Automat. Control* (2021), 1–1. <https://doi.org/10.1109/TAC.2021.3103864>
- [30] Q. Zhang. 2002. Adaptive observer for multiple-input-multiple-output (MIMO) linear time-varying systems. *IEEE Trans. Automat. Control* 47, 3 (2002), 525–529.