A Hybrid Gradient Algorithm for Linear Regression with Hybrid Signals

Adnane Saoud, Mohamed Maghenem, and Ricardo G. Sanfelice

***Abstract*—Given a linear input/output relationship involving unknown parameters, we propose a hybrid gradient descent algorithm to estimate the unknown parameters when the inputs and the outputs are hybrid signals. These signals are allowed to change continuously during ordinary time – or flow – and to change discretely – or jump – at isolated time instances. To estimate the unknown parameters, we develop a gradient descent algorithm that updates the estimates continuously during flows and instantaneously at jumps. The proposed hybrid gradient algorithm generalizes the existing gradient descent algorithms in the continuous-time and the discrete-time settings. Under a relaxed (hybrid) version of the well-known persistence of excitation condition, the proposed hybrid gradient descent algorithm estimates the parameters exponentially fast. An illustrative example is presented, showing the capabilities of our approach while classical algorithms fails to ensure the convergence.**

I. INTRODUCTION

Estimating the parameters of a system is critical in many applications [1]. One of the most popular related problems is linear regression [2], where the relation between the input and the output is linear. For such models, the estimation problem is generally based on the gradient descent algorithm [3], [2]. This algorithm exploits the structure of the system and the available input-output data to online update the estimate of the parameters. An optimality criterion is used to deduce the dynamics of the estimate and to analyze convergence rigorously using Lyapunov techniques [4]. Providing an estimate of the convergence-rate of the gradient algorithm translates into showing exponential stability of the origin for a linear time-varying system, whose state is the estimation error. In the continuous-time setting [5], [6], [7], [8], [9], it is well established that a persistency of excitation condition is necessary and sufficient for uniform exponential stability of the error system. A lower bound on the convergence rate is provided in [6], [5], and [8]. Note that all the aforementioned approaches translate naturally to the discrete-time case [2].

In this paper, the input and the output signals of the linear regression model are hybrid signals; namely, they are allowed to exhibit both continuous and discrete evolution. Note that the classical continuous-time gradient algorithm exploits only the time intervals on which the input signal is continuous. Similarly, the classical discrete-time gradient algorithm exploits only the time instants where the input signal is discontinuous. However, combining both algorithms into one (hybrid) algorithm offers the potential of gathering the benefit of both algorithms. Motivated by this fact, we propose a hybrid gradient algorithm, where the input-output signals are viewed as hybrid arcs. As a result, after passing the error coordinates, we show that the error system is uniformly exponentially stable provided that an appropriate (hybrid) persistence of excitation condition holds. The obtained (hybrid) persistence of excitation condition relaxes the existing (continuous and discrete) ones, where only one behavior of the input-output signal is considered. Interestingly, we are able to give an explicit bound on the convergence rate. We also observe that not all of the mentioned continuous-time and discrete-time approaches can be extended to analyze the proposed hybrid gradient algorithm. In this paper, we are inspired by the approach proposed in [8]. Moreover, we point out the difficulties encountered when trying to extend the results in [6], [5], [9] to the hybrid case.

The remainder of this paper is organized as follows. Preliminaries are in Section II. A general context is in Section III. The motivation is in Section IV. The main result is in Section V. The proof of the main result is in Section VI. A discussion is in Section VII. Due to space constraints, some proofs are omitted and will be published elsewhere.

Notations. Let $\mathbb{R}_{\geq 0} := [0, \infty)$ and $\mathbb{N} := \{0, 1, \ldots, \infty\}$. Given two vectors x and y of the same dimension, m_x denotes the dimension of x, $x^\top$ denotes the transpose of x, $|x|$ denotes the Euclidean norm of x, and $\langle x, y \rangle := x^\top y$ denotes the scalar product of x and y. Given a nonempty set $K \subset \mathbb{R}^{m_x}$, $|x|_K := \inf_{y \in K} |x - y|$ defines the distance between x and the set K, $\text{int}(K)$ denotes the interior of K, and $\text{cl}(K)$ denotes its closure. For a nonempty set $O \subset \mathbb{R}^{m_x}$, $K \backslash O$ denotes the subset of elements of K that are not in O. For a symmetric semi-positive definite matrix $\Gamma \in \mathbb{R}^{n \times n}$, $\lambda_{min}(\Gamma)$ and $\text{tr}(\Gamma)$ denote the smallest eigenvalue of Γ, and the trace of Γ, respectively. Finally, for a function $\phi : \mathbb{R}^n \to \mathbb{R}^m$, $\text{dom}\,\phi$ denotes the domain of definition of ϕ.

II. PRELIMINARIES

A. Hybrid Dynamical Systems

Following [10], we view a hybrid dynamical system $\mathcal{H}$ as the combination of a constrained differential and a constrained difference equations given by

$$\mathcal{H} : \begin{cases} \dot{x} = F(x) & x \in C \\ x^+ = G(x) & x \in D, \end{cases} \tag{1}$$

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 $F : C \to \mathbb{R}^{m_x}$ and $G : D \to \mathbb{R}^{m_x}$, respectively.

A. Saoud and R. G. Sanfelice are with Department of Electrical and Computer Engineering, University of California, Santa Cruz, asaoud,ricardo@ucsc.edu M. Maghenem is with CNRS, GIPSA Lab, Grenoble, France mohamed.maghenem@gipsa-lab.grenoble-inp.fr.
This research has been partially supported by NSF Grants no. ECS-1710621, CNS-1544396, and CNS-2039054, by AFOSR Grants no. FA9550-19-1-0053, FA9550-19-1-0169, and FA9550-20-1-0238, and by CITRIS and the Banatao Institute at the University of California.

A hybrid arc ϕ is defined on a hybrid time domain denoted $\operatorname{dom}\phi \subset \mathbb{R}_{\geq 0} \times \mathbb{N}$. The hybrid arc ϕ is parametrized by an ordinary time variable $t \in \mathbb{R}_{\geq 0}$ and a discrete jump variable $j \in \mathbb{N}$. Its domain of definition $\operatorname{dom}\phi$ is such that for each $(T, J) \in \operatorname{dom}\phi$, $\operatorname{dom}\phi \cap ([0, T] \times \{0, 1, \ldots, J\}) = \cup_{j=0}^{J} ([t_j, t_{j+1}] \times \{j\})$ 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 1 (Concept of solution to $\mathcal{H}$): A hybrid arc $\phi : \operatorname{dom}\phi \to \mathbb{R}^{m_\phi}$ 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 $I^j := \{t : (t, j) \in \operatorname{dom}\phi\}$ has nonempty interior, $t \mapsto \phi(t, j)$ is locally absolutely continuous and

$$\begin{aligned} &\phi(t, j) \in C && \text{for all} && t \in \text{int}(I^j), \\ &\dot{\phi}(t, j) = F(\phi(t, j)) && \text{for almost all} && t \in I^j; \end{aligned} \tag{2}$$

(S2) for all $(t, j) \in \operatorname{dom}\phi$ such that $(t, j + 1) \in \operatorname{dom}\phi$,

$$\phi(t, j) \in D, \qquad \phi(t, j + 1) = G(\phi(t, j)). \tag{3}$$

•

A solution ϕ to $\mathcal{H}$ is said to be maximal if there is no solution ψ to $\mathcal{H}$ such that $\phi(t, j) = \psi(t, j)$ for all $(t, j) \in \operatorname{dom}\phi$ and $\operatorname{dom}\phi$ is a proper subset of $\operatorname{dom}\psi$. It is said to be nontrivial if $\operatorname{dom}\phi$ contains at least two points. It is said to be continuous if it never jumps. It is said to be eventually discrete if $T := \sup_t \operatorname{dom}\phi < \infty$ and $\operatorname{dom}\phi \cap (\{T\} \times \mathbb{N})$ contains at least two points. It is said to be eventually continuous if $J := \sup_j \operatorname{dom}\phi < \infty$ and $\operatorname{dom}\phi \cap (\mathbb{R}_{\geq 0} \times \{J\})$ contains at least two points. It is said to be Zeno if it is complete and $\sup_t \operatorname{dom}\phi < \infty$. The system $\mathcal{H}$ is said to be forward complete if the domain of each maximal solution is unbounded. It is said to be pre-forward complete if the domain of each maximal solution is closed.

B. Hybrid Basic Conditions

Well-posed [10, Definition 6.2] hybrid systems refer to a class of hybrid systems where the solutions enjoy very useful structural properties [10, Chapter 6]. A hybrid system $\mathcal{H} = (C, F, D, G)$ is well-posed if the following conditions, known as the hybrid basic conditions, are satisfied, see [10, Assumption 6.5] and [10, Theorem. 6.8] for more details.

(A1) The sets C and D are closed.
(A2) The flow map $F : C \to \mathbb{R}^n$ is continuous.
(A3) The jump map $G : D \to \mathbb{R}^n$ is continuous.

C. Uniform Exponential Stability in Hybrid Systems

In this section, we recall the notion of uniform exponential stability of a general closed set for hybrid systems [11]. This notion will be used later to characterize the convergence rate of the hybrid gradient algorithm.

Definition 2: Consider the hybrid system $\mathcal{H} := (C, F, D, G)$ and let $\mathcal{A} \subset \mathcal{X}$ be a closed set. The set $\mathcal{A}$ is said to be globally uniformly pre-exponentially stable for $\mathcal{H}$ if there exist $\kappa > 0$ and $\lambda > 0$ such that each solution ϕ to $\mathcal{H}$ satisfies

$$|\phi(t, j)|_{\mathcal{A}} \leq \kappa \exp(-\lambda(t + j))|\phi(0, 0)|_{\mathcal{A}} \quad \forall (t, j) \in \operatorname{dom}\phi. \tag{4}$$

When, additionally, every maximal solution to $\mathcal{H}$ is complete, we say that $\mathcal{A}$ is globally uniformly exponentially stable for $\mathcal{H}$. •

The constant λ in (4) is called the convergence rate, the decay rate, or the rate of the descent of the solutions to $\mathcal{H}$ towards $\mathcal{A}$.

III. General Context

Consider the linear relationship

$$y(t) = \theta^\top \psi(t), \tag{5}$$

where $y : \operatorname{dom} y \to \mathbb{R}$ represents a measured output, $\psi : \operatorname{dom}\psi \to \mathbb{R}^{m_\psi}$ represents a measured input, called regressor, and $\theta \in \mathbb{R}^{m_\theta}$ is a constant vector of unknown parameters to be identified. To estimate θ, one can use the linear estimator of the form

$$\hat{y}(t) = \hat{\theta}(t)^\top \psi(t),$$

where $\hat{y} : \operatorname{dom}\hat{y} \to \mathbb{R}$ is the estimated output and $\hat{\theta} : \operatorname{dom}\hat{\theta} \to \mathbb{R}^{m_\theta}$ is the estimate of the unknown parameter θ. The error between the true and the estimated outputs is given by

$$e(t) := \hat{y}(t) - y(t) = \tilde{\theta}(t)^\top \psi(t), \tag{6}$$

where $\tilde{\theta} := \hat{\theta} - \theta$.

The gradient descent algorithm assigns discrete-time dynamics, when $t \in \mathbb{N}$, or continuous-time dynamics, when $t \in \mathbb{R}_{\geq 0}$, to $\hat{\theta}$ so that it converges to θ, using the knowledge of e and ψ. To do so, the following cost function is introduced:

$$J(e) := \frac{1}{2}e^2. \tag{7}$$

A. The Continuous-Time Gradient Algorithm

In the continuous-time setting, namely, when the regressor signal ψ is viewed as a continuous-time function with $\operatorname{dom}\psi = [0, +\infty)$, the gradient algorithm is given by

$$\dot{\hat{\theta}} = -\gamma \nabla_{\hat{\theta}} J(e(t)) = \gamma \psi(t)(\psi(t)^\top \hat{\theta} - y(t)), \tag{8}$$

where $\gamma > 0$ is a positive constant representing the adaptation rate [3]. As a result, the dynamics of the estimation error is given by

$$\dot{\tilde{\theta}} = -\gamma \psi(t)\psi(t)^\top \tilde{\theta}. \tag{9}$$

Analyzing the convergence of the gradient algorithm can be translated into showing uniform exponential stability of the origin for the time-varying system in (9). It is well known that the following persistency of excitation condition is necessary and sufficient for uniform exponential stability for the origin of (9); see [4].

C1 For the regressor signal $t \mapsto \psi(t) \in \mathbb{R}^n$, there exist $T > 0$, $\mu_1 > 0$, and $\mu_2 > 0$ such that, for each $t_o \geq 0$,

$$\mu_2 I \geq \int_{t_o}^{t_o+T} \psi(s)\psi(s)^\top ds \geq \mu_1 I.$$

Furthermore, under C1, a lower bound on the convergence rate is provided in [6] and [5].

In [8] and [9], condition C1 is replaced by the following slightly more restrictive condition:

C2 For the regressor signal $t \mapsto \psi(t) \in \mathbb{R}^n$, there exist $T > 0$, $\mu > 0$, and $\bar{\psi} > 0$ such that, for each $t_o \geq 0$,

$$\int_{t_o}^{t_o+T} \psi(s)\psi(s)^\top ds \geq \mu I, \quad \operatorname{ess\,sup}\{|\psi(s)| : s \geq 0\} \leq \bar{\psi}.$$

Under C2, a lower bound on the convergence rate is provided in [8] and a strict Lyapunov function is constructed for (9) in [9].

B. The Discrete-Time Gradient Algorithm

In the discrete-time setting, namely, when the regressor signal ψ is viewed as a discrete-time function with $\operatorname{dom}\psi = \mathbb{N}$, the gradient algorithm is given by

$$\hat{\theta}(t+1) = \hat{\theta}(t) - \sigma(t)\nabla_{\hat{\theta}} J(e), \tag{10}$$

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

$$\tilde{\theta}^+ = \left(I - \frac{\gamma\psi(t)\psi(t)^\top}{1+\gamma|\psi(t)|^2}\right)\tilde{\theta}. \tag{11}$$

The existing approaches to study (9) translate naturally to the study of (11). In the later case, conditions C1 and C2 reduce to the following condition:

C3 For the regressor signal $t \mapsto \psi(t) \in \mathbb{R}^n$, there exist $T > 0$, $\mu > 0$, and $\bar{\psi} > 0$ such that, for each $t_o \geq 0$,

$$\sum_{t=t_o}^{t_o+T} \psi(t)\psi(t)^\top \geq \mu I, \quad \sup\{|\psi(t)| : t \in \mathbb{N}\} \leq \bar{\psi}.$$

IV. Motivation

In this section, we motivate the benefit of viewing the regressor function ψ in (5) as a hybrid arc via a simple example.

Example 1: Consider the linear relationship in (5), where the regressor function ψ is given by

$$\psi(t) := \begin{cases} [\sin(t) \quad 0]^\top & \text{if} \quad t \in (2j\pi, 2(j+1)\pi), \quad j \in \mathbb{N} \\ [0.5 \quad 1]^\top & \text{otherwise.} \end{cases}$$

When viewing ψ as a discrete-time function defined on $\{0, 2\pi, ...\}$, we can see that the matrix

$$\psi(t)\psi(t)^\top := \begin{bmatrix} 0.25 & 0.5 \\ 0.5 & 1 \end{bmatrix} \qquad \forall t \in \{0, 2\pi, 4\pi, ...\}$$

is constant and not full rank. Hence, it does not satisfy the persistence of excitation condition in C3.

Similarly, when viewing ψ as a continuous-time function defined on $[0, +\infty)$, we can see that

$$\psi(t)\psi(t)^\top := \begin{cases} \begin{bmatrix} \sin(t)^2 & 0 \\ 0 & 0 \end{bmatrix} & \text{if } t \in (2j\pi, 2(j+1)\pi), \ j \in \mathbb{N} \\ \begin{bmatrix} 0.25 & 0.5 \\ 0.5 & 1 \end{bmatrix} & \text{otherwise.} \end{cases}$$

Note that, for each $t_o > 0$ and $T > 0$, we have

$$\int_{t_o}^{t_o+T} \psi(s)\psi(s)^\top ds = \int_{t_o}^{t_o+T} \begin{bmatrix} \sin(s)^2 & 0 \\ 0 & 0 \end{bmatrix} ds = \begin{bmatrix} \int_{t_o}^{t_o+T} \sin(s)^2 ds & 0 \\ 0 & 0 \end{bmatrix}.$$

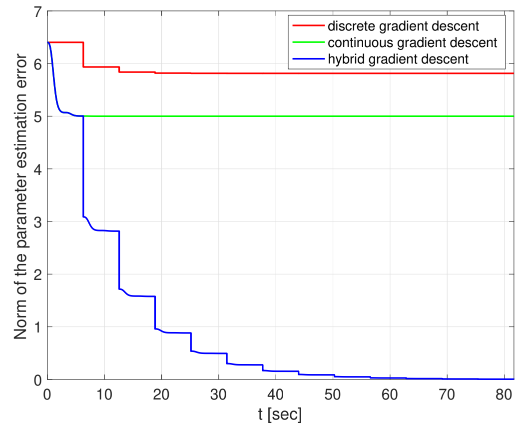

Fig. 1. Evolution of the norm of the parameter error using continuous, discrete and hybrid gradient algorithms.

As a result, the persistence of excitation condition in C2 is not satisfied. □

One of the scenarios where discontinuous regressors, like the one in Example 1, can be encountered is when ψ is function of a solution ϕ to a hybrid system $\mathcal{H} = (C, F, D, G)$ as in (1). In this case, the regressor ψ is a hybrid arc; namely, for some $h : \mathbb{R}^{m_\phi} \to \mathbb{R}^{m_y}$, we have

$$\psi(t, j) := h(\phi(t, j)) \qquad \forall (t, j) \in \operatorname{dom}\phi.$$

Furthermore, the algebraic equation in (5) becomes

$$y(t, j) = \theta^\top \psi(t, j) \qquad \forall (t, j) \in \operatorname{dom}\psi.$$

When viewing the regressor ψ as a hybrid arc exhibiting both flows and jumps, we propose the following hybrid gradient algorithm to update the estimate $\hat{\theta}$ of the unknown parameter θ:

(HG1) Whenever ψ jumps, namely, when $(t, j) \in \operatorname{dom}\psi$ such that $(t, j+1) \in \operatorname{dom}\psi$, we update $\hat{\theta}$ via

$$\hat{\theta}(t, j+1) = \hat{\theta}(t, j) - \frac{\gamma\psi(t, j)(\psi(t, j)^\top\hat{\theta}(t, j) - y(t, j))}{1+\gamma|\psi(t, j)|^2}. \tag{12}$$

(HG2) Whenever ψ flows, namely, when $(t, j) \in \operatorname{dom}\psi$ such that $(t, j+1) \notin \operatorname{dom}\psi$, we update $\hat{\theta}$ via

$$\dot{\hat{\theta}} = -\gamma\psi(t, j)(\psi(t, j)^\top\hat{\theta}(t, j) - y(t, j)). \tag{13}$$

Example 2: Consider the regressor ψ used in Example 1. Figure 1 compares the different gradient algorithms in terms of convergence of the estimation errors. The initial condition for the simulation[1] is $\tilde{\theta}_o := [4 \quad -5]$. The green, the red, and the blue lines represent the evolution of the Eucliden norm of $\tilde{\theta}$ using the continuous-time gradient algorithm in (9), the discrete-time gradient algorithm in (11), and the proposed hybrid gradient algorithm in (12)-(13), respectively. It can be seen that while the continuous and discrete-time algorithms do not allow the convergence of $\tilde{\theta}$ to the origin, our proposed hybrid algorithm ensures this property. □

[1]Files for this simulation can be found at the following address: https://github.com/HybridSystemsLab/HybridGradientDescent

V. Problem Formulation

According to the hybrid gradient algorithm in (HG1)-(HG2), the dynamics of the parameter estimation error $\tilde{\theta} = \hat{\theta} - \theta$ is governed by the hybrid system

$$\mathcal{H}_g : \begin{cases} \begin{bmatrix} \dot{\tilde{\theta}} \\ \dot{t} \\ \dot{j} \end{bmatrix} = \begin{bmatrix} -\gamma\psi(t,j)\psi(t,j)^\top \tilde{\theta} \\ 1 \\ 0 \end{bmatrix} & (\tilde{\theta}, t, j) \in C_g \\ \begin{bmatrix} \tilde{\theta}^+ \\ t^+ \\ j^+ \end{bmatrix} = \begin{bmatrix} \tilde{\theta} - \frac{\gamma\psi(t,j)\psi(t,j)^\top}{1+\gamma|\psi(t,j)|^2}\tilde{\theta} \\ t \\ j+1 \end{bmatrix} & (\tilde{\theta}, t, j) \in D_g, \end{cases} \tag{14}$$

where

$$D_g := \{(\tilde{\theta}, t, j) \in \mathbb{R}^{m_\theta} \times \operatorname{dom}\psi : (t, j+1) \in \operatorname{dom}\psi\}$$

and

$$C_g := \operatorname{cl}((\mathbb{R}^{m_\theta} \times \operatorname{dom}\psi) \backslash D_g).$$

With the hybrid arc ψ given, including t and j as state variables[2], leading to $x = (\tilde{\theta}, t, j)$, makes it possible to write a hybrid system model that is time invariant.

In the rest of the paper, we provide conditions on the regressor ψ to guarantee global uniform exponential stability of the closed set $\mathcal{A} := \{(\tilde{\theta}, t, j) \in \mathbb{R}^{m_\theta} \times \operatorname{dom}\psi : \tilde{\theta} = 0\}$ for $\mathcal{H}_g$. Moreover, our main result provides an estimate of the convergence rate of the hybrid gradient algorithm.

VI. Main Result

To analyze uniform exponential stability of the closed set $\mathcal{A}$ for $\mathcal{H}_g$, we note that $\mathcal{H}_g$ belongs to the following class of hybrid systems with flow and jump maps that are linear in $\tilde{\theta}$:

$$\mathcal{H} : \begin{cases} \dot{x} = \begin{bmatrix} \dot{\tilde{\theta}} \\ \dot{t} \\ \dot{j} \end{bmatrix} = F(x) := \begin{bmatrix} -A(t,j)\tilde{\theta} \\ 1 \\ 0 \end{bmatrix} & x \in C \\ x^+ = \begin{bmatrix} \tilde{\theta}^+ \\ t^+ \\ j^+ \end{bmatrix} =: G(x) = \begin{bmatrix} \tilde{\theta} - B(t,j)\tilde{\theta} \\ t \\ j+1 \end{bmatrix} & x \in D, \end{cases} \tag{15}$$

where $A, B : \operatorname{dom} A = \operatorname{dom} B \to \mathbb{R}^{m_x \times m_x}$ are matrices and $\operatorname{dom} A = \operatorname{dom} B$ is a hybrid time domain, $x := (\tilde{\theta}, t, j) \in \mathcal{X} := \mathbb{R}^{m_\theta} \times \operatorname{dom} A$, $C := \operatorname{cl}(\mathcal{X} \backslash D)$, and

$$D := \{(\tilde{\theta}, t, j) \in \mathcal{X} : (t, j+1) \in \operatorname{dom} A\}.$$

Next, to verify the hybrid basic conditions (A1)-(A3), we consider the following assumption.

Assumption 1 (Regularity properties): For each $j \in \mathbb{N}$, the map $t \mapsto A(t,j)$ is continuous on $I^j := \{t : (t,j) \in \operatorname{dom} A\}$. •

Furthermore, we assume the following structural properties for the matrices A and B to match the properties of the flow and the jump maps in $\mathcal{H}_g$.

[2]Note that the components t an j of the solution x coincide with the hybrid time (t, j) of x.

Assumption 2 (Structural Properties): The matrices A and B satisfy the following properties:

1) For each $(t,j) \in \operatorname{dom} A = \operatorname{dom} B$,
$$A(t,j) = A(t,j)^\top \geq 0 \qquad B(t,j) = B(t,j)^\top \geq 0;$$
2) For each $(t,j) \in \operatorname{dom} A = \operatorname{dom} B$, $|B(t,j)| \leq 1$;
3) There exists $\bar{A} > 0$ such that
$$\operatorname{ess\,sup}\{|A(t,j)| : (t,j) \in \operatorname{dom} A\} \leq \bar{A}.$$

•

Finally, we assume the following hybrid persistency of excitation condition that will enable us to guarantee global uniform exponential stability of the set $\mathcal{A}$ while providing an estimate the hybrid convergence rate.

Assumption 3 (Hybrid Persistence of Excitation): There exist $\bar{k} > 0$ and $\mu > 0$ such that, for each $(t_o, j_o) \in \operatorname{dom} A = \operatorname{dom} B$ and for each hybrid time domain

$$E := \bigcup_{j=j_o}^{J} ([t_j, t_{j+1}] \times \{j\}) \subset \operatorname{dom} A = \operatorname{dom} B \tag{16}$$

with $t_{j_o} := t_o$ and $(t_{J+1} - t_o) + (J - j_o) \geq \bar{k}$, the following holds:

$$\sum_{j=j_o}^{J} \int_{t_j}^{t_{j+1}} A(s,j)ds + \frac{1}{2}\sum_{j=j_o}^{J} B(t_{j+1}, j) \geq \mu I. \tag{17}$$

•

Remark 1: Note that the hybrid system $\mathcal{H}$ in (15) reduces to $\mathcal{H}_g$ in (14) when $A \equiv \gamma\psi\psi^\top$ and $B \equiv \frac{\gamma\psi\psi^\top}{1+\gamma|\psi|^2}$. Furthermore, when the hybrid arc ψ in $\mathcal{H}_g$ is eventually continuous (respectively, eventually discrete or Zeno), Assumptions 2 and 3 reduce to C2 (respectively, C3). •

Remark 2: When the regressor ψ is scalar (i.e, $m_\psi = 1$), then Assumptions 2 and 3 imply that either C2 or C3 holds. However, in the general case when $m_\psi > 1$, it is possible that Assumptions 2 and 3 hold but none of the conditions in C2 and C3 is satisfied as shown in the following example. •

Example 3: For the regressor in Example 1, we note that when $\gamma = 1$ the corresponding maps A and B are given by

$$A(t,j) = \begin{cases} \begin{bmatrix} \sin^2(t) & 0 \\ 0 & 0 \end{bmatrix} & \text{if} \quad t \in (2j\pi, 2(j+1)\pi) \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{if} \quad t = 2j\pi, \end{cases}$$

$$B(t,j) = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{if} \quad t \in (2j\pi, 2(j+1)\pi) \\ \begin{bmatrix} 0.1111 & 0.2222 \\ 0.2222 & 0.4444 \end{bmatrix} & \text{if} \quad t = 2j\pi. \end{cases}$$

Now, one can check that the hybrid persistence of excitation in Assumption 3 holds with $\bar{k} = 2\pi + 1$ and $\mu = 0.21$. □

We have now all the ingredients to state the main result of the paper.

Theorem 1: Given the hybrid system $\mathcal{H}$ in (15), suppose that Assumptions 1, 2, and 3 hold. Then, the closed set $\mathcal{A} := \{(\tilde{\theta}, t, j) \in \mathcal{X} : \tilde{\theta} = 0\}$ is uniformly exponentially stable for $\mathcal{H}$. In particular, for any solution ϕ to $\mathcal{H}$, (4) holds with

$$\lambda := -\log(1-\alpha)/\bar{k}, \tag{18}$$

and

$$\alpha := \frac{2\mu}{\left(1 + (\bar{k}+2)\sqrt{(\bar{A}+2)(1/2 + \bar{A}(\bar{k}+1)^2)}\right)^2}. \tag{19}$$

where $\bar{A}$ is given by Assumption 2 and $\bar{k}$ is given by Assumption 3. □

VII. Proof of Theorem 1

A. Proof of Uniform Stability

To prove uniform stability of the closed set $\mathcal{A}$, we consider the Lyapunov function candidate

$$V(x) := \frac{1}{2}\tilde{\theta}^\top\tilde{\theta} = \frac{1}{2}|x|^2_{\mathcal{A}}. \tag{20}$$

Now, for each $x = (\tilde{\theta}, t, j) \in C$, we have from Assumption 2

$$\langle \nabla V(x), F(x)\rangle = -\tilde{\theta}^\top A(t,j)\tilde{\theta} \le 0.$$

Let us now analyze the variation of V at the jump instants. For each x with $(t, j) \in \operatorname{dom} A$ and $(t, j+1) \in \operatorname{dom} A$, it follows from Assumption 2 that

$$V(G(x)) - V(x) = -\frac{1}{2}\tilde{\theta}^\top B(t,j)\tilde{\theta} \le 0.$$

Hence, for each maximal solution x to $\mathcal{H}$, we conclude that

$$|x(t,j)|_{\mathcal{A}} \le |x(0,0)|_{\mathcal{A}} \qquad \forall (t,j) \in \operatorname{dom} x,$$

which concludes uniform stability of the set $\mathcal{A}$ for $\mathcal{H}$.

B. Proof of Exponential Attractivity

To show exponential stability of the closed set $\mathcal{A}$, we will show that, for each $(t_o, j_o) \in \operatorname{dom} A$ and for each hybrid domain $E := \cup_{j=j_o}^{J}([t_j, t_{j+1}] \times \{j\}) \subset \operatorname{dom} A$ with $(t_{J+1} - t_{j_o}) + (J - j_o) \in [\bar{k}, \bar{k}+1]$, the following inequality is true: For each solution x to $\mathcal{H}$ from $(t,j) = (t_0, j_0)$

$$V(x(t_{J+1}, J)) - V(x(t_o, j_o)) \le (1-\alpha)V(x(t_o, j_o)), \tag{21}$$

where α is in (19). As a consequence, (4) holds with $\lambda := -\log(1-\alpha)/\bar{k}$.

To prove (21), we note that

$$\begin{aligned}\tilde{V} &:= V(x(t_{J+1}, J)) - V(x(t_o, j_o)) \\ &= \sum_{j=j_o}^{J} [V_F(t_j, t_{j+1}, j) + V_G(t_{j+1}, j, j+1)],\end{aligned} \tag{22}$$

where V_F and V_G are given by:

$$V_F(t_j, t_{j+1}, j) := V(x(t_{j+1}, j)) - V(x(t_j, j)) \tag{23}$$

$$V_G(t_{j+1}, j, j+1) := V(x(t_{j+1}, j+1)) - V(x(t_{j+1}, j)). \tag{24}$$

Next, to complete the proof, we will use the following technical lemmas.

Lemma 1: For each $\rho > 0$, the function V_F in (23) satisfies the following inequality for each $j \in \{j_o, \dots, J\}$:

$$\begin{aligned}V_F(t_j, t_{j+1}, j) \le{}& -\rho\bar{A}(\bar{A}+2)(2(j-j_o)+1)(t_{j+1}-t_j)^2\tilde{V} \\ &- \frac{\rho}{1+\rho}\int_{t_j}^{t_{j+1}} \left|A(s,j)^{\frac{1}{2}}\tilde{\theta}(t_o, j_o)\right|^2 ds,\end{aligned}$$

where $\tilde{V}$ is defined in (22) and $\bar{A}$ in Assumption 2. □

Lemma 2: For each $\rho > 0$, the function V_G in (24) satisfies the following inequality for all $j \in \{j_o, \dots, J\}$:

$$\begin{aligned}V_G(t_{j+1}, j, j+1) \le{}& \frac{1}{2}\rho(2(j-j_o)+1)(\bar{A}+2)\tilde{V} \\ &- \frac{1}{2}\frac{\rho}{1+\rho}\left|B(t_{j+1}, j)^{\frac{1}{2}}\tilde{\theta}(t_o, j_o)\right|^2,\end{aligned}$$

where $\tilde{V}$ is defined in (22) and $\bar{A}$ in Assumption 2. □

Combining Lemmas 1 and 2, we obtain the following upper bound on $\tilde{V}$ for each $\rho > 0$:

$$\begin{aligned}\tilde{V} \le\; & -\frac{\rho}{1+\rho}\sum_{j=j_o}^{J}\left|B(t_{j+1}, j)^{\frac{1}{2}}\tilde{\theta}(t_o, j_o)\right|^2 \\ & - \frac{2\rho}{1+\rho}\sum_{j=j_o}^{J}\int_{t_j}^{t_{j+1}}\left|A(s,j)^{\frac{1}{2}}\tilde{\theta}(t_o, j_o)\right|^2 ds \\ & - \frac{\rho}{2}(\bar{A}+2)\tilde{V}\sum_{j=j_o}^{J}(2(j-j_0)+1) \\ & - \rho\bar{A}(\bar{A}+2)\tilde{V}\sum_{j=j_o}^{J}(2(j-j_0)+1)(t_{j+1}-t_j)^2.\end{aligned}$$

Hence,

$$\begin{aligned}\tilde{V} \le & -\frac{2\rho}{1+\rho}V(x(t_o, j_o)) \\ & \times\left(\frac{1}{2}\sum_{j=j_o}^{J}B(t_{j+1}, j) + \sum_{j=j_o}^{J}\int_{t_j}^{t_{j+1}}A(s,j)ds\right) \\ & - \rho(\bar{A}+2)(J-j_0+1)^2(1/2 + \bar{A}(t_{J+1}-t_o)^2)\tilde{V}.\end{aligned}$$

Finally, using Assumption 3, we conclude that

$$\begin{aligned}\tilde{V} \le & -\frac{2\rho\mu}{1+\rho}V(x(t_o, j_o)) \\ & - \rho(\bar{A}+2)(\bar{k}+2)^2(1/2 + \bar{A}(\bar{k}+1)^2)\tilde{V}.\end{aligned}$$

Hence, by choosing

$$\rho := 1/\sqrt{(\bar{A}+2)(\bar{k}+2)^2(1/2 + \bar{A}(\bar{k}+1))}$$

we conclude that

$$V(x(t_{J+1}, J)) - V(x(t_o, j_o)) \le -\alpha V(x(t_o, j_o)),$$

where α is introduced in (19).

VIII. Discussion

Existing approaches to prove exponential stability of the gradient system in (9) include those in [5], [6], [8]. In addition, in [9] an explicit strict Lyapunov function is constructed. Each of the aforementioned approaches allows an explicit estimation of the convergence rate and they extend naturally to the discrete-time version of (9) in (11). However, as we shall show, not all of the aforementioned approaches can be extended to analyze the hybrid gradient system $\mathcal{H}_g$ in (14). We first note that the approach used in this paper is inspired by the result in [8]. Next, we will illustrate the difficulty encountered when trying to extend the results in [5], [6], [9] to the hybrid case.

In [5], the system in (9) is considered under C1. A quadratic Lyapunov function V is used and it was shown that the following solution-based property is satisfied for all $t \geq t_0 \geq 0$:

$$V(x(t)) = V(x(t_0)) - \tilde{\theta}(t_0)^\top \int_{t_0}^{t} \Phi(u,t_0)^\top \psi(u)^\top \psi(u) \Phi(u,t_0) du \tilde{\theta}(t_0) \tag{25}$$

where Φ is the state transition matrix for the system (9). Then using the fact that the matrices $S_0 := \int_{t_0}^{t} \psi(u)^\top \psi(u) du$ and $S := \int_{t_0}^{t} \Phi(u,t_0)^\top \psi(u)^\top \psi(u) \Phi(u,t_0) du$ have the same observability Gramian it can be shown that S is positive definite if and only if S_0 is positive definite. Hence, equation (25) implies that $V(x(t)) \leq (1-\alpha) V(x(t_0))$ with $\alpha = 2\lambda_{\min}(S)$. A similar approach is used to deal with the discrete-time case using the observability Gramian for discrete-time systems [2]. Extending these results to the hybrid case in $\mathcal{H}_g$ is difficult. Indeed, the main difficulty follows from the fact that to define a hybrid observability Gramian, the information about the whole hybrid time domain should be known a priori.

In [6], system (9) is considered under C1. Given $t_o \geq 0$, the following property is established therein:

$$|\tilde{\theta}(t)|^2 - |\tilde{\theta}(t_o)|^2 \geq |\tilde{\theta}(t_o)|^2 f(t,t_o), \tag{26}$$

where

$$f(t,t_o) := \left[\sqrt{\kappa_1(\epsilon_1(t,t_o),\epsilon_2(t,t_o)) + \kappa_2(\epsilon_1(t,t_o))} - \sqrt{\kappa_1(\epsilon_1(t,t_o),\epsilon_2(t,t_o))}\right]^2,$$

$$\epsilon_1(t,t_o) := \lambda_{min}(W(t,t_o)),\ \epsilon_2(t,t_o) := \mathrm{tr}(W(t,t_o)),$$

$W(t,t_o) := \int_{t_o}^{t} \psi(s)\psi(s)^\top ds$, and for particular functions $\kappa_1 : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $\kappa_2 : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to [0,1)$. Note that under C1, it follows that $\epsilon_1(t_o+T,t_o) \geq \mu_1$ and $\epsilon_2(t_o+T,t_o) \leq \mu_2$. Hence, there exists $\gamma \in (0,1)$, such that $|\tilde{\theta}(t_o+T)|^2 - |\tilde{\theta}(t_o)|^2 \geq \gamma |\tilde{\theta}(t_o)|^2$. In the hybrid case, we are able to establish inequalities similar to (26) along the flows and along the jumps separately. That is, given a solution $\tilde{\theta}$ to $\mathcal{H}_g$ and a hybrid domain $E := \cup_{j=j_o}^{J} ([t_j,t_{j+1}] \times \{j\}) \subset \mathrm{dom}\, x$, we can show inequalities of the form

$$|\tilde{\theta}(t_{j+1},j)|^2 - |\tilde{\theta}(t_j,j)|^2 \geq |\tilde{\theta}(t_j,j)|^2 f(t_{j+1},t_j),$$
$$|\tilde{\theta}(t_{j+1},j+1)|^2 - |\tilde{\theta}(t_{j+1},j)|^2 \geq |\tilde{\theta}(t_{j+1},j)|^2 g(j+1,j),$$

where g is, roughly speaking, a discrete version of f. Note that we cannot go any further by using the latter two inequalities due the potential nonlinearities of f and g.

In [9], the system in (9) is considered under C2. The system is shown to admit a continuously differentiable Lyapunov function $V : \mathbb{R}_{\geq 0} \times \mathbb{R}^{m_\theta} \to \mathbb{R}_{\geq 0}$ such that

$$\underline{\alpha}|\tilde{\theta}|^2 \leq V(t,\tilde{\theta}) \leq \bar{\alpha}|\tilde{\theta}|^2, \tag{27}$$

for some $\underline{\alpha} > 0$ and $\bar{\alpha} > 0$. Furthermore, the time derivative of V along the solutions to (9) satisfies, for some $\alpha > 0$, $\dot{V}(t,\tilde{\theta}) \leq -\alpha(\mu/T)|\tilde{\theta}|^2$, where μ and T come from C2. Similarly, for the discrete-time system in (11) and under C3, we can build a Lyapunov function $V : \mathbb{N} \times \mathbb{R}^{m_\theta} \to \mathbb{R}_{\geq 0}$ satisfying (27) and such that, along the solutions to (11),

$$V(j+1,\tilde{\theta}(j+1)) - V(j,\tilde{\theta}(j)) \leq -\alpha \frac{\mu}{J}|\tilde{\theta}|^2, \tag{28}$$

where μ and J come from C3. Now, for the hybrid case in $\mathcal{H}_g$, it is very challenging to find explicitly a scalar Lyapunov function V satisfying (27) that decreases strictly along the flows, even when the excitation is coming from the jumps; and similarly along the jumps, even when the excitation is coming from the flows.

IX. CONCLUSION

This paper proposed a hybrid gradient algorithm to estimate the unknown parameters of an hybrid linear regression problem. When the inputs and the outputs are hybrid arcs, the proposed algorithm updates the estimates continuously during the continuous evolution of the data and instantaneously during their jumps. The proposed hybrid gradient algorithm generalizes the existing continuous and discrete-time gradient descent algorithms. In the sense where a relaxed (hybrid) version of the well-known persistence of excitation condition is shown to be sufficient to guarantee exponential estimation of the parameters. In future work we will develop hybrid parameter estimators for hybrid MRAC models.

REFERENCES

[1] L. Ljung. System identification. *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–19, 1999.

[2] G. Tao. *Adaptive Control Design and Analysis*, volume 37. John Wiley & Sons, 2003.

[3] K. S. Narendra and A. M. Annaswamy. *Stable Adaptive Systems*. Courier Corporation, 2012.

[4] K. S. Narendra and A. M. Annaswamy. Persistent excitation in adaptive systems. *Int. J. of Contr.*, 45(1):127–160, 1987.

[5] B. D. O. Anderson. Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. on Automatic Control,*, 22(1):83–88, Feb 1977.

[6] R. Brockett. The rate of descent for degenerate gradient flows. In *Proc. Math. Theory of Networks and Systems*, 2000.

[7] S. Andersson and S. P. Krishnaprasad. Degenerate gradient flows: a comparison study of convergence rate estimates. In *Proc. of the 41st. IEEE Conf. on Decision and Control*, volume 4, pages 4712–4717, 2002.

[8] A. Loria and E. Panteley. Uniform exponential stability of linear time-varying systems: revisited. *Syst. & Contr. Letters*, 47(1):13–24, 2002.

[9] N. R. Chowdhury, S. Sukumar, M. Maghenem, and A. Loria. On the estimation of algebraic connectivity in graphs with persistently exciting interconnections. *Int. J. of Contr.*, 91(1):132–144, 2018.

[10] R. Goebel, R. G. Sanfelice, and A. R. Teel. *Hybrid Dynamical Systems: Modeling, stability, and robustness*. Princeton University Press, 2012.

[11] R. G. Sanfelice. *Hybrid Feedback Control*. Princeton University Press, 2021.