Asymptotic Tracking and Linear-like Behavior Using Multi-Model Adaptive Control

Mohamad T. Shahab and Daniel E. Miller

Abstract—In this paper, we consider the problem of tracking for a discrete-time plant with unknown plant parameters; we assume knowledge of an upper bound on the plant order, and for each admissible order we assume knowledge of a compact set in which the plant parameters lie. We carry out parameter estimation of an associated auxiliary model; indeed, for each admissible dimension, we cover the set of admissible parameters by a finite number of compact and convex sets and use an original-projection-algorithm-based estimator for each set. At each point in time, we employ a switching algorithm to determine which model and parameter estimates are used in the pole-placement-based control law. We prove that this adaptive controller guarantees desirable linear-like closed-loop behavior: exponential stability, a bounded noise gain in every ᴘ-norm, a convolution bound on the effect of the exogenous inputs, as well as exponential tracking for certain classes of reference and noise signals; this linear-like behavior is leveraged to immediately show tolerance to a degree of plant time-variations and unmodelled dynamics.

I. Introduction

Adaptive control is an approach used to deal with systems with uncertain or time-varying parameters. The classical adaptive controller consists of a linear time-invariant compensator together with a tuning mechanism to adjust the compensator parameters to match the plant; the tuning mechanism often consists of a plant parameter estimator together with a formula to yield the compensator parameters. The first general results of adaptive control came about around 1980, e.g. [14], [16], [35], [39] and [40]. However, these controllers typically do not tolerate unmodeled dynamics, time-variations, and/or noise/disturbances very well, see e.g. [44]; furthermore, they put stringent assumptions on *a priori* information about the plant. Over the following two decades, there was a good deal of effort to address these shortcomings. A common approach was to make small controller design changes, such as the use of σ-modification, signal normalization, or a deadzone, e.g. see [21], [28], [24], and [29]; another approach imposes a convexity assumption on the set of admissible parameters, which is utilized in the estimation process to restrain the estimates of the plant parameters to the convex set, e.g. see [38], [53], [52], [51], and [27]. Although these controllers provide a degree of tolerance to noise, unmodelled dynamics

and/or slow time-variations, in general, they do not provide exponential stability nor a bounded noise gain[1]; clearly, it is desirable that the closed-loop system exhibits such linear-like properties. While we can prove a form of exponential stability if a persistent excitation condition is satisfied, e.g. see [2], this places a stringent requirement on exogenous inputs, which we would like to avoid.

There are various non-classical approaches to adaptive control that provide some linear-like system properties, like those that include multiple estimators, multiple controller designs and/or a switching mechanism.

(i) Pre-routed switching between a list of candidate controllers was used in [15] and [32]; while exponential stability is proven, a bounded gain on the noise is not proven and transient behavior can be poor.

(ii) *Supervisory Control* [36], [37], [11], [20], [19] and [50] provides a more efficient way to switch between candidate controllers and is shown to provide more desirable properties; however, the complexity of the approach grows with the size of plant uncertainty. In certain circumstances a bounded gain on the noise is proven, and some progress has been made to deal with time-variation, e.g. see [50], though crisp bounds on closed-loop behavior are not proven.

(iii) Another approach which utilizes multi-controllers is based on the concept of unfalsified control–see [45], [49], [8], [6], [10] and [9]; the goal here is to switch any destabilizing controller out of the loop completely. In [9], time-varying plants are considered as well. In some cases, an ISS-type of stability is proven.

(iv) A different falsification-based approach is discussed in [54] proving exponential stability and some form of tolerance to noise; however, a bounded gain on the noise is not proven.

(v) Another family of non-classical approaches are ones that use multiple parallel parameter estimators, like in [41], [42]; here a switching algorithm is used to choose which parameter estimate to use in the control law; however, exponential stability is not explicitly proven.

(vi) In a closely-related approach to the previous one, a weighted combination of multiple controllers is used, e.g. [26], [7] and [5], or a weighted combination of multiple parameter estimates is used, e.g. [43] and [18]; while improvement in closed-loop behavior is shown, convexity of the uncertainty set is strictly assumed.

(vii) There are also Lyapunov-based approaches to switching adaptive control which provide some form of exponential stability, e.g. [12], and bounded-input bounded-state behavior, e.g. [3] and [4]; however, *a priori* knowledge of suitable

M.T. Shahab was with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada, and is with the Computer, Electrical and Mathematical Science and Engineering (CEMSE) Division, King Abdullah University of Science & Technology (KAUST), Thuwal 23955, Saudi Arabia. (email: mohamad.shahab@uwaterloo.ca)

D.E. Miller is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada. (email: miller@uwaterloo.ca)

---

[1]An exception is the work of Ydstie [53]: a bounded gain is proven, but a crisp bound on the effects of the initial condition is not provided.

Lyapunov functions is required as part of the controller.
In all of these approaches, a linear-like convolution bound on the closed-loop behavior is not proven.

Recently, an approach was developed which guarantees a linear-like convolution bound on the closed-loop behavior, in both the one-step-ahead control setting [30] and [34], and in the pole-placement control setting [31] and [33]; this yields exponential stability as well as a bounded gain on the noise (in every p-norm); this linear-like behavior is leveraged to immediately show tolerance to a degree of plant time-variations and unmodelled dynamics. As far as the authors know, such linear-like bounds have never before been proven in the adaptive setting. The key idea is to employ an estimator based on the original (ideal) Projection Algorithm together with projection of the parameter estimates onto a given compact and convex set. In [33], the convexity assumption is weakened slightly (without completely removing it) and stability is proven, but not tracking. In [47], the same linear-like result is proven without any convexity assumption but only for the case of 1st-order one-step-ahead adaptive control.

In this paper, we consider the problem of tracking the sum of a finite number of sinusoids of known frequency in the presence of plant uncertainty. We assume knowledge of an upper bound on the plant order, and for each admissible order we assume knowledge of a compact set in which the plant parameters lie; although we impose some natural technical assumptions on the sets, we do not assume that they are convex. To facilitate the tracking requirement, rather than directly estimating the plant parameters, we instead estimate the parameters of a suitably defined auxiliary plant model. We use the compactness of the parameter uncertainty set for each admissible order to prove that it is contained in a finite union of compact and convex sets; we construct a parameter estimator for each of these compact and convex sets, based on the original projection algorithm. A switching algorithm is used to determine which estimates are used in the controller at a given point in time. We prove that **the desired linear-like convolution bounds are achieved**, and if the reference and disturbance signals belong to the aforementioned class of sinusoids then the **tracking error goes exponentially to zero**. We would like to point out that a preliminary version of this paper appears in [48]; however, it deals only with the problem of step tracking with a known plant order, and no proofs are provided.

We now provide an outline of the paper. In Section II, we discuss the unknown plant, the auxiliary plant to be used for estimation, and the uncertainty sets. We introduce the multi-model adaptive controller consisting of the estimator, the control law, and the switching algorithm in Section III. In Section IV we provide the main result of the paper, which shows that the closed-loop behavior satisfies the desired convolution bound, and exponential tracking for certain classes of reference and noise signals is provided. In Section V, we show, briefly, that this result holds in the presence of plant time-variations and unmodelled dynamics. We will provide an illustrative simulation example in Section VI. Finally, summary and conclusions are provided in Section VII.

Notation. We use standard notation throughout the paper.

We denote $\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{Z}$, $\mathbb{Z}^+$ and $\mathbb{N}$ as the set of real numbers, nonnegative real numbers, integers, nonnegative integers and natural numbers, respectively. Let $\lceil \cdot \rceil$ denote the ceiling function. We will denote the Euclidean-norm of a vector and the induced norm of a matrix by the subscript-less default notation $\| \cdot \|$. For a square matrix A, let $\det(A)$ denote the determinant of A. Also, $\boldsymbol{\ell}_\infty(\mathbb{R}^p)$ denotes the set of $\mathbb{R}^p$-valued bounded sequences; we use the simple notation of $\boldsymbol{\ell}_\infty$ to denote the special case of $\boldsymbol{\ell}_\infty(\mathbb{R})$. For a signal $f \in \boldsymbol{\ell}_\infty$, define the ∞-norm by $\|f\|_\infty := \sup_{t \in \mathbb{Z}} |f(t)|$. For a closed and convex set $\Omega \subset \mathbb{R}^p$, let the function $\underset{\Omega}{\text{Proj}} \{ \cdot \} : \mathbb{R}^p \mapsto \Omega$ denote the projection onto the set Ω; because the set Ω is closed and convex, the function $\underset{\Omega}{\text{Proj}}$ is well-defined. If $\Omega \subset \mathbb{R}^p$ is a compact (closed and bounded) set, we define $\|\Omega\| := \max_{x \in \Omega} \|x\|$. Define the normal vector $\mathbf{e}_j \in \mathbb{R}^p$ of appropriate length p as

$$\mathbf{e}_j := \begin{bmatrix} \overbrace{0 \quad \cdots \quad 0}^{j-1 \text{ elements}} & 1 & 0 & \cdots & 0 \end{bmatrix}^\top.$$

Let $\mathbf{0}_{p \times q}$ and $\mathbf{0}_p$ denote the $p \times q$ matrix and the $p \times 1$ vector whose entries are all zeros, respectively. Let I_p denote the identity matrix of size p.

II. THE SETUP

A. The Plant

We consider the nth-order linear time-invariant discrete-time plant

$$y(t+1) = \sum_{j=1}^n a_j y(t-j+1) + \sum_{j=1}^n b_j u(t-j+1) + w(t),$$
$$t \in \mathbb{Z}, \tag{1}$$

with $y(t), u(t), w(t) \in \mathbb{R}$ denoting the measured output, the control input, and the disturbance/noise input, respectively. Such a plant can be expressed in the (two-sided) z-transform form as

$$\mathbf{A}(z^{-1})Y(z) = \mathbf{B}(z^{-1})U(z) + z^{-1}W(z), \tag{2}$$

with the corresponding polynomials defined as

$$\mathbf{A}(z^{-1}) := 1 - a_1 z^{-1} - a_2 z^{-2} \cdots - a_n z^{-n}, \quad \text{and}$$

$$\mathbf{B}(z^{-1}) := b_1 z^{-1} + b_2 z^{-2} \cdots + b_n z^{-n},$$

with $Y(z), U(z)$ and $W(z)$ denoting the z-transform of $y(t), u(t)$ and $w(t)$, respectively. The plant can be represented by the strictly proper transfer function $\frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}$. We can represent the plant model by the vector of parameters

$$\theta = \begin{bmatrix} a_1 & a_2 & \cdots & a_n & b_1 & b_2 & \cdots & b_n \end{bmatrix}^\top \in \mathbb{R}^{2n}.$$

The objective is to control the system when θ is unknown but lies in a set of admissible parameters. Since our goal is to provide uniform bounds, we shall require that this set be compact; unlike in our earlier work [33], we will not insist on convexity. Since we will be using a pole-placement approach, we will require that $z^n \mathbf{A}(z^{-1})$ and $z^n \mathbf{B}(z^{-1})$ be coprime. Indeed, it turns out that our approach works if the plant order

n is not known exactly, but rather we have an upper bound $\bar{n}$ on n. To this end, for each $n \in \{1, 2, \ldots, \bar{n}\}$ we let

$$\boldsymbol{\Theta}_n \subset \mathbb{R}^{2n}$$

denote the set of admissible parameters, and impose

Assumption 1. For every $n \in \{1, 2, \ldots, \bar{n}\}$:

(1) the set $\boldsymbol{\Theta}_n$ is compact[2] and
(2) for every $\theta \in \boldsymbol{\Theta}_n$, the corresponding polynomials $z^n \mathbf{A}(z^{-1})$ and $z^n \mathbf{B}(z^{-1})$ are coprime.

B. The Control Objective

The objective is to prove an exponential form of stability, a convolution bound on the noise and on a general reference signal, and tracking (and disturbance rejection) of the sum of a finite number of sinusoids of known frequency. To characterize this class, we use the standard approach in the literature, e.g. see Goodwin and Sin [17, pp. 155-157]. Let $y^*(t) \in \mathbb{R}$ be the reference signal. If the goal is to track reference signals which are a weighted sum of sinusoids with distinct frequencies of $\{\omega_1, \omega_2, \ldots, \omega_{g_1}\} \subset (0, \pi)$, then define

$$\overline{\mathbf{Q}}(z^{-1}) := \prod_{p=1}^{g_1} [(1 - e^{j\omega_p} z^{-1})(1 - e^{-j\omega_p} z^{-1})];$$

if we wish to track set-points as well, then multiply this by $1 - z^{-1}$, and if we wish to track signals of the form $(-1)^t$ as well, then multiply this by $1 + z^{-1}$; we label the resulting polynomial $\mathbf{Q}_1(z^{-1})$. So in the z-domain, this class of reference signals is described by

$$\mathbf{Q}_1(z^{-1}) Y^*(z) = 0,$$

with $Y^*(z)$ as the z-transform of $y^*(t)$. In a similar way, we can form a (possibly different) polynomial $\mathbf{Q}_2(z^{-1})$ with distinct roots on the unit circle so that the class of disturbances to be rejected is described by

$$\mathbf{Q}_2(z^{-1}) W(z) = 0.$$

The gth-order polynomial

$$\mathbf{Q}(z^{-1}) = 1 - q_1 z^{-1} - q_2 z^{-2} \cdots - q_g z^{-g} \quad (3a)$$

$$:= \text{least common multiple of } \{\mathbf{Q}_1(z^{-1}), \mathbf{Q}_2(z^{-1})\} \quad (3b)$$

will be used in the controller synthesis. Note that if the goal is only stability (without tracking or disturbance rejection), then $\mathbf{Q}(z^{-1}) = 1$.

As we know from classical continuous-time control, if the plant has a zero at the origin then we cannot design an LTI stabilizing controller which ensures that the plant tracks steps. To this end, we impose the following natural assumption:

Assumption 2. For each $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta \in \boldsymbol{\Theta}_n$, the corresponding polynomial $z^n \mathbf{B}(z^{-1})$ and the polynomial $z^g \mathbf{Q}(z^{-1})$ are coprime.

Remark 1. *Here we allow the plant to have zeros outside the open unit desk as long as they are not roots of $z^g \mathbf{Q}(z^{-1})$. Hence, the plant could be unstable and non-minimum phase, which makes it challenging to control.*

[2]It could very well be that $\boldsymbol{\Theta}_n$ is empty for some n.

C. The Auxiliary Plant

If n is known and the set of admissible parameters $\boldsymbol{\Theta}_n$ is convex, then the classical approach is to carry out system identification of the plant in the usual way, e.g. [17], and design the pole-placement based control law in such a way as to force an "internal model of $\mathbf{Q}(z^{-1})$" into the controller; this has been shown to be quite effective in classical approaches which prove asymptotic stability, e.g. see [17], as well as in our recent work [33] where we prove exponential stability and step tracking. If, however, the set of admissible parameters is not convex, which can be the case here, the standard trick is to replace it with its closed convex hull. Unfortunately, often that set will contain models that violate coprimeness, so we need another approach. The compactness of the set of admissible parameters can be utilized to easily prove that it is contained in a finite union of convex sets with desirable properties; we can then use an estimator for each convex set and from time to time switch between estimates for use in the control law. We have extended our approach of [33] to the case of multi-estimators: we can achieve exponential stability and a convolution bound on the effect of the exogenous signals, but we have failed to prove asymptotic tracking (although it could very well be true). This is due, in large part, to the fact that the approach does not rule out persistent switching[3]. We will deal with this difficulty by doing system identification on a related auxiliary plant model rather than the original plant model.

With $y^*, w \in \boldsymbol{\ell}_\infty$, let us define the tracking error ε by

$$\varepsilon(t) := y(t) - y^*(t); \quad (4)$$

also define an auxiliary control input $v(t) \in \mathbb{R}$ and its z-transform counterpart by

$$V(z) := \mathbf{Q}(z^{-1}) U(z) \quad (5a)$$

$$\Leftrightarrow v(t) = u(t) - q_1 u(t-1) - \cdots - q_g u(t-g). \quad (5b)$$

If we multiply both sides of the z-transformed counterpart of the plant model (2) by $\mathbf{Q}(z^{-1})$ and use the definition in (5), then we end up with

$$\mathbf{Q}(z^{-1}) \mathbf{A}(z^{-1}) Y(z) = \mathbf{B}(z^{-1}) V(z) + z^{-1} \mathbf{Q}(z^{-1}) W(z);$$

denoting the z-transform of $\varepsilon(t)$ by $\mathcal{E}(z)$, if we subtract $\mathbf{Q}(z^{-1}) \mathbf{A}(z^{-1}) Y^*(z)$ from both sides of the above equation then we obtain the auxiliary plant model

$$\underbrace{\mathbf{Q}(z^{-1}) \mathbf{A}(z^{-1})}_{=:\bar{\mathbf{A}}(z^{-1})} \mathcal{E}(z) = \mathbf{B}(z^{-1}) V(z) +$$

$$z^{-1} \underbrace{\mathbf{Q}(z^{-1}) \left[W(z) - z\mathbf{A}(z^{-1}) Y^*(z) \right]}_{=:\bar{W}(z)}, \quad (6)$$

or in other words

$$\bar{\mathbf{A}}(z^{-1}) \mathcal{E}(z) = \mathbf{B}(z^{-1}) V(z) + z^{-1} \bar{W}(z). \quad (7)$$

To proceed, we now examine the polynomial $\bar{\mathbf{A}}(z^{-1})$ carefully; we have

$$\bar{\mathbf{A}}(z^{-1}) =: 1 - \bar{a}_1 z^{-1} - \bar{a}_2 z^{-2} \cdots - \bar{a}_{n+g} z^{-(n+g)}$$

[3]Persistent switching is also not ruled out in the approaches of [4], [9], [10], [36] and [37].

$$= \mathbf{Q}(z^{-1})\mathbf{A}(z^{-1}) = \left(1 - \sum_{j=1}^{g} q_j z^{-j}\right)\left(1 - \sum_{j=1}^{n} a_j z^{-j}\right);$$

it is easy to see that there exists a matrix $\mathcal{V}(\mathbf{Q}, n) \in \mathbb{R}^{(n+g)\times(n+1)}$, depending solely on $\mathbf{Q}(z^{-1})$, so that we can rewrite $\bar{\mathbf{A}}(z^{-1})$ as

$$\bar{\mathbf{A}}(z^{-1}) = 1 - \begin{bmatrix} z^{-1} & z^{-2} & \cdots & z^{-(n+g)} \end{bmatrix} \mathcal{V}(\mathbf{Q}, n) \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

We see that the parameters of $\bar{\mathbf{A}}(z^{-1})$ are determined in a simple way from those of $\mathbf{A}(z^{-1})$. Indeed, for each pair of n and $\mathbf{Q}(z^{-1})$, we can form a matrix $\bar{\mathcal{V}}(\mathbf{Q}, n) \in \mathbb{R}^{(2n+g)\times(2n+1)}$ defined by

$$\bar{\mathcal{V}}(\mathbf{Q}, n) := \begin{bmatrix} \mathcal{V}(\mathbf{Q}, n) \\ & I_n \end{bmatrix} \tag{8}$$

so that the parameters of $\bar{\mathbf{A}}(z^{-1})$ and $\mathbf{B}(z^{-1})$ of (7) are given by

$$\bar{\mathcal{V}}(\mathbf{Q}, n) \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \\ b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \\ \vdots \\ \bar{a}_{n+g} \\ b_1 \\ \vdots \\ b_n \end{bmatrix} =: \theta^*.$$

So the set of admissible parameters of (7) is given by

$$\tilde{\boldsymbol{\Theta}}_n := \left\{ \bar{\mathcal{V}}(\mathbf{Q}, n) \begin{bmatrix} 1 \\ \theta \end{bmatrix} : \theta \in \boldsymbol{\Theta}_n \right\} \subset \mathbb{R}^{2n+g}. \tag{9}$$

Using this notation, the auxiliary plant (7) can now be put into regressor form:

$$\varepsilon(t+1) = \psi(t)^\top \theta^* + \bar{w}(t), \tag{10}$$

with $\bar{w}(t)$ as the inverse z-transform of $\bar{W}(z)$, $\psi(t) \in \mathbb{R}^{2n+g}$ defined as

$$\psi(t) := \begin{bmatrix} \varepsilon(t) & \varepsilon(t-1) & \cdots & \varepsilon(t-n-g+1) \\ & v(t) & v(t-1)\cdots & v(t-n+1) \end{bmatrix}^\top$$

and $\theta^* \in \tilde{\boldsymbol{\Theta}}_n$. As in the case of the original plant (1), the order is not known, though it is known that $n \in \{1, 2, \ldots, \bar{n}\}$; hence, while the dimension of $\psi(t)$ clearly depends on n, to enhance readability this will not be made explicit.

Remark 2. *Carrying out system identification on the auxiliary plant* (10)*, instead of the original plant* (1)*, makes it easier to analyze and prove asymptotic tracking. The reason is simple: the tracking error is a sub-state of the regression vector of* (10) *but not of* (1)*. That being said, the proof is quite involved (see Theorem 1 and Proposition 3).*

D. Uncertainty sets

Since for every $n \in \{1, 2, \ldots, \bar{n}\}$, $\boldsymbol{\Theta}_n$ is compact, it follows that $\tilde{\boldsymbol{\Theta}}_n$ is as well; also, because of Assumptions 1 and 2 we see that for every $\theta^* \in \tilde{\boldsymbol{\Theta}}_n$, the corresponding polynomials $z^{n+g}\bar{\mathbf{A}}(z^{-1})$ and $z^n \mathbf{B}(z^{-1})$ are coprime. Of course, if we

were to replace $\tilde{\boldsymbol{\Theta}}_n$ by its convex hull, then those properties may fail to hold. This brings us to the following result. We show that for any $n \in \{1, 2, \ldots, \bar{n}\}$, $\tilde{\boldsymbol{\Theta}}_n$ can be approximated by a finite number of convex sets with desired properties.

Proposition 1. *For every* $n \in \{1, 2, \ldots, \bar{n}\}$ *and* $\mu > 0$*, there exist a finite number of convex, compact sets* $\tilde{\boldsymbol{\Theta}}_n^i \subset \mathbb{R}^{2n+g}$ $(i = 1, 2, \ldots, m_n)$ *that satisfy*

(i) $\tilde{\boldsymbol{\Theta}}_n \subset \bigcup_{i=1}^{m_n} \tilde{\boldsymbol{\Theta}}_n^i$,

(ii) *for every* $\theta^* \in \bigcup_{i=1}^{m_n} \tilde{\boldsymbol{\Theta}}_n^i$ *there exists a* $\tilde{\theta}^* \in \tilde{\boldsymbol{\Theta}}_n$ *that satisfy* $\|\tilde{\theta}^* - \theta^*\| \leq \mu$.

Furthermore, if $\mu > 0$ *is sufficiently small, then we can choose the* $\tilde{\boldsymbol{\Theta}}_n^i$*'s to have an additional property as well:*

(iii) *for every* $\theta^* \in \bigcup_{i=1}^{m_n} \tilde{\boldsymbol{\Theta}}_n^i$*, the corresponding pair of polynomials* $z^{n+g}\bar{\mathbf{A}}(z^{-1})$ *and* $z^n \mathbf{B}(z^{-1})$ *are coprime.*

Proof. This is equivalent to Proposition 1 in [48]; the proof is there, where it utilizes the Heine-Borel Theorem. ∎

In general, finding a set of m_n $\tilde{\boldsymbol{\Theta}}_n^i$'s which satisfy the desired properties of Proposition 1 for which m_n is small and $\tilde{\boldsymbol{\Theta}}_n^i$ has "nice[4] structure" is not easy. However, this is not the focus of our paper. This *covering* problem is an open research problem—e.g. see [1], [13] and [23]. So at this point we assume that this process has been done for each $n \in \{1, 2, \ldots, \bar{n}\}$; we will show an example on how to do this in Section VI.

The idea here is to use a parameter estimator for each compact and convex set, and at each point in time we choose which one to use in constructing the control law. At this point, for every $n \in \{1, 2, \ldots, \bar{n}\}$ we have at hand m_n compact and convex parameter uncertainty sets (they can be disjoint or overlapping) that correspond to models of nth-order plants:

$$\tilde{\boldsymbol{\Theta}}_1^1, \tilde{\boldsymbol{\Theta}}_1^2, \ldots, \tilde{\boldsymbol{\Theta}}_1^{m_1}, \tilde{\boldsymbol{\Theta}}_2^1, \tilde{\boldsymbol{\Theta}}_2^2, \ldots, \tilde{\boldsymbol{\Theta}}_2^{m_2}, \ldots, \tilde{\boldsymbol{\Theta}}_{\bar{n}}^1, \tilde{\boldsymbol{\Theta}}_{\bar{n}}^2 \ldots, \tilde{\boldsymbol{\Theta}}_{\bar{n}}^{m_{\bar{n}}},$$

yielding a total of $m := m_1 + m_2 + \cdots + m_{\bar{n}}$ sets. For ease of notation, we re-label these sets as

$$\mathcal{S}_i \subset \mathbb{R}^{2n_i+g}, \quad i = 1, 2, \ldots m;$$

here $n_i \in \{1, 2, \ldots, \bar{n}\}$ represents the plant order of the associated model.

Now define the index set

$$\mathcal{I}^* := \{1, 2, \ldots, m\}.$$

For each $\theta^* \in \mathcal{S}_i$, $i = 1, 2, \ldots, m$, we define

$$i^*(\theta^*) = \min\{i \in \mathcal{I}^* : \theta^* \in \mathcal{S}_i\};$$

when there is no ambiguity, we will drop the argument and simply write i^*. Before proceeding, define $\bar{\mathbf{s}} := \max_i \|\mathcal{S}_i\|$.

III. THE MULTI-MODEL ADAPTIVE CONTROLLER

In this section we present the proposed adaptive controller; we discuss parameter multi-estimators, the associated switching control law, and the switching algorithm. The proposed controller is illustrated in the block diagram of the closed-loop system given in Figure 1.

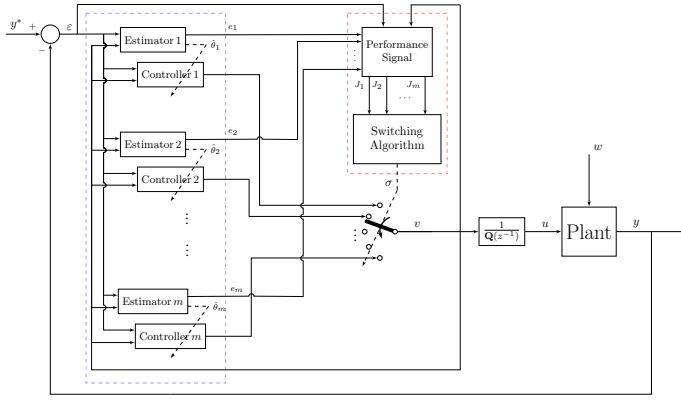[4]Nice in the sense that it is computationally easy to project onto it.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3052745, IEEE Transactions on Automatic Control

SHAHAB AND MILLER : ASYMPTOTIC TRACKING AND LINEAR-LIKE BEHAVIOR USING MULTI-MODEL ADAPTIVE CONTROL 5

Fig. 1.   Block diagram of the closed-loop system; enclosed inside the dashed boxes are the multiple estimators/controllers (blue), and the switching mechanism (red).

A. Parameter Estimation

First, for each $i \in \mathcal{I}^*$, the corresponding regressor vector is defined by $\psi_i(t) \in \mathbb{R}^{2n_i + g}$:

$$\psi_i(t) := \begin{bmatrix} \varepsilon(t) & \varepsilon(t-1) & \cdots & \varepsilon(t-n_i-g+1) \\ & v(t) & v(t-1) & \cdots & v(t-n_i+1) \end{bmatrix}^\top.$$

So we know that the auxiliary plant (10) can be rewritten as

$$\varepsilon(t+1) = \psi_{i^*}(t)^\top \theta^* + \bar{w}(t). \tag{11}$$

Given an estimate $\hat{\theta}_{i^*}(t) \in \mathcal{S}_{i^*}$ at time t, we can now define the prediction error associated with this model by

$$e_{i^*}(t+1) := \varepsilon(t+1) - \psi_{i^*}(t)^\top \hat{\theta}_{i^*}(t). \tag{12}$$

A common way to obtain the next parameter estimate is to solve the optimization problem

$$\underset{\theta}{\operatorname{argmin}} \left\{ \|\theta - \hat{\theta}_{i^*}(t)\| : \varepsilon(t+1) = \psi_{i^*}(t)^\top \theta \right\}$$

yielding the original **Projection Algorithm**

$$\hat{\theta}_{i^*}(t+1) = \begin{cases} \hat{\theta}_{i^*}(t) & \text{if } \psi_{i^*}(t) = 0 \\ \hat{\theta}_{i^*}(t) + \frac{\psi_{i^*}(t)}{\|\psi_{i^*}(t)\|^2} e_{i^*}(t+1) & \text{otherwise.} \end{cases}$$

Of course, we do not know the value of i^*; so we simply generalize the above steps by defining

$$e_i(t+1) := \varepsilon(t+1) - \psi_i(t)^\top \hat{\theta}_i(t), \qquad i \in \mathcal{I}^* \tag{13}$$

and with $\hat{\theta}_i(t) \in \mathcal{S}_i$, we adopt the projection algorithm for each $i \in \mathcal{I}^*$:

$$\hat{\theta}_i(t+1) = \begin{cases} \hat{\theta}_i(t) & \text{if } \psi_i(t) = 0 \\ \hat{\theta}_i(t) + \frac{\psi_i(t)}{\|\psi_i(t)\|^2} e_i(t+1) & \text{otherwise.} \end{cases} \tag{14}$$

It is common in the literature to modify this algorithm by adding a positive constant to the denominator of the second term; while this prevents numerical problems when $\psi_i(t)$ is close to zero, when used in an adaptive controller it turns out that exponential stability is lost—see [33] and [30]. The concern about numerical issues is addressed in an alternate way in [33] by turning off the estimation if it is clear that the noise is swamping the estimation error. Here we do the same thing; however, to make our proof work, we also redefine

the denominator in the estimation update law. To proceed, we make three modifications to (14):

1) To restrict $\hat{\theta}_i(t)$ to $\mathcal{S}_i$, after applying the suitably modified version of (14), we project it onto $\mathcal{S}_i$.
2) We define the longest data vector by $\bar{\psi}(t) \in \mathbb{R}^{2\bar{n}+g}$:

$$\bar{\psi}(t) := \begin{bmatrix} \varepsilon(t) & \varepsilon(t-1) & \cdots & \varepsilon(t-\bar{n}-g+1) \\ & v(t) & v(t-1) & \cdots & v(t-\bar{n}+1) \end{bmatrix}^\top; \tag{15}$$

observe that $\|\psi_i(t)\| \leq \|\bar{\psi}(t)\|$ for all i. We will replace $\psi_i(t)$ by $\bar{\psi}(t)$ in the denominator of (14).
3) By examining (12), we see that

$$e_{i^*}(t+1) = \psi_{i^*}(t)^\top \left[\theta^* - \hat{\theta}_{i^*}(t) \right] + \bar{w}(t), \tag{16}$$

which means that $|e_{i^*}(t+1)| \leq 2\bar{s}\|\psi_{i^*}(t)\| + |\bar{w}(t)|$. So if $|e_{i^*}(t+1)| > 2\bar{s}\|\psi_{i^*}(t)\|$, then the disturbance may be overwhelming the data, so we turn off the estimator if the gap is too large.

To this end, with $\delta \in (0, \infty]$ and with $\bar{\psi}(t)$ replacing $\psi_i(t)$, we define $\rho_i : \mathbb{Z} \mapsto \{0, 1\}$ by

$$\rho_i(t) := \begin{cases} 1 & \text{if } |e_i(t+1)| < (2\bar{s}+\delta)\|\bar{\psi}(t)\| \\ 0 & \text{otherwise,} \end{cases} \tag{17}$$

which is used to determine when to turn off the algorithm; the larger the value of δ, the larger that we allow $\bar{w}(t)$ to be (versus $\|\bar{\psi}(t)\|$) before we turn off the estimator. This leads to our proposed "vigilant" estimator: the estimator i updates are computed as follows:

$$\check{\theta}_i(t+1) = \hat{\theta}_i(t) + \rho_i(t) \frac{\psi_i(t)}{\|\bar{\psi}(t)\|^2} e_i(t+1) \tag{18a}$$

$$\hat{\theta}_i(t+1) = \underset{\mathcal{S}_i}{\operatorname{Proj}} \left\{ \check{\theta}_i(t+1) \right\}. \tag{18b}$$

The function ρ_i is used to alleviate the concern of dividing by zero without losing the desired properties of the original projection algorithm. In the case of $\delta = \infty$, we will adopt the understanding that $\infty \times 0 = 0$, in which case the formula in (18a) collapses to the original algorithm (14) with ψ_i in the denominator replaced by $\bar{\psi}$. If $\delta < \infty$, we can be assured the update term is bounded above by $2\bar{s}+\delta$, which alleviates concerns about having infinite gain.

Define the (correct) parameter estimation error $\tilde{\theta}_{i^*}(t) := \hat{\theta}_{i^*}(t) - \theta^*$. The following lists properties of the estimation algorithm (18). These properties are similar to ones found in Proposition 1 and 3 of [33]; the difference arises when the order is unknown, in which case $\psi_i(t)$ and $\bar{\psi}(t)$ differ for some i.

Proposition 2. *For every $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta^* \in \tilde{\Theta}_n$, and every $t_0 \in \mathbb{Z}$, $t_2 > t_1 \geq t_0$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i \, (i \in \mathcal{I}^*)$ and $w, y^* \in \ell_\infty$, when the estimation algorithm in (18) is applied to the corresponding auxiliary plant (10), the following holds:*
*1) for **every estimator** $i = 1, 2, \ldots, m$,*

$$\|\hat{\theta}_i(t_2) - \hat{\theta}_i(t_1)\| \leq \sum_{j=t_1}^{t_2-1} \rho_i(j) \frac{|e_i(j+1)|}{\|\bar{\psi}(j)\|}. \tag{19}$$

2) for the **correct estimator** i^*,

$$\|\tilde{\theta}_{i^*}(t_2)\|^2 \leq \|\tilde{\theta}_{i^*}(t_1)\|^2 +$$

$$\sum_{j=t_1}^{t_2-1} \rho_{i^*}(j) \left[-\frac{1}{2} \frac{e_{i^*}(j+1)^2}{\|\bar{\psi}(j)\|^2} + 2 \frac{\bar{w}(j)^2}{\|\bar{\psi}(j)\|^2} \right]. \quad (20)$$

Proof. See the Appendix. ∎

B. Switching Control Law

For each i, the parameter estimate $\hat{\theta}_i(t)$ is partitioned naturally as

$$\hat{\theta}_i(t) =: \begin{bmatrix} \hat{a}_{i,1}(t) & \hat{a}_{i,2}(t) & \cdots & \hat{a}_{i,n_i+g}(t) \\ & \hat{b}_{i,1}(t) & \hat{b}_{i,2}(t) & \cdots & \hat{b}_{i,n_i}(t) \end{bmatrix}^\top;$$

associated with these estimates are the polynomials

$$\hat{\mathbf{A}}_i(t, z^{-1}) = 1 - \hat{a}_{i,1}(t)z^{-1} - \hat{a}_{i,2}(t)z^{-2} \cdots - \hat{a}_{i,n_i+g}(t)z^{-(n_i+g)},$$

$$\hat{\mathbf{B}}_i(t, z^{-1}) = \hat{b}_{i,1}(t)z^{-1} + \hat{b}_{i,2}(t)z^{-2} \cdots + \hat{b}_{i,n_i}(t)z^{-n_i}.$$

Next we design a $(n_i + g)^{\text{th}}$-order strictly proper controller; we choose the following polynomials

$$\hat{\mathbf{L}}_i(t, z^{-1}) = 1 + \hat{l}_{i,1}(t)z^{-1} + \hat{l}_{i,2}(t)z^{-2} + \cdots + \hat{l}_{i,n_i}(t)z^{-n_i},$$

$$\hat{\mathbf{P}}_i(t, z^{-1}) = \hat{p}_{i,1}(t)z^{-1} + \hat{p}_{i,2}(t)z^{-2} + \cdots + \hat{p}_{i,n_i+g}(t)z^{-(n_i+g)}$$

so as to place all closed-loop poles at $z = 0$:

$$\hat{\mathbf{A}}_i(t, z^{-1})\hat{\mathbf{L}}_i(t, z^{-1}) + \hat{\mathbf{B}}_i(t, z^{-1})\hat{\mathbf{P}}_i(t, z^{-1}) = 1. \quad (21)$$

Since $z^{n_i+g}\hat{\mathbf{A}}_i(t, z^{-1})$ and $z^{n_i}\hat{\mathbf{B}}_i(t, z^{-1})$ are coprime by design, we know that there exist unique $\hat{\mathbf{L}}_i(t, z^{-1})$ and $\hat{\mathbf{P}}_i(t, z^{-1})$ which satisfy this equation—see [22, Theorem 2.3.1]; this entails solving a linear equation. It is also easy to prove that the coefficients of $\hat{\mathbf{L}}_i(t, z^{-1})$ and $\hat{\mathbf{P}}_i(t, z^{-1})$ are analytic functions of $\hat{\theta}_i(t) \in \mathcal{S}_i$. For a suitable choice of $i \in \mathcal{I}^*$ at time t, we define the control input by

$$\hat{\mathbf{L}}_i(t-1, z^{-1})V(z) = -\hat{\mathbf{P}}_i(t-1, z^{-1})\mathcal{E}(z). \quad (22)$$

This can be written in terms of the data vector $\psi_i(t)$: to this end, we define the control gains $\hat{K}_i(t) \in \mathbb{R}^{2n_i+g}$ by

$$\hat{K}_i(t) := \begin{bmatrix} -\hat{p}_{i,1}(t) & -\hat{p}_{i,2}(t) & \cdots & -\hat{p}_{i,n_i+g}(t) \\ & -\hat{l}_{i,1}(t) & -\hat{l}_{i,2}(t) & \cdots & -\hat{l}_{i,n_i}(t) \end{bmatrix} \quad (23)$$

so that (22) becomes

$$v(t) = \hat{K}_i(t-1)\psi_i(t-1).$$

We will use a switching signal $\sigma : \mathbb{Z} \to \mathcal{I}^*$ to denote the index i: $\sigma(t)$ denotes the index of the controller to use at time t.

In our earlier work [33], we considered the problem of closed-loop stability (but not tracking) in the case of switching between 2 estimators of the same dimension. Unfortunately, the approach does not extend in a simple way to the case of $m > 2$ estimators, so we will need a different algorithm. As we will soon see, our closed-loop system behavior will in large part be determined by a time-varying matrix $\mathcal{A}_{\sigma(t)}(t) \in \mathbb{R}^{2\bar{n}+g}$ (see (38)); at all times this matrix will be deadbeat, i.e. all of its eigenvalues will be at zero. However, its product

$$\mathcal{A}_{\sigma(t)}(t) \times \mathcal{A}_{\sigma(t-1)}(t-1) \times \cdots \times \mathcal{A}_{\sigma(t_0)}(t_0), \qquad t \geq t_0$$

will not usually be deadbeat. A natural solution to this problem is to update the estimators every $2\bar{n} + g$ steps; the problem with this idea is that we end up with no information about $e_i(t+1)$ between the updates, so the closed-loop system is not amenable to analysis. So our solution procedure will need to be different: we update $\sigma(t)$ only every $N \geq 2\bar{n} + g$ steps; however, we keep the estimators running and the control gains updating; the aforementioned product of matrices is still not deadbeat, but it is close to being so, in a sense which will be apparent from the proof. To this end, we define a sequence of switching times as follows: we initialize $\hat{t}_0 := t_0$ and then define

$$\hat{t}_\ell := \hat{t}_0 + \ell N, \qquad \ell \in \mathbb{N}.$$

So the switching signal is piecewise constant of the form

$$\sigma(t) = \sigma(\hat{t}_\ell), \qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \quad \ell \in \mathbb{Z}^+; \quad (24)$$

the algorithm to compute $\sigma(\hat{t}_\ell)$ will be introduced shortly. We propose the choice of the control law

$$v(t) = \hat{K}_{\sigma(t-1)}(t-1)\psi_{\sigma(t-1)}(t-1), \qquad t > t_0, \quad (25)$$

which generates the auxiliary control input; this is combined with (5) to yield the plant control input

$$u(t) = v(t) + \sum_{j=1}^{g} q_j u(t-j), \qquad t > t_0. \quad (26)$$

What remains to be defined is the choice of the switching signal $\sigma(\hat{t}_\ell)$, which we will do in the next subsection.

C. Switching Algorithm

With $N \in \mathbb{N}$, define the set of switching times by

$$\mathcal{T}_N := \left\{ \hat{t}_\ell \geq \hat{t}_0 \ : \ \hat{t}_\ell = \hat{t}_0 + \ell N, \ \ell \in \mathbb{Z}^+ \right\}. \quad (27)$$

To proceed, for each $i \in \mathcal{I}^*$ we define a performance signal $J_i : \mathcal{T}_N \to \mathbb{R}^+$ by

$$J_i(\hat{t}_\ell) := \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \rho_i(j) \frac{|e_i(j+1)|}{\|\bar{\psi}(j)\|}, \qquad \ell \in \mathbb{Z}^+; \quad (28)$$

this quantity is an upper bound on the amount of change in $\hat{\theta}_i(t)$ on the interval $[\hat{t}_\ell, \hat{t}_{\ell+1}]$. We may expect the estimator with the least amount of update to be the best one, which would lead to a switching signal of the form

$$\sigma(\hat{t}_{\ell+1}) = \underset{i \in \mathcal{I}^*}{\operatorname{argmin}} \ J_i(\hat{t}_\ell).$$

Although this rule works in every simulation that we have tried, a proof remains elusive; a potential problem is that the switching signal could oscillate between two bad choices, and never (or rarely) choose a "correct" one. Instead, we propose a different approach. At each switching time $\hat{t}_\ell$ we have an admissible set $\mathcal{I}(\hat{t}_\ell)$: we initialize $\mathcal{I}(\hat{t}_0) = \mathcal{I}^*$, and we obtain $\mathcal{I}(\hat{t}_{\ell+1})$ from $\mathcal{I}(\hat{t}_\ell)$ by removing all $j \in \mathcal{I}(\hat{t}_\ell)$ satisfying

$$J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \leq J_j(\hat{t}_\ell),$$

i.e. we keep all models in the admissible index set for which the performance signal is better (i.e. smaller) than the one we are currently using; clearly $j = \sigma(\hat{t}_\ell)$ satisfies this bound, but

more j's may as well; if this results in $\mathcal{I}(\hat{t}_{\ell+1})$ being empty, then we **reset** $\mathcal{I}(\hat{t}_{\ell+1})$ to be $\mathcal{I}^*$. This *Switching Algorithm* is summarized as follows: with $\sigma(\hat{t}_0) = \sigma_0$ and $\mathcal{I}(\hat{t}_0) = \mathcal{I}^*$:

$$\hat{\mathcal{I}}(\hat{t}_\ell) = \left\{ i \in \mathcal{I}^* \ : \ J_i(\hat{t}_\ell) < J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\}, \tag{29a}$$

$$\mathcal{I}(\hat{t}_{\ell+1}) = \begin{cases} \mathcal{I}^* & \text{if } \mathcal{I}(\hat{t}_\ell) \cap \hat{\mathcal{I}}(\hat{t}_\ell) = \varnothing \\ \mathcal{I}(\hat{t}_\ell) \cap \hat{\mathcal{I}}(\hat{t}_\ell) & \text{otherwise,} \end{cases} \tag{29b}$$

$$\sigma(\hat{t}_{\ell+1}) = \operatorname*{argmin}_{i \in \mathcal{I}(\hat{t}_{\ell+1})} J_i(\hat{t}_\ell), \qquad \ell \in \mathbb{Z}^+. \tag{29c}$$

Remark 3. *We define the **index set reset times** as those $\hat{t}_\ell$, $\ell \in \mathbb{Z}^+$, for which $\mathcal{I}(\hat{t}_\ell) = \mathcal{I}^*$.*

We refer to Figure 2 for an illustration about time steps, switching times, and index set reset times on the timeline. Now we present a desirable property of the switching algorithm (29).

Lemma 1. *When the controller* (18), (23)–(25), *and* (27)–(29) *is applied to the auxiliary plant* (10), *for every* $n \in \{1, 2, \ldots, \bar{n}\}$ *and* $\theta^* \in \tilde{\mathbf{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $N \geq 1$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$ *and* $w, y^* \in \boldsymbol{\ell}_\infty$, *if* $\hat{t}_{\underline{\ell}}$ *and* $\hat{t}_{\bar{\ell}}$ *are two consecutive index set reset times, then there exists a* $\ell^* \in [\underline{\ell}, \bar{\ell})$ *such that:*

$$J_{\sigma(\hat{t}_{\ell^*})}(\hat{t}_{\ell^*}) \leq J_{i^*}(\hat{t}_{\ell^*}). \tag{30}$$

Remark 4. *Lemma 1 says that, between every two index set resets, there is an interval of the form $[\hat{t}_{\ell^*}, \hat{t}_{\ell^*+1})$ for which the performance associated with the chosen index is equal to or better than that of the performance associated with the correct index.*

Proof of Lemma 1. Let $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta^* \in \tilde{\mathbf{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $N \geq 1$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary. Let $\hat{t}_{\underline{\ell}}$ and $\hat{t}_{\bar{\ell}}$ be two consecutive index set reset times.

We prove (30) by contradiction; assume that

$$J_{\sigma(\hat{t}_j)}(\hat{t}_j) > J_{i^*}(\hat{t}_j), \quad \text{for all } j \in [\underline{\ell}, \bar{\ell}). \tag{31}$$

Then, according to (29a), we should have

$$i^* \in \hat{\mathcal{I}}(\hat{t}_j), \qquad j \in [\underline{\ell}, \bar{\ell}). \tag{32}$$

We know by the definition of index resets that for all $j \in (\underline{\ell}, \bar{\ell})$ we have $\mathcal{I}(\hat{t}_j) \neq \mathcal{I}^*$, which means that by (29b)

$$\mathcal{I}(\hat{t}_j) = \mathcal{I}(\hat{t}_{j-1}) \cap \hat{\mathcal{I}}(\hat{t}_{j-1}), \qquad j \in (\underline{\ell}, \bar{\ell});$$

then by induction we see that

$$\mathcal{I}(\hat{t}_j) = \mathcal{I}(\hat{t}_{\underline{\ell}}) \cap \hat{\mathcal{I}}(\hat{t}_{\underline{\ell}}) \cap \hat{\mathcal{I}}(\hat{t}_{\underline{\ell}+1}) \cap \cdots \cap \hat{\mathcal{I}}(\hat{t}_{j-2}) \cap \hat{\mathcal{I}}(\hat{t}_{j-1}),$$
$$j \in (\underline{\ell}, \bar{\ell}).$$

But $\mathcal{I}(\hat{t}_{\underline{\ell}}) = \mathcal{I}^*$, so using (32) in the above, we see that

$$i^* \in \mathcal{I}(\hat{t}_j), \qquad j \in [\underline{\ell}, \bar{\ell}] \tag{33}$$

as well. So according to this and to (32) we have $i^* \in \mathcal{I}(\hat{t}_{\bar{\ell}-1}) \cap \hat{\mathcal{I}}(\hat{t}_{\bar{\ell}-1})$. However, we know by the definition of index resets and (29b) that $\mathcal{I}(\hat{t}_{\bar{\ell}-1}) \cap \hat{\mathcal{I}}(\hat{t}_{\bar{\ell}-1}) = \varnothing$, which is a contradiction, so it must be that (31) does not hold. ∎

In the above we do not make any claim that $\theta^* \in \mathcal{S}_{\sigma(t)}$ at any time; it only makes an indirect statement about the size of the prediction error. It turns out that this is enough to ensure that the desired closed-loop behavior is attained.

IV. THE MAIN RESULT

We will define a vector $\bar{\phi}(t) \in \mathbb{R}^{2(\bar{n}+g)}$

$$\begin{aligned} \bar{\phi}(t) := \big[y(t) \quad y(t-1) \quad \cdots \quad y(t-\bar{n}-g+1) \\ u(t) \quad u(t-1) \quad \cdots \quad u(t-\bar{n}-g+1) \big]^\top \end{aligned}$$

to serve as the "plant state"; while this is longer than what is needed for a minimal state representation of (1), the choice will facilitate our analysis. Recall that the vectors ψ_i, $i \in \mathcal{I}^*$, and $\bar{\psi}$ contain values of the tracking error and the auxiliary control input, while the vector $\bar{\phi}$ contains values of the plant input and output. Before proceeding, it is convenient to define a weighted sum of past values of y^*:

$$\tilde{y}^*(t) := \sum_{j=0}^{\bar{n}+g-1} |y^*(t-j)|. \tag{34}$$

Theorem 1. *For every* $\lambda \in (0, 1)$, $\delta \in (0, \infty]$ *and* $N \geq 2\bar{n}+g$, *there exists a constant* $\gamma > 0$ *so that when the adaptive controller consisting of the parameter estimators* (18), *control gains* (23), *switching signal* (24) *with switching times* (27), *performance signal* (28), *switching algorithm* (29), *and control law* (25) *and* (26), *is applied to the plant* (1), *for every* $n \in \{1, 2, \ldots, \bar{n}\}$ *and* $\theta \in \mathbf{\Theta}_n$, $t_0 \in \mathbb{Z}$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$, *and* $w, y^* \in \boldsymbol{\ell}_\infty$,
i) the following bound holds:

$$\|\bar{\phi}(t)\| \leq \gamma \lambda^{t-\tau} \|\bar{\phi}(\tau)\| +$$
$$\gamma \sum_{j=\tau}^{t-1} \lambda^{t-1-j} (|w(j)| + |\tilde{y}^*(j+1)|), \ t > \tau \geq t_0; \tag{35}$$

ii) if $\mathbf{Q}(z^{-1}) Y^*(z) = 0$ *and* $\mathbf{Q}(z^{-1}) W(z) = 0$, *then* $y(t) \to y^*(t)$ *exponentially fast, in the sense that*

$$|\varepsilon(t)| \leq \gamma \lambda^{t-t_0} \left(\|\bar{\phi}(t_0)\| + \|y^*\|_\infty \right), \qquad t \geq t_0.$$

Remark 5. *The above result shows that the closed-loop system experiences linear-like behavior. There is a **uniform exponential decay** bound on the effect of the initial condition, and a **convolution** bound on the effect of the exogenous inputs. This implies that the system has a bounded gain (from w and y^* to y) in **every** p-norm. For example, for $p = \infty$, we see from the above bound that*

$$\|\bar{\phi}(t)\| \leq \tfrac{\gamma(\bar{n}+g)}{1-\lambda} \left(\lambda^{t-t_0} \|\bar{\phi}(t_0)\| + \|w\|_\infty + \|y^*\|_\infty \right), \ t \geq t_0.$$

Hence, if $w, y^* \in \boldsymbol{\ell}_\infty$, then $y, u \in \boldsymbol{\ell}_\infty$, so ε, v, e_i $(i \in \mathcal{I}^*)$ lie in $\boldsymbol{\ell}_\infty$ as well.

Proving Theorem 1 requires two steps:

- First, we analyze the adaptive control system to obtain a desired bound on the key quantity $\bar{\psi}(t)$ (which consists of present and past values of the tracking error and the auxiliary input), which plays a key role in the auxiliary model (10), the parameter estimator (18), and the control law (25); this requires a careful analysis of the closed-loop system.
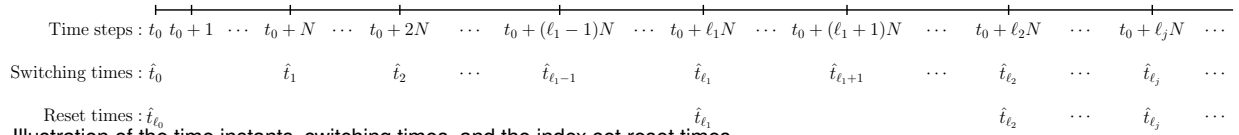
Fig. 2. Illustration of the time instants, switching times, and the index set reset times.

- Second, we use linear system theory to translate the bound on $\bar{\psi}$ to a bound on $\bar{\phi}$ (which consists of present and past values of the plant's input and output).

To enhance readability and to focus the reader's attention on the most important aspects of the approach, we will present the first part in the form of a Proposition, whose proof we place in the main body, while the proof of the second part is placed in the Appendix.

Proposition 3. *For every $\lambda \in (0,1)$, $\delta \in (0,\infty]$ and $N \geq 2\bar{n}+g$, there exists a constant $c > 0$ so that when the adaptive controller* (18), (23)–(25), *and* (27)–(29) *is applied to the auxiliary plant* (10), *for every $n \in \{1,2,\ldots,\bar{n}\}$ and $\theta^* \in \tilde{\mathbf{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$, the following holds*

$$\|\bar{\psi}(t)\| \leq c\lambda^{t-\tau}\|\bar{\psi}(\tau)\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1}|\bar{w}(j)|, \quad t > \tau \geq t_0.$$

Before presenting the proof of this Proposition, we need a crude bound on the closed-loop behavior; the proof of this is in the Appendix.

Lemma 2. *For every $p \geq 0$, there exist constants $\bar{c}_1, \bar{c}_2 \geq 1$ so that when the adaptive controller* (18), (23)–(29) *is applied to the plant* (1), *for every $n \in \{1,2,\ldots,\bar{n}\}$ and $\theta \in \mathbf{\Theta}_n$, $t_0 \in \mathbb{Z}$, $t \geq t_0$, $N \geq 1$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$, the following hold:*

i) $\|\bar{\psi}(t+p)\| \leq \bar{c}_1\|\bar{\psi}(t)\| + \bar{c}_1 \sum_{j=0}^{p-1} |\bar{w}(t+j)|.$

ii) $\|\bar{\phi}(t+p)\| \leq \bar{c}_2\|\bar{\phi}(t)\| + \bar{c}_2 \sum_{j=0}^{p-1} (|w(t+j)| + |\tilde{y}^*(t+j)|).$

Proof of Proposition 3. Fix $\lambda \in (0,1)$, $\delta \in (0,\infty]$ and $N \geq 2\bar{n}+g$. Let $n \in \{1,2,\ldots,\bar{n}\}$, $\theta^* \in \tilde{\mathbf{\Theta}}_n$, $t_0 \in \mathbb{Z}$, $\bar{\psi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \boldsymbol{\ell}_\infty$ be arbitrary. We denote the sequence of index set reset times by $\hat{t}_{\ell_0}, \hat{t}_{\ell_1}, \hat{t}_{\ell_2}, \ldots$ (see Figure 2).

Step 1: Obtain a state-space model describing $\bar{\psi}(t)$ for $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$.

It will be convenient for analysis to have all of the parameter estimates and controller gains to be of the same length. To this end, we will pad $\hat{\theta}_i(t)$ and $\hat{K}_i(t)$ by zeros in the appropriate locations: we define $\hat{\Theta}_i(t) \in \mathbb{R}^{2\bar{n}+g}$ by

$$\hat{\Theta}_i(t) := \begin{bmatrix} \hat{a}_{i,1}(t) & \hat{a}_{i,2}(t) & \cdots & \hat{a}_{i,n_i+g}(t) & \mathbf{0}_{\bar{n}-n_i}^\top \\ \hat{b}_{i,1}(t) & \hat{b}_{i,2}(t) & \cdots & \hat{b}_{i,n_i}(t) & \mathbf{0}_{\bar{n}-n_i}^\top \end{bmatrix}^\top$$

and $\hat{K}_i(t) \in \mathbb{R}^{2\bar{n}+g}$ by

$$\hat{K}_i(t) := \begin{bmatrix} -\hat{p}_{i,1}(t) & -\hat{p}_{i,2}(t) & \cdots & -\hat{p}_{i,n_i+g}(t) & \mathbf{0}_{\bar{n}-n_i}^\top \\ -\hat{l}_{i,1}(t) & -\hat{l}_{i,2}(t) & \cdots & -\hat{l}_{i,n_i}(t) & \mathbf{0}_{\bar{n}-n_i}^\top \end{bmatrix};$$

so by definition of the prediction error (13) and from the

control law in (25) we have

$$\begin{aligned} \varepsilon(t+1) &= \hat{\theta}_{\sigma(t)}(t)^\top \psi_{\sigma(t)}(t) + e_{\sigma(t)}(t+1) \\ &= \hat{\Theta}_{\sigma(t)}(t)^\top \bar{\psi}(t) + e_{\sigma(t)}(t+1) \end{aligned} \tag{36}$$

$$\begin{aligned} v(t+1) &= \hat{K}_{\sigma(t)}(t)\psi_{\sigma(t)}(t) \\ &= \hat{\bar{K}}_{\sigma(t)}(t)\bar{\psi}(t). \end{aligned} \tag{37}$$

Next, define the matrix $\mathcal{A}_i(t) \in \mathbb{R}^{(2\bar{n}+g)\times(2\bar{n}+g)}$ by

$$\mathcal{A}_i(t) := \begin{bmatrix} \hat{\Theta}_i(t)^\top \\ \begin{bmatrix} I_{\bar{n}+g-1} & \mathbf{0}_{(\bar{n}+g-1)\times(\bar{n}+1)} \end{bmatrix} \\ \hat{\bar{K}}_i(t) \\ \begin{bmatrix} \mathbf{0}_{(\bar{n}-1)\times(\bar{n}+g)} & I_{\bar{n}-1} & \mathbf{0}_{\bar{n}-1} \end{bmatrix} \end{bmatrix}; \tag{38}$$

using (21) it is easy to verify that the characteristic equation of the matrix $\mathcal{A}_i(t)$ (for frozen time t) satisfies

$$\begin{aligned} &\det\big(zI_{2\bar{n}+g} - \mathcal{A}_i(t)\big) \\ &= z^{2\bar{n}+g}[\hat{\mathbf{A}}_i(t,z^{-1})\hat{\mathbf{L}}_i(t,z^{-1}) + \hat{\mathbf{B}}_i(t,z^{-1})\hat{\mathbf{P}}_i(t,z^{-1})] \\ &= z^{2\bar{n}+g}. \end{aligned}$$

This means that for every $i \in \mathcal{I}^*$, for each time t the matrix $\mathcal{A}_i(t)$ has all of its eigenvalues at zero. Also define $B_1 := \mathbf{e}_1 \in \mathbb{R}^{2\bar{n}+g}$ and

$$\Delta_i(t) := \rho_i(t)\frac{e_i(t+1)}{\|\bar{\psi}(t)\|^2}B_1\bar{\psi}(t)^\top \tag{39}$$

so we have

$$B_1 e_i(t+1) = \Delta_i(t)\bar{\psi}(t) + B_1 \underbrace{[1 - \rho_i(t)]e_i(t+1)}_{=:\eta_i(t)}.$$

From (36) and (37), the fact that the switching signal is constant on $[\hat{t}_\ell, \hat{t}_{\ell+1})$, and the definition of $\bar{\psi}$, we have that

$$\begin{aligned} \bar{\psi}(t+1) &= \mathcal{A}_{\sigma(t)}(t)\bar{\psi}(t) + B_1 e_{\sigma(t)}(t+1) \\ &= \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\bar{\psi}(t) + \big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\big]\bar{\psi}(t) + \\ &\quad B_1 e_{\sigma(\hat{t}_\ell)}(t+1) \\ &= \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\bar{\psi}(t) + \\ &\quad \big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) + \Delta_{\sigma(\hat{t}_\ell)}(t)\big]\bar{\psi}(t) + B_1\eta_{\sigma(\hat{t}_\ell)}(t), \\ &\qquad t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \ \ell \in \mathbb{Z}^+. \end{aligned} \tag{40}$$

Step 2: Obtain a bound on $\|\bar{\psi}(\hat{t}_{\ell+1})\|$ in terms of $\|\bar{\psi}(\hat{t}_\ell)\|$.

We now are going to analyze the key equation (40) in detail; we make the following observations. For $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$, we have $\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \in \mathbb{R}^{(2\bar{n}+g)\times(2\bar{n}+g)}$ to be a constant matrix with all eigenvalues equal to zero; since $N \geq 2\bar{n}+g$,

$$\big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\big]^{\hat{t}_{\ell+1}-\hat{t}_\ell} = \big[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\big]^N = 0. \tag{41}$$

Next, note that for $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$ we have that the dimension of $\hat{\theta}_{\sigma(t)}(t)$ is constant; by utilizing part 1) of Proposition 2 to provide a bound on the difference between parameter estimates

at two different point in time, the fact that the controller gains are analytic functions of the parameter estimates, and (38), we conclude that there exists a constant c_1 such that

$$
\begin{aligned}
& \left\| \mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\| \\
& \leq \left\| \hat{\Theta}_{\sigma(\hat{t}_\ell)}(t) - \hat{\Theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\| + \left\| \hat{\bar{K}}_{\sigma(\hat{t}_\ell)}(t) - \hat{\bar{K}}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\| \\
& = \left\| \hat{\theta}_{\sigma(\hat{t}_\ell)}(t) - \hat{\theta}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\| + \left\| \hat{K}_{\sigma(\hat{t}_\ell)}(t) - \hat{K}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\| \\
& \leq (1 + c_1) \sum_{j=\hat{t}_\ell}^{t-1} \rho_{\sigma(\hat{t}_\ell)}(j) \frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|}, \\
& \hspace{2cm} t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \ \ell \in \mathbb{Z}^+.
\end{aligned}
\tag{42}
$$

From (39) we obtain

$$
\|\Delta_{\sigma(\hat{t}_\ell)}(t)\| = \rho_{\sigma(\hat{t}_\ell)}(t) \frac{|e_{\sigma(\hat{t}_\ell)}(t+1)|}{\|\bar{\psi}(t)\|}.
\tag{43}
$$

So from (42), (43) and the definition of the performance signal (28), there exists a constant c_2 so that for all $t \in [\hat{t}_\ell, \hat{t}_{\ell+1})$:

$$
\begin{aligned}
& \left\| \mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) + \Delta_{\sigma(\hat{t}_\ell)}(t) \right\| \\
& \leq \left\| \mathcal{A}_{\sigma(\hat{t}_\ell)}(t) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right\| + \|\Delta_{\sigma(\hat{t}_\ell)}(t)\| \\
& \leq (1 + c_1) \left(\sum_{j=\hat{t}_\ell}^{t-1} \rho_{\sigma(\hat{t}_\ell)}(j) \frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|} \right) + \\
& \hspace{1.5cm} \rho_{\sigma(\hat{t}_\ell)}(t) \frac{|e_{\sigma(\hat{t}_\ell)}(t+1)|}{\|\bar{\psi}(t)\|} \\
& \leq \underbrace{(1 + c_1)}_{=:c_2} \sum_{j=\hat{t}_\ell}^{t} \rho_{\sigma(\hat{t}_\ell)}(j) \frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|} \\
& \leq c_2 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \rho_{\sigma(\hat{t}_\ell)}(j) \frac{|e_{\sigma(\hat{t}_\ell)}(j+1)|}{\|\bar{\psi}(j)\|} \\
& = c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell), \hspace{1cm} t \in [\hat{t}_\ell, \hat{t}_{\ell+1}), \ \ell \in \mathbb{Z}^+.
\end{aligned}
\tag{44}
$$

To proceed, we need a bound on $\eta_i(t)$.

Claim 1. *There exists a c_3 such that for all $i \in \mathcal{I}^*$:*

$$
|\eta_i(t)| \leq c_3 |\bar{w}(t)|, \hspace{1cm} t \geq t_0.
\tag{45}
$$

Proof of Claim 1. If $\rho_i(t) = 1$, then $\eta_i(t) = 0$. If $\rho_i(t) = 0$, then $\eta_i(t) = e_i(t+1)$ and from the estimator definition

$$
|e_i(t+1)| \geq (2\bar{\mathbf{s}} + \delta)\|\bar{\psi}(t)\|;
$$

but notice that substituting (11) into (13) yields

$$
\begin{aligned}
& e_i(t+1) = \psi_{i^*}(t)^\top \theta^* - \psi_i(t)^\top \hat{\theta}_i(t) + \bar{w}(t) \\
& \Rightarrow |e_i(t+1)| \leq 2\bar{\mathbf{s}}\|\bar{\psi}(t)\| + |\bar{w}(t)|.
\end{aligned}
$$

Combining the above two statements:

$$
2\bar{\mathbf{s}}\|\bar{\psi}(t)\| + |\bar{w}(t)| \geq (2\bar{\mathbf{s}} + \delta)\|\bar{\psi}(t)\| \Rightarrow \|\bar{\psi}(t)\| \leq \frac{1}{\delta}|\bar{w}(t)|;
$$

this means that $|e_i(t+1)| \leq \frac{2\bar{\mathbf{s}}}{\delta}|\bar{w}(t)| + |\bar{w}(t)|$, so define $c_3 := \frac{2\bar{\mathbf{s}}}{\delta} + 1$. $\qquad\square$

Now we return to analyzing the key equation (40). Solving for $\bar{\psi}(\hat{t}_{\ell+1})$ yields

$$
\begin{aligned}
\bar{\psi}(\hat{t}_{\ell+1}) = {} & \left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right]^{\hat{t}_{\ell+1}-\hat{t}_\ell} \bar{\psi}(\hat{t}_\ell) + \\
& \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right]^{\hat{t}_{\ell+1}-j-1} \left(\left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(j) - \mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) + \right. \right. \\
& \left. \left. \Delta_{\sigma(\hat{t}_\ell)}(j) \right] \bar{\psi}(j) + B_1 \eta_{\sigma(\hat{t}_\ell)}(j) \right).
\end{aligned}
\tag{46}
$$

It follows from the compactness of the $\mathcal{S}_i$'s that $\left\| \left[\mathcal{A}_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right]^j \right\|$, $j = 0, 1 \ldots, N-1$, is bounded above by a constant, which we label c_4. So incorporating this and the observations of (41), (44) and (45) into (46), we obtain

$$
\begin{aligned}
\|\bar{\psi}(\hat{t}_{\ell+1})\| & \leq c_4 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \left(c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\|\bar{\psi}(j)\| + c_3|\bar{w}(j)| \right) \\
& = c_4 c_2 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \|\bar{\psi}(j)\| + c_4 c_3 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} |\bar{w}(j)|.
\end{aligned}
\tag{47}
$$

It follows from Lemma 2 (applied for $p = 1, 2, \ldots, N-1$) that there exists a constant c_5 so that the following holds:

$$
\sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} \|\bar{\psi}(j)\| \leq c_5 \|\bar{\psi}(\hat{t}_\ell)\| + c_5 \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-2} |\bar{w}(j)|;
\tag{48}
$$

so substituting (48) into (47) it follows that there exists a constant c_6 so that for all $\ell \in \mathbb{Z}^+$:

$$
\begin{aligned}
\|\bar{\psi}(\hat{t}_{\ell+1})\| \leq {} & c_6 J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell)\|\bar{\psi}(\hat{t}_\ell)\| + \\
& c_6 \left(1 + J_{\sigma(\hat{t}_\ell)}(\hat{t}_\ell) \right) \sum_{j=\hat{t}_\ell}^{\hat{t}_{\ell+1}-1} |\bar{w}(j)|.
\end{aligned}
\tag{49}
$$

Step 3: Obtain a bound on $\bar{\psi}$ between index set reset times which depend solely on J_{i^*}.

Let $\hat{t}_{\ell_j}$ be an arbitrary reset time. From Lemma 1 we know that there exists an ℓ^* satisfying $\ell_j \leq \ell^* < \ell_{j+1}$ such that

$$
J_{\sigma(\hat{t}_{\ell^*})}(\hat{t}_{\ell^*}) \leq J_{i^*}(\hat{t}_{\ell^*}).
\tag{50}
$$

So here we will analyze the closed-loop behavior for

$$
[\hat{t}_{\ell_j}, \hat{t}_{\ell_{j+1}}) = [\hat{t}_{\ell_j}, \hat{t}_{\ell^*}) \cup [\hat{t}_{\ell^*}, \hat{t}_{\ell^*+1}) \cup [\hat{t}_{\ell^*+1}, \hat{t}_{\ell_{j+1}}).
$$

The behavior on $[\hat{t}_{\ell^*}, \hat{t}_{\ell^*+1})$ can be analyzed by combining (49) with (50):

$$
\begin{aligned}
\|\bar{\psi}(\hat{t}_{\ell^*+1})\| \leq {} & c_6 J_{i^*}(\hat{t}_{\ell^*})\|\bar{\psi}(\hat{t}_{\ell^*})\| + \\
& c_6 \left(1 + J_{i^*}(\hat{t}_{\ell^*}) \right) \sum_{j=\hat{t}_{\ell^*}}^{\hat{t}_{\ell^*+1}-1} |\bar{w}(j)|.
\end{aligned}
\tag{51}
$$

The behavior for the other two intervals can be analyzed by utilizing Lemma 2. To this end, from the switching algorithm, it is clear that

$$
\ell_{j+1} - \ell_j \leq m, \hspace{1cm} j \in \mathbb{Z}^+;
$$

so $\hat{t}_{\ell_{j+1}} - \hat{t}_{\ell_j} \leq Nm$, which means that $\hat{t}_{\ell^*} - \hat{t}_{\ell_j} \leq Nm$

and $\hat{t}_{\ell_{j+1}} - \hat{t}_{\ell^*+1} \leq Nm$ as well; so we can utilize Lemma 2 with $p \leq Nm$: in particular, there exists a constant c_7 so that $\|\bar{\psi}(\hat{t}_{\ell_{j+1}})\| \leq c_7\|\bar{\psi}(\hat{t}_{\ell^*+1})\| + c_7 \sum_{j=\hat{t}_{\ell^*+1}}^{\hat{t}_{\ell_{j+1}}-1} |\bar{w}(j)|$ and $\|\bar{\psi}(\hat{t}_{\ell^*})\| \leq c_7\|\bar{\psi}(\hat{t}_{\ell_j})\| + c_7 \sum_{j=\hat{t}_{\ell_j}}^{\hat{t}_{\ell^*}-1} |\bar{w}(j)|$. Now define

$$\alpha(\ell_j) := \max_{\tau \in [\ell_j, \ell_{j+1})} J_{i^*}(\hat{t}_\tau) \tag{52}$$

and $\tilde{w}(j) := \sum_{q=\hat{t}_{\ell_j}}^{\hat{t}_{\ell_{j+1}}-1} |\bar{w}(q)|$; by combining these bounds with (51) we conclude that there exists a constant c_8 such that

$$\|\bar{\psi}(\hat{t}_{\ell_{j+1}})\| \leq c_8\alpha(\ell_j)\|\bar{\psi}(\hat{t}_{\ell_j})\| + c_8(1 + \alpha(\ell_j))\tilde{w}(j), \quad j \in \mathbb{Z}^+. \tag{53}$$

Step 4: Analyze the first-order difference inequality (53).

We will analyze (53) to obtain a bound on the closed-loop behavior of the system for the whole time horizon. The first step is to analyze the square sum of $\alpha(\cdot)$ over an interval; from the definition of $\alpha(\cdot)$ and the Cauchy-Schwarz property, we have for all $j_2 > j_1 \geq 0$:

$$\sum_{q=j_1}^{j_2-1} \alpha(\ell_q)^2 = \sum_{q=j_1}^{j_2-1} \left(\max_{p \in [\ell_q, \ell_{q+1})} J_{i^*}(\hat{t}_p) \right)^2$$

$$\leq \sum_{q=j_1}^{j_2-1} \left(\sum_{p=\ell_q}^{\ell_{q+1}-1} J_{i^*}(\hat{t}_p) \right)^2$$

$$\leq m \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} J_{i^*}(\hat{t}_p)^2$$

$$= m \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} \left(\sum_{\tau=\hat{t}_p}^{\hat{t}_{p+1}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|}{\|\bar{\psi}(\tau)\|} \right)^2$$

$$\leq Nm \sum_{q=j_1}^{j_2-1} \sum_{p=\ell_q}^{\ell_{q+1}-1} \sum_{\tau=\hat{t}_p}^{\hat{t}_{p+1}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|^2}{\|\bar{\psi}(\tau)\|^2}$$

$$= Nm \sum_{\tau=\hat{t}_{\ell_{j_1}}}^{\hat{t}_{\ell_{j_2}}-1} \rho_{i^*}(\tau) \frac{|e_{i^*}(\tau+1)|^2}{\|\bar{\psi}(\tau)\|^2}.$$

Using the above together with part 2) of Proposition 2, we obtain, for any $p > q \geq 0$:

$$\sum_{j=q}^{p-1} \alpha(\ell_j)^2 \leq 2Nm\|\tilde{\theta}_{i^*}(\hat{t}_{\ell_q})\|^2 + 4Nm \sum_{j=\hat{t}_{\ell_q}}^{\hat{t}_{\ell_p}-1} \rho_{i^*}(j) \frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2}. \tag{54}$$

To proceed, let $\tau \geq t_0$ be arbitrary. We define $\lambda_1 := \frac{\lambda^{Nm}}{\max\{1, c_8\}} \in (0, 1)$. We now partition the timeline into two parts: one in which $\bar{w}(\cdot)$ is small versus $\bar{\psi}(\cdot)$ and one where it is not; with

$$\nu := \left(\frac{\lambda_1}{4Nm} \right)^2, \tag{55}$$

we define

$$S_{good} = \left\{ j \geq \tau : \bar{\psi}(j) \neq 0 \text{ and } \frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} < \nu \right\},$$

$$S_{bad} = \left\{ j \geq \tau : \bar{\psi}(j) = 0 \text{ or } \frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} \geq \nu \right\};$$

clearly $\{j \in \mathbb{Z} : j \geq \tau\} = S_{good} \cup S_{bad}$. Notice that if $\bar{w} = 0$, then S_{good} could be the whole timeline $[\tau, \infty)$. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ which satisfy:
(i) $k_0 = \tau$ serves as the initial instant of the first interval;
(ii) $[k_l, k_{l+1})$ either belongs to S_{good} or S_{bad}; and
(iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{good} then $[k_{l+1}, k_{l+2})$ belongs to S_{bad}, and vice versa.

Now we analyze the behavior during each interval.
Step 4.1: $[k_l, k_{l+1}) \subset S_{bad}$.

Let $j \in [k_l, k_{l+1})$ be arbitrary. In this case $\frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} \geq \nu$ or $\|\bar{\psi}(j)\| = 0$; in either case

$$\|\bar{\psi}(j)\| \leq \frac{1}{\sqrt{\nu}} |\bar{w}(j)|.$$

Also, applying Lemma 2 for one step, there exists a constant c_9 so that

$$\|\bar{\psi}(j+1)\| \leq c_9\|\bar{\psi}(j)\| + c_9|\bar{w}(j)|$$
$$\leq c_9\frac{1}{\sqrt{\nu}}|\bar{w}(j)| + c_9|\bar{w}(j)|, \quad j \in [k_l, k_{l+1}).$$

This, in turn, implies that

$$\|\bar{\psi}(j)\| \leq \begin{cases} \frac{1}{\sqrt{\nu}}|\bar{w}(j)| & j = k_l \\ c_9(\frac{1}{\sqrt{\nu}} + 1)|\bar{w}(j-1)| & j = k_l + 1, \ldots, k_{l+1}. \end{cases} \tag{56}$$

Step 4.2: $[k_l, k_{l+1}) \subset S_{good}$.

First suppose that $k_{l+1} - k_l \leq 2Nm$; then by Lemma 2 it can be easily proven that there exists a constant c_{10} so that

$$\|\bar{\psi}(t)\| \leq c_{10}\lambda^{t-k_l}\|\bar{\psi}(k_l)\| + c_{10} \sum_{j=k_l}^{t-1} \lambda^{t-j-1}|\bar{w}(j)|, \ t \in [k_l, k_{l+1}].$$

Now suppose that $k_{l+1} - k_l > 2Nm$. This means that in the interval of interest, namely $[k_l, k_{l+1})$, there are at least two reset times: there exist $q < p$ so that

$$k_l \leq \hat{t}_{\ell_q} < \hat{t}_{\ell_p} \leq k_{l+1};$$

in fact, there may be many choices of q and p; so let $\underline{q}$ be the smallest such q and $\bar{p}$ be the largest such p. To proceed, observe that $\|\bar{\psi}(j)\| \neq 0$ and $\frac{|\bar{w}(j)|^2}{\|\bar{\psi}(j)\|^2} < \nu$ for $j \in [k_l, k_{l+1})$. Using this bound which holds on $[k_l, k_{l+1})$, together with the fact that $\|\tilde{\theta}_{i^*}(\hat{t}_{\ell_q})\| \leq 2\|\mathcal{S}_{i^*}\| \leq 2\bar{s}$, we rewrite (54) to yield

$$\sum_{j=q}^{p-1} \alpha(\ell_j)^2 \leq 8Nm\bar{s}^2 + 4Nm(\hat{t}_{\ell_p} - \hat{t}_{\ell_q})\nu$$

$$= 8Nm\bar{s}^2 + 4N^2m(\ell_p - \ell_q)\nu$$
$$\leq 8Nm\bar{s}^2 + 4N^2m^2(p - q)\nu, \quad \underline{q} \leq q < p \leq \bar{p}. \tag{57}$$

From the definition of ν in (55), the above bound can be simplified to

$$\sum_{j=q}^{p-1} \alpha(\ell_j)^2 \leq 8Nm\bar{s}^2 + (p - q)\frac{\lambda_1^2}{4}, \quad \underline{q} \leq q < p \leq \bar{p}. \tag{58}$$

Now we will analyze the difference inequality in (53). First, we use (58) to bound the second occurrence of α in (53); from (58) we see that $\alpha(\ell_j) \leq \sqrt{8Nm\bar{s}^2 + \lambda_1^2/4} =: c_{11}$, $\underline{q} \leq j \leq \bar{p}$. So we can rewrite (53) to yield

$$\|\bar{\psi}(\hat{t}_{\ell_{j+1}})\| \leq c_8 \alpha(\ell_j)\|\bar{\psi}(\hat{t}_{\ell_j})\| + \underbrace{c_8(1+c_{11})}_{=:c_{12}}\tilde{w}(j). \quad (59)$$

We now proceed to solve the above difference inequality; we will utilize the "inequality of arithmetic and geometric means", or the AM-GM inequality in short.

Claim 2. *There exists a constant $\gamma_1 > 1$ such that*

$$\prod_{j=q}^{p-1} \alpha(\ell_j) \leq \gamma_1 \lambda_1^{p-q}, \qquad \underline{q} \leq q < p \leq \bar{p}. \quad (60)$$

Proof of Claim 2. Let $q, p \in \mathbb{Z}^+$ be arbitrary such that $\underline{q} \leq q < p \leq \bar{p}$. By the AM-GM inequality and the fact that $\alpha(\ell_j) \geq 0$, we obtain

$$\prod_{j=q}^{p-1} \alpha(\ell_j) \leq \left[\frac{1}{p-q}\sum_{j=q}^{p-1}\alpha(\ell_j)^2\right]^{\frac{p-q}{2}}.$$

Using (58) we obtain

$$\prod_{j=q}^{p-1} \alpha(\ell_j) \leq \left[\frac{8Nm\bar{s}^2}{p-q} + \frac{\lambda_1^2}{4}\right]^{\frac{p-q}{2}}.$$

So it is enough to prove that there exists a constant γ_1 so that

$$\left(\underbrace{\left[\frac{8Nm\bar{s}^2}{j} + \frac{\lambda_1^2}{4}\right]^{\frac{1}{2}}}_{=:\beta(j)}\right)^j \leq \gamma_1\lambda_1^j, \quad j > 0.$$

We can easily show that with $\bar{j} := \left\lceil\left(\frac{\bar{s}}{\lambda_1}\right)^2\right\rceil \times 16Nm$, we have

$$\frac{8Nm\bar{s}^2}{\bar{j}} \leq \frac{\lambda_1^2}{2},$$

which means that

$$\beta(j)^j \leq \lambda_1^j \leq 1, \qquad j \geq \bar{j}.$$

Since $\beta(j)$ decreases as $j \geq 1$ increases, we conclude that if we define $\gamma_1 := \max\left\{1, \left(\frac{\beta(1)}{\lambda_1}\right)^{\bar{j}}\right\}$, then

$$\beta(j)^j \leq \gamma_1\lambda_1^j, \qquad j = 1, 2, \ldots, \bar{j},$$

as well, so the claim holds. $\qquad\square$

Using the bound in (60) and the definition of λ_1 we obtain

$$\prod_{j=q}^{p-1}[c_8\alpha(\ell_j)] \leq \gamma_1\lambda_1^{p-q}c_8^{p-q},$$

$$\leq \gamma_1\lambda^{Nm(p-q)}, \qquad \underline{q} \leq q < p \leq \bar{p}. \quad (61)$$

We can now proceed to solve (59) iteratively; if we use the bound in (61), we see that

$$\|\bar{\psi}(\hat{t}_{\ell_p})\| \leq \gamma_1\lambda^{Nm(p-q)}\|\bar{\psi}(\hat{t}_{\ell_q})\|+$$

$$\sum_{j=q}^{p-1}\gamma_1 c_{12}\left(\lambda^{Nm}\right)^{p-j-1}\tilde{w}(j), \qquad \underline{q} \leq q < p \leq \bar{p}.$$

We can now use Lemma 2 (for no more than Nm steps at a time):

- to provide a bound on $\|\bar{\psi}(t)\|$ between consecutive index set reset times, i.e. between $\hat{t}_{\ell_j}$ and $\hat{t}_{\ell_{j+1}}$;
- to provide a bound on $\|\bar{\psi}(t)\|$ on the beginning part of the interval $[k_l, k_{l+1})$, until we get to the first admissible index set reset time $\hat{t}_{\ell_{\underline{q}}}$;
- to provide a bound on $\|\bar{\psi}(t)\|$ on the last part of the interval $[k_l, k_{l+1})$, after the last admissible index set reset time $\hat{t}_{\ell_{\bar{p}}}$.

After simplification, we conclude that there exists a constant $\gamma_2 \geq c_{10}$ so that

$$\|\bar{\psi}(t)\| \leq \gamma_2\lambda^{t-k_l}\|\bar{\psi}(k_l)\| + \gamma_2\sum_{j=k_l}^{t-1}\lambda^{t-j-1}|\bar{w}(j)|,$$

$$t \in [k_l, k_{l+1}]. \quad (62)$$

Step 4.3: Combining the bounds on S_{good} and S_{bad}.

Now we combine Step 4.1 and Step 4.2 into a general bound on $\bar{\psi}$: we glue the bounds of Step 4.1 and Step 4.2 together. Define $\bar{\gamma} := \max\{\gamma_2, c_9(1 + \frac{1}{\sqrt{\nu}}), \gamma_2 c_9(1 + \frac{1}{\sqrt{\nu}})\}$.

Claim 3. *The following bound holds:*

$$\|\bar{\psi}(t)\| \leq \bar{\gamma}\lambda^{t-\tau}\|\bar{\psi}(\tau)\| + \sum_{j=\tau}^{t-1}\bar{\gamma}\lambda^{t-j-1}|\bar{w}(j)|, \ \ t \geq \tau. \quad (63)$$

Proof of the Claim 3. If $[k_0, k_1) = [\tau, k_1) \subset S_{good}$, then (63) is true for $t \in [k_0, k_1]$ by (62). If $[k_0, k_1) \subset S_{bad}$, then from (56) we obtain

$$\|\bar{\psi}(j)\| \leq \begin{cases} \|\bar{\psi}(\tau)\| & j = k_0 = \tau \\ c_9(1 + \frac{1}{\sqrt{\nu}})|\bar{w}(j-1)| & j = k_0 + 1, \ldots, k_1. \end{cases}$$

which means that (63) holds on $[k_0, k_1]$ for this case as well.

We now use induction: suppose that (63) is true for $t \in [k_0, k_l]$; we need to prove it holds for $t \in (k_l, k_{l+1}]$ as well. If $k \in [k_l, k_{l+1}) \subset S_{bad}$, then from (56) we see that

$$\|\bar{\psi}(j)\| \leq c_9(1 + \frac{1}{\sqrt{\nu}})|\bar{w}(j-1)|, \ j = k_l + 1, k_l + 2, \ldots, k_{l+1},$$

which means (63) holds on $(k_l, k_{l+1}]$. On the other hand, if $[k_l, k_{l+1}) \subset S_{good}$, then $k_l - 1 \in S_{bad}$; from (56) we have that

$$\|\psi(k_l)\| \leq c_9(1 + \frac{1}{\sqrt{\nu}})|\bar{w}(k_l - 1)|.$$

Using (62) to analyze the behavior on $[k_l, k_{l+1}]$, we have

$$\|\bar{\psi}(k)\| \leq \gamma_2\lambda^{k-k_l}[c_9(1 + \frac{1}{\sqrt{\nu}})|\bar{w}(k_l - 1)|]$$

$$+ \sum_{j=k_l}^{k-1}\gamma_2\lambda^{k-j-1}|\bar{w}(j)|,$$

$$\leq \bar{\gamma}\sum_{j=k_l-1}^{k-1}\lambda^{k-j-1}|\bar{w}(j)|, \qquad k \in [k_l, k_{l+1}], \quad (64)$$

which implies that (63) holds. $\qquad\square$

Finally, as $\tau \geq t_0$ is arbitrary, it follows that the proof of Proposition 3 is concluded. $\qquad\blacksquare$

Proof of Theorem 1 (Outline). A brief summary of the proof of Theorem 1 is as follows. From the Proposition 3, we directly get a bound on the y's and prove exponential tracking. We obtain a bound on the u's based on the observability of the augmentation of plant and (26). Finally we simplify and get the desired bound (35) on $\bar{\phi}(t)$. The details can be found in the Appendix. ∎

V. ROBUSTNESS RESULTS

It turns out that the exponential stability property and the convolution bounds proven in the main result of this paper will guarantee robustness to a degree of time-variations and unmodelled dynamics. In this way, the approach has a lot in common with LTI systems, which also enjoys this feature. Indeed, we have recently proven in [46] that this is true even in a more general situation than one considered here; the results there are a generalization of the robustness results in [33]. Here, we briefly discuss the setup and provide the key result, while we refer you to that paper for full details of the proof. We will show that if the time-variations are slow enough and the size of the unmodelled dynamics is small enough, then the closed-loop system retains the desired linear-like properties.

To proceed, we consider a time-varying version of the plant (1). In order to apply [46], we would like the plant model to incorporate the vector $\bar{\phi}(t)$ regardless of the value of $n \in \{1, 2, \ldots, \bar{n}\}$, so we will pad θ with zeros in the obvious spots and then write the time-varying version of plant (1) as

$$y(t+1) = \overline{\theta}(t)^\top \bar{\phi}(t) + w(t), \quad t \in \mathbb{Z}; \tag{65}$$

we define $\overline{\mathbf{\Theta}}_n \subset \mathbb{R}^{2(\bar{n}+g)}$ to represent the padded elements of $\mathbf{\Theta}_n$, and define $\overline{\mathbf{\Theta}} := \cup_{n=1}^{\bar{n}} \overline{\mathbf{\Theta}}_n$, which is clearly compact. We adopt a common model of acceptable time-variations used in adaptive control (e.g. see [25]).

Definition 1. For c_0 and $\epsilon \geq 0$, let $\mathcal{S}\left(\overline{\mathbf{\Theta}}, c_0, \epsilon\right)$ denote the subset of $\ell_\infty\left(\mathbb{R}^{2(\bar{n}+g)}\right)$ whose elements $\overline{\theta}$ satisfy $\overline{\theta}(t) \in \overline{\mathbf{\Theta}}$ for every $t \in \mathbb{Z}$ and

$$\sum_{t=t_1}^{t_2-1} \|\overline{\theta}(t+1) - \overline{\theta}(t)\| \leq c_0 + \epsilon(t_2 - t_1), \quad t_2 > t_1, \ t_1 \in \mathbb{Z}.$$

Remark 6. *The above model encompasses both slow-drift variations ($c_0 = 0$) and/or occasional jumps.*

We now consider the time-varying plant (65) with the term $w_\Delta(t) \in \mathbb{R}$ added to represent the unmodelled dynamics:

$$y(t+1) = \overline{\theta}(t)^\top \bar{\phi}(t) + w(t) + w_\Delta(t), \quad t \in \mathbb{Z}. \tag{66}$$

We also adopt a common model of unmodelled dynamics used in adaptive control: with $\beta \in (0, 1)$ and $\mu > 0$,

$$\mathfrak{w}(t+1) = \beta\mathfrak{w}(t) + \beta\|\bar{\phi}(t)\|, \ \mathfrak{w}(t_0) = \mathfrak{w}_0 \tag{67a}$$

$$|w_\Delta(t)| \leq \mu\mathfrak{w}(t) + \mu\|\bar{\phi}(t)\|; \tag{67b}$$

see [33] for a more detailed explanation about this model.

Theorem 2. *For every $\delta \in (0, \infty]$, $N \geq 2\bar{n}+g$, $\beta \in (0, 1)$ and $\bar{c}_0 \geq 0$, there exist $\bar{\epsilon} > 0$, $\mu > 0$, $\tilde{\lambda} \in (\beta, 1)$ and $\tilde{\gamma} > 0$ such that when the adaptive controller (18), (23)–(29) is applied to the time-varying plant (66) with w_Δ satisfying (67), for*
every $t_0 \in \mathbb{Z}$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\overline{\theta} \in \mathcal{S}\left(\overline{\mathbf{\Theta}}, \bar{c}_0, \bar{\epsilon}\right)$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \ell_\infty$, the following holds:

$$\left\|\begin{bmatrix} \bar{\phi}(t) \\ \mathfrak{w}(t) \end{bmatrix}\right\| \leq \tilde{\gamma}\tilde{\lambda}^{t-t_0}\left\|\begin{bmatrix} \bar{\phi}(t_0) \\ \mathfrak{w}_0 \end{bmatrix}\right\| +$$
$$\sum_{j=t_0}^{t-1} \tilde{\gamma}\tilde{\lambda}^{t-j-1}(|w(j)| + |\tilde{y}^*(j+1)|), \ t \geq t_0.$$

Proof. We apply Theorem 1 to prove a convolution bound on the plant (1), and then apply Theorems 1–2 of [46]. ∎

VI. A SIMULATION EXAMPLE

In this example, we will show the efficacy of the proposed approach, mainly in dealing with plant changes, noise and unmodelled dynamics. We have the upper bound on the order of the plant to be $\bar{n} = 2$; consider the following family of plants: (i) first-order plants with an uncertainty set of

$$\mathbf{\Theta}_1 = \left\{\begin{bmatrix} a_1 & b_1 \end{bmatrix}^\top \in \mathbb{R}^2 : a_1 \in [1, \tfrac{3}{2}], b_1 \in [-2, -1] \cup [1, 2]\right\}$$

and (ii) second-order plants with an uncertainty set of

$$\mathbf{\Theta}_2 = \left\{\begin{bmatrix} a_1 & a_2 & b_1 & b_2 \end{bmatrix}^\top \in \mathbb{R}^4 : \right.$$
$$\left. a_1 = \tfrac{3}{2}, a_2 \in \left\{\tfrac{-3}{2}, \tfrac{3}{2}\right\}, b_1 \in [-1, 0], b_2 \in [-5, -3]\right\}.$$

It is obvious that each of the sets above is compact as required; also, the coprimeness requirement is satisfied. You can see that all potential models are unstable. Also, the 2nd-order models are all nonminimum phase.

The goal is to track reference signals of frequency $\frac{\pi}{25}$: so we set $\mathbf{Q}(z^{-1}) = 1 - 2\cos(\frac{\pi}{25})z^{-1} + z^{-2}$, i.e.

$$q_1 = 2\cos(\tfrac{\pi}{25}), \qquad q_2 = -1, \qquad g = 2.$$

Observe that 2nd-order plant models have a real zero that can lie in $[3, \infty)$, i.e. the associated $\mathbf{B}(z^{-1})$ and $\mathbf{Q}(z^{-1})$ are coprime as required. Next, with $n \in \{1, 2\}$ we use the definition in (9) to construct the uncertainty sets of the associated auxiliary plant; clearly these sets are compact. We see that $\tilde{\mathbf{\Theta}}_1$ and $\tilde{\mathbf{\Theta}}_2$ are not convex; the convex hull of each could violate the coprimeness requirement; for example notice that the convex hull of $\tilde{\mathbf{\Theta}}_1$ includes the case of $b_1 = 0$, which corresponds to a non-stabilizable system, violating the coprimeness assumption. For each of $\tilde{\mathbf{\Theta}}_1$ and $\tilde{\mathbf{\Theta}}_2$, we will need a set of compact and convex sets so that their union contain $\tilde{\mathbf{\Theta}}_1$ and $\tilde{\mathbf{\Theta}}_2$ respectively and satisfy the coprimeness requirement. There is a natural choice: define

$$\mathcal{S}_1 = \left\{\begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & \bar{a}_4 & b_1 & b_2 \end{bmatrix}^\top \in \mathbb{R}^6 : \bar{a}_1 = \tfrac{3}{2} + q_1, \right.$$
$$\bar{a}_2 = -1 - \tfrac{3}{2}(q_1 + 1), \bar{a}_3 = \tfrac{3}{2}(q_1 + 1), \bar{a}_4 = -\tfrac{3}{2},$$
$$\left. b_1 \in [-1, 0], b_2 \in [-5, -3]\right\},$$

$$\mathcal{S}_2 = \left\{\begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & \bar{a}_4 & b_1 & b_2 \end{bmatrix}^\top \in \mathbb{R}^6 : \bar{a}_1 = \tfrac{3}{2} + q_1, \right.$$
$$\bar{a}_2 = -1 + \tfrac{3}{2}(1 - q_1), \bar{a}_3 = \tfrac{3}{2}(1 - q_1), \bar{a}_4 = \tfrac{3}{2},$$
$$\left. b_1 \in [-1, 0], b_2 \in [-5, -3]\right\},$$

$$\mathcal{S}_3 = \left\{ \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & b_1 \end{bmatrix}^\top \in \mathbb{R}^4 : \bar{a}_1 \in [1 + q_1, \tfrac{3}{2} + q_1], \right.$$

$$\left. \bar{a}_2 \in [-1 - \tfrac{3q_1}{2}, -(1 + q_1)], \bar{a}_3 \in [1, \tfrac{3}{2}], b_1 \in [-2, -1] \right\},$$

$$\mathcal{S}_4 = \left\{ \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \bar{a}_3 & b_1 \end{bmatrix}^\top \in \mathbb{R}^4 : \bar{a}_1 \in [1 + q_1, \tfrac{3}{2} + q_1], \right.$$

$$\left. \bar{a}_2 \in [-1 - \tfrac{3q_1}{2}, -(1 + q_1)], \bar{a}_3 \in [1, \tfrac{3}{2}], b_1 \in [1, 2] \right\};$$

clearly $\tilde{\Theta}_1 \subset \mathcal{S}_3 \cup \mathcal{S}_4$ and $\tilde{\Theta}_2 \subset \mathcal{S}_1 \cup \mathcal{S}_2$. The auxiliary plant associated with the 2$^{\text{nd}}$-order plant models has potential poles of either complex ones, or real ones of values 2.186 or -0.686, while it has a zero that lies in $[3, \infty)$; this means that the coprimeness requirement is satisfied as well. So we are going to estimate parameters using 4 parallel estimators; however, we see that for $\mathcal{S}_1$ and $\mathcal{S}_2$, each has only one value for parameters $\bar{a}_1, \bar{a}_2, \bar{a}_3$ and $\bar{a}_4$, which means that the estimation of those parameters is trivial.

For this simulation, we set the plant to

$$y(t+1) = \begin{cases} \tfrac{3}{2}\big[y(t) - y(t-1)\big] \\ \quad - \tfrac{3}{4}u(t) - 4u(t-1) + w(t), & t \le 500 \\ \tfrac{5}{4}\big[y(t) + u(t)\big] + w(t), & t > 500. \end{cases}$$

We set the reference signal to $y^*(t) = 2\sin(\tfrac{\pi}{25}t)$ and the noise to

$$w(t) = \begin{cases} 0.05\cos(45t), & 250 \le t < 750 \\ 0 & \text{otherwise.} \end{cases}$$

We also consider some unmodelled dynamics enter the plant; we use a term of the form discussed in Section V:

$$\mathfrak{w}(t+1) = \tfrac{3}{4}\mathfrak{w}(t) + \tfrac{3}{4}\|\bar{\phi}(t)\|, \ \mathfrak{w}(0) = 0$$

$$w_\Delta(t) = \begin{cases} 0 & t < 700 \\ 0.075\mathfrak{w}(t) + 0.075\|\bar{\phi}(t)\| & \text{otherwise.} \end{cases}$$

We apply the proposed controller (18) and (23)–(29); we choose $N = 6$ and $\delta = \infty$. We set the plant initial conditions to $y(0) = y(-1) = y(-2) = y(-3) = 1$ and $u(0) = u(-1) = u(-2) = u(-3) = 0$; we also set

$$\hat{\theta}_1(0) = \begin{bmatrix} \tfrac{3}{2} + q_1 & -1 + \tfrac{3}{2}(1 + q_1) & \tfrac{3}{2}(q_1 + 1) & -\tfrac{3}{2} & -\tfrac{1}{2} & -4.5 \end{bmatrix}^\top,$$

$$\hat{\theta}_2(0) = \begin{bmatrix} \tfrac{3}{2} + q_1 & -1 + \tfrac{3}{2}(1 - q_1) & \tfrac{3}{2}(1 - q_1) & \tfrac{3}{2} & -\tfrac{1}{2} & -4.5 \end{bmatrix}^\top,$$

$$\hat{\theta}_3(0) = \begin{bmatrix} 3 & -3 & 1 & -2 \end{bmatrix}^\top, \hat{\theta}_4(0) = \begin{bmatrix} 3 & -3 & 1 & 2 \end{bmatrix}^\top,$$

and $\sigma_0 = 3$. The results are in Figure 3. We see that the controller provides good tracking performance. The performance worsens temporarily when noise is added, when plant change happens, and when unmodelled dynamics enter the system; however, tracking recovers quickly.

VII. Summary and Conclusion

In this paper, we consider the problem of tracking for a discrete-time plant with unknown order; we assume knowledge of an upper bound on the order, and the uncertainty set of parameters for each admissible order must lie in a compact set, subject to a coprimeness requirement. Rather than directly estimating the plant parameters, we instead estimate the parameters of a suitably defined auxiliary plant model. We use compactness to prove that for each admissible order, the uncertainty set is contained in a finite union of convex sets;
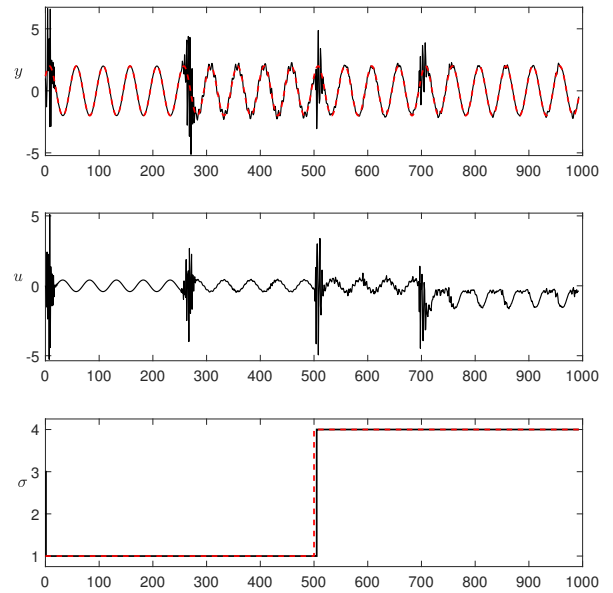

Fig. 3. The upper plot shows both the reference (dashed) and the output (solid); the next plot shows the plant control input; the bottom plot shows the switching signal (solid) and the correct index (dashed).

we use a projection-algorithm based estimator for each convex set. At each point in time, we employ a switching algorithm to determine which model and parameter estimates are used in the control law. We prove that this adaptive controller guarantees very desirable **linear-like closed-loop behavior**: exponential stability, a bounded noise gain in every p-norm, and convolution bounds on the input-output behavior, as well as exponential tracking for certain classes of reference and noise signals. Furthermore, the linear-like properties of the closed-loop behavior is leveraged to show tolerance to a degree of small plant time-variations and unmodelled dynamics. An example of sinusoid tracking is presented to illustrate the results.

We would like to further investigate the performance of the transient behavior which would be helpful for potential design issues. Furthermore, our switching algorithm includes a comparison of behavior over a short interval; a very natural, simpler, switching algorithm, based on instantaneous comparison, works in simulation, and we are striving to show that it works in theory.

Appendix

Proof of Proposition 2. Fix $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta^* \in \tilde{\Theta}_n$; let $t_0 \in \mathbb{Z}$, $t_2 > t_1 \ge t_0$, $\bar{\psi}(t_0) \in \mathbb{R}^{2\bar{n}+g}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ $(i \in \mathcal{I}^*)$ and $w, y^* \in \ell_\infty$ be arbitrary.

For every estimator i, projection does not make the parameter estimate worse; for $t \ge t_0$, it follows from (18) that if $\rho_i(t) = 0$, then $\|\hat{\theta}_i(t+1) - \hat{\theta}_i(t)\| = 0$, and if $\rho_i(t) = 1$, then

$$\|\hat{\theta}_i(t+1) - \hat{\theta}_i(t)\| \le \|\check{\theta}_i(t+1) - \hat{\theta}_i(t)\| \le \frac{|e_i(t+1)|}{\|\bar{\psi}(t)\|}.$$

We conclude that part 1) follows by iteration.

Next, define $\tilde{\theta}_{i^*}(t) := \check{\theta}_{i^*}(t) - \theta^*$; for $t \ge t_0$ when $\rho_{i^*}(t) = 0$, then $\hat{\theta}_{i^*}(t+1) = \hat{\theta}_{i^*}(t)$, so we have

$$\|\tilde{\theta}_{i^*}(t+1)\|^2 = \|\tilde{\theta}_{i^*}(t)\|^2; \tag{69}$$

on the other hand, when $\rho_{i*}(t) = 1$, we have

$$\check{\tilde{\theta}}_{i*}(t+1) = \tilde{\theta}_{i*}(t) + \frac{\psi_{i*}(t)}{\|\bar{\psi}(t)\|^2} e_{i*}(t+1)$$

$$\Rightarrow \|\check{\tilde{\theta}}_{i*}(t+1)\|^2 = \|\tilde{\theta}_{i*}(t)\|^2 + \frac{\|\psi_{i*}(t)\|^2 |e_{i*}(t+1)|^2}{\|\bar{\psi}(t)\|^4} +$$

$$2\frac{\tilde{\theta}_{i*}(t)^\top \psi_{i*}(t)e_{i*}(t+1)}{\|\bar{\psi}(t)\|^2}$$

$$\leq \|\tilde{\theta}_{i*}(t)\|^2 + \frac{|e_{i*}(t+1)|^2}{\|\bar{\psi}(t)\|^2} + 2\frac{\tilde{\theta}_{i*}(t)^\top \psi_{i*}(t)e_{i*}(t+1)}{\|\bar{\psi}(t)\|^2}.$$

$$(70)$$

Using (16) to obtain a representation for $\tilde{\theta}_{i*}(t)^\top \psi_{i*}(t)$ in (70) we obtain

$$\|\check{\tilde{\theta}}_{i*}(t+1)\|^2 \leq \|\tilde{\theta}_{i*}(t)\|^2 - \frac{|e_{i*}(t+1)|^2}{\|\bar{\psi}(t)\|^2} + \frac{2\bar{w}(t)e_{i*}(t+1)}{\|\bar{\psi}(t)\|^2}$$

$$\leq \|\tilde{\theta}_{i*}(t)\|^2 - \frac{|e_{i*}(t+1)|^2}{2\|\bar{\psi}(t)\|^2} + \frac{2\bar{w}(t)^2}{\|\bar{\psi}(t)\|^2}.$$

Since projection does not make the parameter estimate worse, it follows that

$$\|\tilde{\theta}_{i*}(t+1)\|^2 \leq \|\tilde{\theta}_{i*}(t)\|^2 - \frac{|e_{i*}(t+1)|^2}{2\|\bar{\psi}(t)\|^2} + \frac{2\bar{w}(t)^2}{\|\bar{\psi}(t)\|^2}. \quad (71)$$

If we combine the above bound for the case of $\rho_{i*}(t) = 1$ with (69) for the case when $\rho_{i*}(t) = 0$, and iterate, then we obtain (20). ∎

Proof of Lemma 2. Fix $p \geq 0$. Let $n \in \{1, 2, \ldots, \bar{n}\}$ and $\theta \in \Theta_n$, $t_0 \in \mathbb{Z}$, $t \geq t_0$, $N \geq 1$, $\sigma_0 \in \mathcal{I}^*$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$) and $w, y^* \in \ell_\infty$ be arbitrary.

From the associated auxiliary plant (10) we see that $|\varepsilon(t+1)| \leq \|\theta^*\|\|\psi(t)\| + |\bar{w}(t)| \leq \bar{s}\|\bar{\psi}(t)\| + |\bar{w}(t)|$. From (25) and compactness, we have that there exists a constant γ_1 so that $|v(t+1)| \leq \gamma_1 \|\bar{\psi}(t)\|$. From the definition of $\|\bar{\psi}(t+1)\|$ given in (15), we have that $\|\bar{\psi}(t+1)\| \leq \|\bar{\psi}(t)\| + |\varepsilon(t+1)| + |v(t+1)|$. Combining these three bounds, we end up with

$$\|\bar{\psi}(t+1)\| \leq (1 + \bar{s} + \gamma_1)\|\bar{\psi}(t)\| + |\bar{w}(t)|.$$

We solve this iteratively for p steps and put $\bar{c}_1 := (1 + \bar{s} + \gamma_1)^p$ to conclude the proof of part (i).

The proof of part (ii) is similar; we omit the details for space considerations. ∎

Proof of Theorem 1. Fix $\lambda \in (0, 1)$, $\delta \in (0, \infty]$ and $N \geq 2\bar{n} + g$. Let $n \in \{1, 2, \ldots, \bar{n}\}$, $\theta \in \Theta_n$, $t_0 \in \mathbb{Z}$, $\bar{\phi}(t_0) \in \mathbb{R}^{2(\bar{n}+g)}$, $\sigma_0 \in \mathcal{I}^*$, $\hat{\theta}_i(t_0) \in \mathcal{S}_i$ ($i \in \mathcal{I}^*$), and $w, y^* \in \ell_\infty$ be arbitrary.

Let $\tau \geq t_0$ be arbitrary. Then applying the adaptive controller to the associated auxiliary plant (10), by Proposition 3 there exists a constant c such that

$$\|\bar{\psi}(t)\| \leq c\lambda^{t-\tau}\|\bar{\psi}(\tau)\| + \sum_{j=\tau}^{t-1} c\lambda^{t-j-1}|\bar{w}(j)|, \quad t > \tau. \quad (72)$$

Step 1: Finding a bound on $y(\cdot)$.

It turns out to be easy to leverage (72) to provide first a desired bound on the output y and its past values. Using the definition of $\bar{\psi}$ given in (15), the definition of $\tilde{y}^*$ given in (34), and the fact that $y(t) = \varepsilon(t) + y^*(t)$, with a change in the indexes it follows from (72) that

$$\left\| \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-\bar{n}-g+1) \end{bmatrix} \right\| \leq c\lambda^{t-\tau-\bar{n}-g}\|\bar{\psi}(\tau+\bar{n}+g)\| +$$

$$|\tilde{y}^*(t)| + \sum_{j=\tau+\bar{n}+g}^{t-1} c\lambda^{t-j-1}|\bar{w}(j)|, \quad t > \tau + \bar{n} + g; \quad (73)$$

the reason for choosing a starting time of $\tau + \bar{n} + g$ rather than τ will become clear later in the proof. Although this provides a bound on the top part of $\bar{\phi}$, the quantity on the RHS differs from that on the RHS of the desired bound (35). We will now proceed to get a similar kind of bound on the bottom part of $\bar{\phi}$, after which we convert the quantity in the RHS to one of the desired form.

Step 2: Finding a bound on $u(\cdot)$.

We start by constructing a state-space model of the plant (1); we will choose one of dimension n which is in controllable canonical form:

$$x(t+1) = Ax(t) + Bu(t) \quad (74a)$$
$$y(t) = Cx(t) + w(t-1). \quad (74b)$$

Corresponding to our coprimeness and compactness assumptions, the set of all such (A, B, C) triples lies in a compact set.

From the plant control input defined in (26), we can view u as the output of the following g^{th}-order system. In fact, with $Q \in \mathbb{R}^{g \times g}$ defined by

$$Q := \begin{bmatrix} q_1 & q_2 & \cdots & q_{g-1} & q_g \\ 1 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad B_c := \mathbf{e}_1, \quad C_c := \mathbf{e}_1^\top,$$

$$\xi(t) := \begin{bmatrix} u(t) & u(t-1) & \cdots & u(t-g+1) \end{bmatrix}^\top,$$

it follows that

$$\xi(t+1) = Q\xi(t) + B_c v(t+1), \quad (75a)$$
$$u(t) = C_c \xi(t). \quad (75b)$$

Since $\varepsilon(t) = y(t) - y^*(t)$, by combining (74) with (75) we obtain the augmented $(n+g)^{\text{th}}$-order state-space system

$$\begin{bmatrix} x(t+1) \\ \xi(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} A & BC_c \\ \mathbf{0}_{g \times n} & Q \end{bmatrix}}_{=:\bar{A}} \underbrace{\begin{bmatrix} x(t) \\ \xi(t) \end{bmatrix}}_{=:\bar{x}(t)} + \underbrace{\begin{bmatrix} 0 \\ B_c \end{bmatrix}}_{=:\bar{B}} v(t+1) \quad (76a)$$

$$\varepsilon(t) = \underbrace{\begin{bmatrix} C & \mathbf{0}_g^\top \end{bmatrix}}_{=:\bar{C}} \begin{bmatrix} x(t) \\ \xi(t) \end{bmatrix} + w(t-1) - y^*(t). \quad (76b)$$

Since (74) is controllable and observable and does not have common zeros with the eigenvalues of Q (Assumption 2), it follows that $(\bar{C}, \bar{A})$ is observable; hence, there exists a unique $\bar{H}$ such that the eigenvalues of $\bar{A} + \bar{H}\bar{C}$ are all zero and it is well-known that $\bar{H}$ is a continuous function of $\bar{A}$ and $\bar{C}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3052745, IEEE Transactions on Automatic Control

SHAHAB AND MILLER : ASYMPTOTIC TRACKING AND LINEAR-LIKE BEHAVIOR USING MULTI-MODEL ADAPTIVE CONTROL 15

Now rewrite (76) as

$$\bar{x}(t+1) = \left[\bar{A} + \bar{H}\bar{C}\right]\bar{x}(t)+$$
$$- \bar{H}\varepsilon(t) + \bar{B}v(t+1) + \bar{H}w(t-1) - \bar{H}y^*(t);$$

noting that $[\bar{A} + \bar{H}\bar{C}]^j = 0$ for all $j \geq n + g$, the solution of the above equation is

$$\bar{x}(t) = \sum_{j=1}^{n+g} \left[\bar{A} + \bar{H}\bar{C}\right]^{j-1} \left(\left[-\bar{H}\varepsilon(t-j) + \bar{B}v(t-j+1)\right] + \right.$$
$$\left. \bar{H}[w(t-j-1) - y^*(t-j)]\right), \qquad t \geq \tau + n + g.$$

We now want to analyze the behavior of $\bar{x}$ in terms of $\bar{\psi}$. But

$$\varepsilon(t-j) = \mathbf{e}_j^\top \bar{\psi}(t-1), \qquad j = 1, 2, \ldots, n+g,$$
$$v(t-j+1) = \begin{cases} \mathbf{e}_{\bar{n}+g+j}^\top \bar{\psi}(t), & j = 1, 2, \ldots, n \\ \mathbf{e}_{2\bar{n}+g}^\top \bar{\psi}(t+n-j), & j = n+1, \ldots, n+g, \end{cases}$$

and $\xi(t)$ is part of $\bar{x}(t)$, then there exists a constant γ_1 so that

$$\|\xi(t)\| \leq \gamma_1 \sum_{j=0}^{g} \|\bar{\psi}(t-j)\| + \gamma_1 \sum_{j=t-n-g}^{t-1} |w(j-1)| +$$
$$\gamma_1 |\tilde{y}^*(t-1)|, \quad t \geq \tau + n + g. \qquad (77)$$

Step 3: Finding a bound on $\bar{\phi}$.

We will use (72) to provide a bound on the $\bar{\psi}(\cdot)$'s on the RHS of (77) in terms of $\bar{\psi}(\tau + \bar{n} + g)$. While $\xi(t)$ contains $u(t), u(t-1), \ldots, u(t-g+1)$, the vector $\bar{\phi}(t)$ contains $u(t), u(t-1), \ldots, u(t-\bar{n}-g+1)$. However, we see that

$$\left\| \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-\bar{n}-g+1) \end{bmatrix} \right\| \leq \sum_{j=0}^{\bar{n}} \|\xi(t-j)\|. \qquad (78)$$

Now we apply the bound in (77) to obtain bounds on $\xi(t), \xi(t-1), \ldots, \xi(t-\bar{n})$; after simplification, we combine the resulting bound with (73) and conclude that there exists a constant γ_2 so that

$$\left\|\bar{\phi}(t)\right\| \leq \gamma_2 \lambda^{t-\tau-\bar{n}-g} \|\bar{\psi}(\tau+\bar{n}+g)\| + \gamma_2 \sum_{j=0}^{\bar{n}+1} |\tilde{y}^*(t-j)| +$$
$$\gamma_2 \sum_{j=t-2\bar{n}-g}^{t-1} |w(j-1)| + \gamma_2 \sum_{j=\tau+\bar{n}+g}^{t-1} \lambda^{t-j-1}|\bar{w}(j)|,$$
$$t \geq \tau + 2\bar{n} + 2g. \qquad (79)$$

The bound in (79) looks very similar to the desired bound, except for the use of $\bar{w}$, the use of $\bar{\psi}$ instead of $\bar{\phi}$, and the starting point. We deal with these issues in that order.

(i) We first replace $\bar{w}(\cdot)$ with its constituent signals: we see from the definition of $\bar{w}$ in (6) that $\bar{w}(j)$ is a weighted sum of $w(j), w(j-1), \ldots, w(j-g)$ and $y^*(j+1), y^*(j), y^*(t-1) \ldots, y^*(j-n-g+1)$.

(ii) Next we want to have a bound on $\|\bar{\psi}(\tau+\bar{n}+g)\|$ in terms of $\|\bar{\phi}(\tau+\bar{n}+g)\|$. From the definition of $\bar{\psi}$ and the definition of the auxiliary input v, we see that $\bar{\psi}(t)$ consists of $y(t) - y^*(t), y(t-1) - y^*(t-1), \ldots, y(t-\bar{n}-g+1) - y^*(t-\bar{n}-g+1)$ and weighted sums of $u(t), u(t-1), \ldots, u(t-\bar{n}-g+1)$; so by the definition of $\bar{\phi}(\cdot)$ we conclude that there exists a

constant γ_3 such that

$$\|\bar{\psi}(t)\| \leq \gamma_3 \|\bar{\phi}(t)\| + \sum_{j=0}^{\bar{n}+g-1} |y^*(t-j)|, \quad t \geq t_0. \qquad (80)$$

(iii) We now obtain a bound on $\|\bar{\phi}(\tau+\bar{n}+g)\|$ in terms of $\|\bar{\phi}(\tau)\|$ by utilizing Lemma 2 with $p = \bar{n} + g$.

If we incorporate the above three observations into (79) and simplify, then we conclude that there exists a constant γ_4 so that

$$\left\|\bar{\phi}(t)\right\| \leq \gamma_4 \lambda^{t-\tau} \|\bar{\phi}(\tau)\| +$$
$$\gamma_4 \sum_{j=\tau}^{t-1} \lambda^{t-j-1}(|w(j)| + |\tilde{y}^*(j+1)|),$$
$$t \geq \tau + 2\bar{n} + 2g + 1. \qquad (81)$$

Last of all, we can apply Lemma 2 directly to obtain the desired bound for $t \in [\tau, \tau + 2\bar{n} + 2g]$, which we can combine it with (81) to conclude that (35) holds.

Step 4: Proving exponential tracking.

Observe that if y^* and w satisfy $\mathbf{Q}(z^{-1})Y^*(z) = 0$ and $\mathbf{Q}(z^{-1})W(z) = 0$, then from the definition of $\bar{W}(z)$ in (6) we have $\bar{w}(t) = 0$. So from (72) and the definition of $\bar{\psi}$, we see that

$$|\varepsilon(t)| \leq \|\bar{\psi}(t)\| \leq c\lambda^{t-t_0}\|\bar{\psi}(t_0)\|, \qquad t \geq t_0.$$

Using (80) to obtain a bound on $\|\bar{\psi}(t_0)\|$, we obtain the desired bound on the tracking error. ∎

REFERENCES

[1] B. D. O. Anderson, T. S. Brinsmead, F. De Bruyne, J. Hespanha, D. Liberzon, and A. S. Morse, "Multiple model adaptive control. part 1: Finite controller coverings," *International Journal of Robust and Nonlinear Control*, vol. 10, no. 11-12, pp. 909–929, 2000.

[2] B. D. O. Anderson and C. R. Johnson, "Exponential convergence of adaptive identification and control algorithms," *Automatica*, vol. 18, no. 1, pp. 1 – 13, 1982.

[3] D. Angeli and E. Mosca, "Lyapunov-based switching supervisory control of nonlinear uncertain systems," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 500–505, 2002.

[4] D. Angeli and E. Mosca, "Adaptive switching supervisory control of nonlinear systems with no prior knowledge of noise bounds," *Automatica*, vol. 40, no. 3, pp. 449 – 457, 2004.

[5] S. Baldi, P. Ioannou, and E. Mosca, "Multiple Model Adaptive Mixing Control: The Discrete-Time Case," *IEEE Trans. Autom. Control*, vol. 57, no. 4, pp. 1040–1045, Apr 2012.

[6] S. Baldi, G. Battistelli, E. Mosca, and P. Tesi, "Multi-model unfalsified adaptive switching supervisory control," *Automatica*, vol. 46, no. 2, pp. 249–259, 2010.

[7] S. Baldi, P. A. Ioannou, and E. B. Kosmatopoulos, "Adaptive mixing control with multiple estimators," *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 8, pp. 800–820, Aug 2012.

[8] G. Battistelli, E. Mosca, M. G. Safonov, and P. Tesi, "Stability of unfalsified adaptive switching control in noisy environments," *IEEE Trans. Autom. Control*, vol. 55, no. 10, pp. 2424–2429, 10 2010.

[9] G. Battistelli, J. P. Hespanha, E. Mosca, and P. Tesi, "Model-Free Adaptive Switching Control of Time-Varying Plants," *IEEE Trans. Autom. Control*, vol. 58, no. 5, pp. 1208–1220, May 2013.

[10] G. Battistelli, E. Mosca, and P. Tesi, "Adaptive memory in multi-model switching control of uncertain plants," *Automatica*, vol. 50, no. 3, pp. 874–882, 2014.

[11] D. Borrelli, A. S. Morse, and E. Mosca, "Discrete-time supervisory control of families of two-degrees-of-freedom linear set-point controllers," *IEEE Trans. Autom. Control*, vol. 44, no. 1, pp. 178–181, Jan 1999.

[12] W. Chen, C. Wen, and J. Wu, "Global exponential/finite-time stability of nonlinear adaptive switching systems with applications in controlling systems with unknown control direction," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2738–2744, 2018.

[13] S. Fekri, M. Athans, and A. Pascoal, "Issues, progress and new results in robust adaptive control," *International Journal of Adaptive Control and Signal Processing*, vol. 20, no. 10, pp. 519–579, Dec 2006.

[14] A. Feuer and A. S. Morse, "Adaptive control of single-input, single-output linear systems," *IEEE Trans. Autom. Control*, vol. 23, no. 4, pp. 557–569, Aug 1978.

[15] M. Fu and B. Barmish, "Adaptive stabilization of linear systems via switching control," *IEEE Trans. Autom. Control*, vol. 31, no. 12, pp. 1097–1103, Dec 1986.

[16] G. C. Goodwin, P. Ramadge, and P. Caines, "Discrete-time multivariable adaptive control," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 449–456, Jun 1980.

[17] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. New York, NY, USA: Dover Publications, Inc., 1984.

[18] Z. Han and K. S. Narendra, "New Concepts in Adaptive Control Using Multiple Models," *IEEE Trans. Autom. Control*, vol. 57, no. 1, pp. 78–89, Jan 2012.

[19] J. P. Hespanha, D. Liberzon, and A. S. Morse, "Hysteresis-based switching algorithms for supervisory control of uncertain systems," *Automatica*, vol. 39, no. 2, pp. 263–272, Feb 2003.

[20] J. P. Hespanha, D. Liberzon, and A. S. Morse, "Overcoming the limitations of adaptive control by means of logic-based switching," *Systems & Control Letters*, vol. 49, no. 1, pp. 49–65, May 2003.

[21] P. A. Ioannou and K. S. Tsakalis, "A robust direct adaptive controller," *IEEE Trans. Autom. Control*, vol. 31, no. 11, pp. 1033–1043, Nov 1986.

[22] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[23] S. Kersting and M. Buss, "How to systematically distribute candidate models and robust controllers in multiple-model adaptive control: A coverage control approach," *IEEE Trans. Autom. Control*, vol. 63, no. 4, pp. 1075–1089, Apr 2018.

[24] G. Kreisselmeier and B. D. O. Anderson, "Robust model reference adaptive control," *IEEE Trans. Autom. Control*, vol. 31, no. 2, pp. 127–133, Feb 1986.

[25] G. Kreisselmeier, "Adaptive control of a class of slowly time-varying plants," *Systems & Control Letters*, vol. 8, no. 2, pp. 97–103, Dec 1986.

[26] M. Kuipers and P. A. Ioannou, "Multiple Model Adaptive Control With Mixing," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1822–1836, Aug 2010.

[27] Y. Li and H.-F. Chen, "Robust adaptive pole placement for linear time-varying systems," *IEEE Trans. Autom. Control*, vol. 41, no. 5, pp. 714–719, May 1996.

[28] R. H. Middleton and G. C. Goodwin, "Adaptive control of time-varying linear systems," *IEEE Trans. Autom. Control*, vol. 33, no. 2, pp. 150–155, 1988.

[29] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne, "Design issues in adaptive control," *IEEE Trans. Autom. Control*, vol. 33, no. 1, pp. 50–58, Jan 1988.

[30] D. E. Miller, "A parameter adaptive controller which provides exponential stability: The first order case," *Systems & Control Letters*, vol. 103, pp. 23–31, May 2017.

[31] D. E. Miller, "Classical discrete-time adaptive control revisited: Exponential stabilization," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, Aug 2017, pp. 1975–1980.

[32] D. E. Miller and E. J. Davison, "An adaptive controller which provides Lyapunov stability," *IEEE Trans. Autom. Control*, vol. 34, no. 6, pp. 599–609, Jun 1989.

[33] D. E. Miller and M. T. Shahab, "Classical pole placement adaptive control revisited: linear-like convolution bounds and exponential stability," *Math. Control Signals Syst*, vol. 30, no. 4, p. 19, Nov 2018.

[34] D. E. Miller and M. T. Shahab, "Classical d-step-ahead adaptive control revisited: Linear-like convolution bounds and exponential stability," in *2019 American Control Conf.*, July 2019, pp. 417–422.

[35] A. S. Morse, "Global stability of parameter-adaptive control systems," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 433–439, Jun 1980.

[36] A. S. Morse, "Supervisory control of families of linear set-point controllers–Part 1: Exact matching," *IEEE Trans. Autom. Control*, vol. 41, no. 10, pp. 1413–1431, 1996.

[37] A. S. Morse, "Supervisory Control of Families of Linear Set-Point Controllers–Part 2: Robustness," *IEEE Trans. Autom. Control*, vol. 42, no. 11, pp. 1500–1515, 1997.

[38] S. M. Naik, P. R. Kumar, and B. E. Ydstie, "Robust continuous-time adaptive control by parameter projection," *IEEE Trans. Autom. Control*, vol. 37, no. 2, pp. 182–197, 1992.

[39] K. Narendra and Y.-H. Lin, "Stable discrete adaptive control," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 456–461, Jun 1980.

[40] K. Narendra, Y.-H. Lin, and L. Valavani, "Stable adaptive controller design, part II: Proof of stability," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 440–448, Jun 1980.

[41] K. S. Narendra and J. Balakrishnan, "Improving transient response of adaptive control systems using multiple models and switching," *IEEE Trans. Autom. Control*, vol. 39, no. 9, pp. 1861–1866, 1994.

[42] K. S. Narendra and J. Balakrishnan, "Adaptive control using multiple models," *IEEE Trans. Autom. Control*, vol. 42, no. 2, pp. 171–187, 1997.

[43] K. Narendra and Z. Han, "A new approach to adaptive control using multiple models," *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 8, pp. 778–799, Aug 2012.

[44] C. Rohrs, L. Valavani, M. Athans, and G. Stein, "Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics," *IEEE Trans. Autom. Control*, vol. 30, no. 9, pp. 881–889, Sep 1985.

[45] M. G. Safonov and T.-C. Tsao, "The unfalsified control concept and learning," *IEEE Trans. Autom. Control*, vol. 42, no. 6, pp. 843–847, Jun 1997.

[46] M. T. Shahab and D. E. Miller, "The inherent robustness of a new approach to adaptive control," in *2020 IEEE Conf. on Control Technology and Applications (CCTA)*, 2020, pp. 510–515.

[47] M. T. Shahab and D. E. Miller, "Multi-estimator based adaptive control which provides exponential stability: The first-order case," in *2018 IEEE Conference on Decision and Control (CDC)*, Dec 2018, pp. 2223–2228.

[48] M. T. Shahab and D. E. Miller, "Adaptive Set-Point Regulation using Multiple Estimators," in *2019 IEEE Conference on Decision and Control (CDC)*, Dec 2019, pp. 84–89.

[49] M. Stefanovic and M. G. Safonov, "Safe Adaptive Switching Control: Stability and Convergence," *IEEE Trans. Autom. Control*, vol. 53, no. 9, pp. 2012–2021, Oct 2008.

[50] L. Vu and D. Liberzon, "Supervisory Control of Uncertain Linear Time-Varying Systems," *IEEE Trans. Autom. Control*, vol. 56, no. 1, pp. 27–42, Jan 2011.

[51] C. Wen, "A robust adaptive controller with minimal modifications for discrete time-varying systems," *IEEE Trans. Autom. Control*, vol. 39, no. 5, pp. 987–991, May 1994.

[52] C. Wen and D. J. Hill, "Global boundedness of discrete-time adaptive control just using estimator projection," *Automatica*, vol. 28, no. 6, pp. 1143–1157, Nov 1992.

[53] B. E. Ydstie, "Transient performance and robustness of direct adaptive control," *IEEE Trans. Autom. Control*, vol. 37, no. 8, pp. 1091–1105, 1992.

[54] P. V. Zhivoglyadov, R. H. Middleton, and M. Fu, "Localization based switching adaptive control for time-varying discrete-time systems," *IEEE Trans. Autom. Control*, vol. 45, no. 4, pp. 752–755, Apr 2000.

Mohamad T. Shahab (S'18) received his Ph.D. degree in Electrical and Computer Engineering from the University of Waterloo, Waterloo, Ontario, Canada in 2020, and his M.Sc. degree in Systems Engineering and B.Sc. degree in Control and Instrumentation Systems Engineering from King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia, in 2007 and 2015, respectively. After earning his Ph.D. he joined, as a Postdoctoral Fellow, King Abdullah University of Science & Technology (KAUST), Thuwal, Saudi Arabia. From Jan 2012 to Dec 2015, he was a Research Engineer with Baker Hughes. His research interests include adaptive control, adaptive switching control, optimization & control, and their applications to robotic systems.

Daniel E. Miller (S'82, M'90, SM'03) received the B.Sc. degree in Electrical Engineering from the University of New Brunswick in 1984, and the M.A.Sc. and Ph.D. degrees in Electrical Engineering from the University of Toronto in 1986 and 1990, respectively. During 1990 to 1991 he held postdoctoral positions at Cambridge University, U.K., the University of Groningen, The Netherlands, and the Massachusetts Institute of Technology. Since 1991 he has been a faculty member in the Department of Electrical and Computer Engineering, University of Waterloo, where he currently holds the rank of Full Professor. He has held visiting positions at the University of Alberta and the University of Newcastle, Australia, and has been an Associate Editor of the *IEEE Transactions on Automatic Control*. His research interests include adaptive control, decentralized control, time-varying control, delay margins, and performance limitations.