Multi-Estimator Based Adaptive Control Which Provides Exponential Stability: The First-Order Case

Mohamad T. Shahab and Daniel E. Miller

The final published version can be found in the proceedings of the IEEE 57th Conference on Decision and Control (CDC 2018). DOI: 10.1109/CDC.2018.8619630

***Abstract*— Classical adaptive controllers provide asymptotic stabilization; neither exponential stability nor a bounded noise gain is typically proven. In recent work it is shown that these desired properties can be achieved by using an estimator based on the original ideal Projection Algorithm (together with a restriction of the parameter estimates to a given compact convex set), rather than the commonly used modified classical algorithm. Here the goal is to remove the convexity requirement. To this end, we consider the first-order case with unknown plant parameters belonging to a compact uncertainty set of controllable pairs. The first step of our approach is to observe that the compact uncertainty set can be covered by a finite number of convex compact sets, each of controllable pairs. For each of the convex compact sets, we design an estimator together with the corresponding one-step-ahead controller, and apply a switching logic to choose between them. We prove that the resulting controller guarantees linear-like convolution bounds on the closed-loop behavior, which implies exponential stability and a bounded noise gain.**

I. Introduction

Adaptive control is an approach used to deal with systems with uncertain or time-varying parameters. The first general results of adaptive control came about around 1980, e.g. [3], [5], [16], [21] and [22]. However, these controllers typically do not tolerate unmodeled dynamics, time-variations, and/or noise/disturbances very well, see e.g. [26]; furthermore, they put stringent assumptions on *a priori* information about the structure of the plant, e.g. the sign of the high-frequency gain. Over the following two decades, there was a great deal of effort to address these shortcomings.

First of all, attempts were made to handle unmodeled dynamics, slow time-variations and noise/disturbance. A common approach was to make small controller design changes, such as σ-modifications, signal normalization, and deadzones, e.g. see [9], [10], and [12]. An alternative powerful approach shows that adaptive controllers incorporating an estimator which uses projection (onto a convex set of admissible parameters) provides desirable properties–see [20], [28], [27] and [30].

Second of all, a great deal of work was focused on removing (or reducing) the assumptions on structural information of the plant, e.g. see [4], [13], [19], [25] and [29]. In all of these results, a pre-routed switching controller was used. A side effect of this is that they typically yield poor transient behavior and/or a large control input signal.

Later, non-pre-routed logic-based switching approaches to adaptive control emerged, such as Supervisory Control [17], [18], [7] and [8]; here the complexity of the approach grows with the size of the plant uncertainty. Other similar approaches are found in the literature: in [23] and [24] the authors use multiple models and argue that the transient behavior improved. In another approach, labelled Adaptive Mixing Control [11], [1] and [2], results are presented which have a tolerance to noise and unmodeled dynamics; a convexity assumption is enforced.

In the recent papers [14] and [15] by the second co-author, an approach is provided which guarantees a linear-like convolution bound on the closed-loop behavior; this yields exponential stability as well as a bounded gain on the noise (in the context of the first-order one-step-ahead control paradigm [14] and the pole-placement stability paradigm [15]). The approach employs a classical design (though the proofs are completely new); namely, the estimator is based on the original ideal Projection Algorithm (together with projecting parameter estimates onto a given compact convex set), rather than the commonly-used modified classical algorithm. The requirement of convexity on the set of uncertainty is shown to play a crucial role in getting nice closed-loop properties.

Since convexity is a very restrictive requirement, the main objective of this paper is to extend [14] by replacing the assumption of a convex and compact uncertainty set with the assumption of a compact uncertainty set only. We first show how to cover the compact set of uncertainty by a finite number of convex sets; then we design an estimator for each convex set, and construct some logic to switch between the corresponding LTI compensators. A key facet of the approach is the switching logic and the use of multi-estimation.

We use standard notation throughout the paper. We denote $\mathbb{Z}$, $\mathbb{Z}^+$ and $\mathbb{N}$ as the sets of integers, non-negative integers and natural numbers, respectively. Let $\lceil\cdot\rceil$ denote the ceiling function. We will denote the Euclidean-norm by the subscript-less default notation $\|\cdot\|$. Also, ℓ_∞ denotes the set of real-valued bounded sequences. If $\Omega \subset \mathbb{R}^p$ is a convex and compact (closed and bounded) set, we define $\|\Omega\| := \max_{x\in\Omega}\|x\|$.

II. The Setup

Here we consider the first order system

$$\begin{aligned} y(k+1) &= \underbrace{\begin{bmatrix} y(k) \\ u(k) \end{bmatrix}^\top}_{=:\phi(k)^\top} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{=:\theta^*} + w(k), \\ y(k_0) &= y_0, \qquad k \geq k_0, \end{aligned} \tag{1}$$

where $y(k) \in \mathbb{R}$ is the state, $u(k) \in \mathbb{R}$ is the control input, $w(k) \in \mathbb{R}$ is the noise (or disturbance). We assume that θ^* is unknown and belongs to a **closed and bounded (compact)** set $\mathcal{S} \subset \mathbb{R}^2$ satisfying a controllability assumption: $\begin{bmatrix} a \\ 0 \end{bmatrix} \notin \mathcal{S}$ for every $a \in \mathbb{R}$. Here we have an exogenous reference signal

M. T. Shahab and D. E. Miller are with Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1. Emails: {m4shahab,miller}@uwaterloo.ca

This research is supported by an NSERC Discovery Grant.

and the control objective is to track it asymptotically while providing a strong notion of closed-loop stability.

As discussed earlier, the property of convexity on the set of uncertainty is shown to play a crucial role in getting nice closed-loop properties. However, here we impose no such assumption. If the set of admissible parameters is not convex, the standard trick in adaptive control is to replace it with its closed convex hull. Unfortunately, often that set contains uncontrollable models (i.e. $b = 0$ in case of 1st-order plants). Here the key idea is to cover the compact set of admissible parameters $\mathcal{S}$ by a finite number of convex sets: the following proposition illustrates that we can always obtain a cover with two convex sets.

Proposition 1. *For any compact set* $\mathcal{S} \subset \left\{\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : b \neq 0\right\}$, *there exist compact and convex sets* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *which also lie in* $\left\{\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : b \neq 0\right\}$ *such that* $\mathcal{S} \subset \mathcal{S}_1 \cup \mathcal{S}_2$.

Proof. For a given $\mathcal{S}$, define

$$\begin{aligned} \mathcal{S}_1 &:= \text{convex hull of } \left\{\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S} : b > 0\right\}, \\ \mathcal{S}_2 &:= \text{convex hull of } \left\{\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S} : b < 0\right\}. \end{aligned}$$

The result follows immediately. ■

Remark. If a convex set is complicated, it may be difficult (numerically) to project onto it. If we define $\bar{a} := \max\left\{|a| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}\right\}$, $\bar{b} := \max\left\{|b| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}\right\}$ and $\underline{b} := \min\left\{|b| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}\right\}$, then Proposition 1 also holds if we define $\mathcal{S}_1 := \left\{\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [-\bar{a}, \bar{a}],\ b \in [\underline{b}, \bar{b}]\right\}$ and $\mathcal{S}_2 := \left\{\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [-\bar{a}, \bar{a}],\ b \in [-\bar{b}, -\underline{b}]\right\}$, which are rectangles.

If $\mathcal{S}_1$ and $\mathcal{S}_2$ are large, it may be beneficial to have more than two, but smaller, convex sets. The usefulness of the above discussion will be obvious when we discuss parameter estimation in the next section. At this point we assume that

$$\mathcal{S} \subset \bigcup_{i=1}^{m} \mathcal{S}_i$$

and each set $\mathcal{S}_i$ is compact and convex and satisfies $\left[\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right] \notin \mathcal{S}_i$ for every $a \in \mathbb{R}$.

A. Parameter Estimation

For each $\mathcal{S}_i$ and $\hat{\theta}_i(k_0) \in \mathcal{S}_i$, we design a Projection Algorithm estimator which generates an estimate $\hat{\theta}_i(k) \in \mathcal{S}_i$ at each $k > k_0$. The associated prediction error is defined as

$$e_i(k+1) = y(k+1) - \phi(k)^\top \hat{\theta}_i(k). \tag{2}$$

The parameter estimation algorithm is as follows: a new parameter estimate is found by solving the optimization problem

$$\operatorname*{argmin}_{\theta} \left\{ \|\theta - \hat{\theta}(k)\| : y(k+1) = \phi(k)^\top \theta \right\}$$

yielding

$$\check{\theta}_i(k+1) = \begin{cases} \hat{\theta}_i(k) & \text{if } \phi(k) = 0 \\ \hat{\theta}_i(k) + \frac{\phi(k)}{\|\phi(k)\|^2} e_i(k+1) & \text{otherwise;} \end{cases} \tag{3}$$

with the function $\operatorname*{Proj}_{\mathcal{S}_i}\{\cdot\} : \mathbb{R}^2 \mapsto \mathcal{S}_i$ denoting the projection onto the set $\mathcal{S}_i$, we set

$$\hat{\theta}_i(k+1) = \operatorname*{Proj}_{\mathcal{S}_i}\left\{\check{\theta}_i(k+1)\right\}. \tag{4}$$

(Because the set $\mathcal{S}_i$ is closed and convex, the projection function is well-defined.) Furthermore, it is convenient to parametrize $\hat{\theta}_i(k)$ as $\hat{\theta}_i(k) =: \begin{bmatrix} \hat{a}_i(k) \\ \hat{b}_i(k) \end{bmatrix}$. Also, define the associated Lyapunov function, $V_i(k) := [\theta^* - \hat{\theta}_i(k)]^\top [\theta^* - \hat{\theta}_i(k)]$.

It is common in the literature to replace the above algorithm with the classical projection algorithm [6], [5] (with $0 < a < 2, c > 0$):

$$\hat{\theta}_i(k+1) = \hat{\theta}_i(k) + \frac{a\phi(k)}{c + \phi(k)^\top \phi(k)} e_i(k+1). \tag{5}$$

The addition of the c term will prevent numerical problems when $\phi(k)$ is close to zero in (3). However, the gain of the update law of the classical algorithm is small when $\phi(k)$ is small, which is the reason why the closed-loop behavior in the adaptive control context is asymptotic rather than exponential (in general) when an estimator of this form is used.

The following proposition lists properties of the estimation algorithm (3)–(4):

Proposition 2 [15]. *For every* $k_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\hat{\theta}_i(k_0), \theta^* \in \mathcal{S}_i$, $w \in \ell_\infty$, *when the estimation algorithm in* (3) *and* (4) *is applied to the plant* (1), *the following holds:*

$$V_i(k) \leq V_i(k_0) - \frac{1}{2} \sum_{j=k_0, \phi(j)\neq 0}^{k-1} \frac{|e_i(j+1)|^2}{\|\phi(j)\|^2} + 2 \sum_{j=k_0, \phi(j)\neq 0}^{k-1} \frac{|w(j)|^2}{\|\phi(k)\|^2}, \qquad k \geq k_0 + 1. \tag{6}$$

B. Switching Controller

Let $r(\cdot)$ be the reference signal to be tracked. We assume that it is known one step ahead. If we invoke the Certainty Equivalence Principle there is a natural choice for the one-step-ahead adaptive control law associated with the i^{th} estimator:

$$u(k) = -\frac{\hat{a}_i(k)}{\hat{b}_i(k)} y(k) + \frac{1}{\hat{b}_i(k)} r(k+1),$$

which ensures that $r(k+1) = \phi(k)^\top \hat{\theta}_i(k)$. Here, of course, we do not know which $\mathcal{S}_i$ contains θ^*. In fact, θ^* may lie in more than one such set.

Let us define the index set $\mathcal{I} = \{1, 2, \ldots, m\}$. To this end, we define a switching signal $\sigma : \mathbb{Z} \mapsto \mathcal{I}$ that decides which controller to use at any given point in time, i.e. we set

$$u(k) = -\frac{\hat{a}_{\sigma(k)}(k)}{\hat{b}_{\sigma(k)}(k)} y(k) + \frac{1}{\hat{b}_{\sigma(k)}(k)} r(k+1). \tag{7}$$

Next we define the tracking error by $\varepsilon(k) := y(k) - r(k)$; when the above control law is applied, it is easy to see that

$$\varepsilon(k+1) = e_{\sigma(k)}(k+1). \tag{8}$$

III. Exponential Stabilization for $m = 2$

We begin with the case of two uncertainty sets, i.e. we have $\mathcal{I} = \{1, 2\}$. Here we adopt the following simple switching rule: with an initial condition of $\sigma(k_0) = \sigma_0$,

$$\sigma(k) = \operatorname*{argmin}_{i \in \mathcal{I}} |e_i(k)|, \quad k \geq k_0 + 1. \tag{9}$$

This logic chooses the model with the minimum prediction error: it is memoryless and is a function only of signals at the same instant. For the case when $|e_1(k)| = |e_2(k)|$, we (somewhat arbitrarily) select $\sigma(k)$ to be 1. Now, we present the main result of this paper.

Theorem 1. *Consider the plant* (1) *with* $\mathcal{I} = \{1,2\}$ *and suppose that the controller consisting of the estimator* (3) *and* (4)*, the control law* (7)*, and the switching rule* (9) *is applied. For every* $\lambda \in (0,1)$*, there exists a constant* $\gamma > 0$ *such that for every* $k_0 \in \mathbb{Z}$*,* $y_0 \in \mathbb{R}$*,* $\sigma_0 \in \mathcal{I}$*,* $\theta^* \in \mathcal{S}$*,* $\hat{\theta}_i(k_0) \in \mathcal{S}_i$ *(*$i \in \mathcal{I}$*), and* $r, w \in \ell_\infty$*, the closed-loop system satisfies*

$$\|\phi(k)\| \le \gamma\lambda^{k-k_0}|y_0| + \sum_{j=k_0}^{k} \gamma\lambda^{k-1-j}|r(j+1)| + \sum_{j=k_0}^{k-1} \gamma\lambda^{k-1-j}|w(j)|, \quad k \ge k_0. \quad (10)$$

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 sum bound on the effect of the exogenous signals. If the initial condition is zero, there is a bounded gain on the map from the exogenous signals (the noise and reference signal) to ϕ; in classical adaptive control this is rarely the case. This is analogous to the result in [14] which deals with one convex uncertainty set. Furthermore, while the choice in (9) seems obvious, as far as we are aware there is no proof of stability in the literature for the situation in which the classical algorithm (5) is used in conjunction with the control law (7).

Before proving Theorem 1 we first show that the simple logic in (9) yields a very desirable closed-loop property.

Lemma 1. *Consider the plant* (1) *with* $m = 2$*, and suppose that the controller consisting of the estimator* (3) *and* (4)*, the control law* (7)*, and the switching rule* (9) *is applied. Then for every* $k_0 \in \mathbb{Z}$*,* $y_0 \in \mathbb{R}$*,* $\sigma_0 \in \{1,2\}$*,* $\theta^* \in \mathcal{S}$*, and* $\hat{\theta}_i(k_0) \in \mathcal{S}_i$ *(*$i = 1,2$*) and* $r, w \in \ell_\infty$*, for every* $j \in \{1,2\}$ *and* $k \ge k_0 + 1$ *we have that*

(a) $|\varepsilon(k)| \le |e_j(k)|$ *or*
(b) $|\varepsilon(k+1)| \le |e_j(k+1)|$.

Proof. Fix $k_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \{1,2\}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(k_0) \in \mathcal{S}_i$ $(i = 1,2)$, and $r, w \in \ell_\infty$, and let $j \in \{1,2\}$ and $k \ge k_0+1$ be arbitrary. Let $\bar{j}$ be the element of $\{1,2\}$ which is not j. Suppose that (b) fails to hold; in view of (8) it must be that $\sigma(k) = \bar{j}$; from (9) this means that $|e_{\bar{j}}(k)| \le |e_j(k)|$. Since $\varepsilon(k) \in \{e_1(k), e_2(k)\}$, we conclude that $|\varepsilon(k)| \le |e_j(k)|$, i.e. (a) holds. ■

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

We now state a technical result to be used in the proof of Theorem 1.

Lemma 2. [14] *(i) With* $n \in \mathbb{N} \cup \{\infty\}$*, suppose that* $a_j \in \mathbb{R}$ *and* $c > 0$ *satisfy*

$$\sum_{j=0}^{n} a_j^2 \le c.$$

Then for every $\lambda \in (0,1)$*, if we define* $\gamma := c^{\frac{c+1}{2}}\left(\frac{1}{\lambda}\right)^{\frac{c}{\lambda^2}+1}$*, then the following holds:*

$$\left|\prod_{l=0}^{j-1} a_l\right| \le \gamma\lambda^j, \qquad j = 0,1,\dots,n.$$

(ii) With $n < p \le \infty$*, suppose that* $a_j \in \mathbb{R}$ *and* $c_1 > 0$ *satisfy*

$$\sum_{l=j}^{j+n} a_l^2 \le c_1, \qquad j = 0,1,\dots,p-n.$$

Then for every $\lambda \in (0,1)$*, if* c_1, λ *and* n *satisfy*

$$n \ge \frac{\frac{c_1+1}{2}\ln(c_1) + (4\frac{c_1}{\lambda^2}+1)(\ln(2) - \ln(\lambda))}{\ln(2)}$$

and $\gamma_1 := c_1^{\frac{c_1+1}{2}}\left(\frac{2}{\lambda}\right)^{\frac{4c_1}{\lambda^2}+1}$*, then*

$$\left|\prod_{l=0}^{j-1} a_l\right| \le \gamma_1\lambda^j, \qquad j = 0,1,\dots,p.$$

Also, before proving the main result, define the constants $\bar{\mathbf{s}} := \max_i \|\mathcal{S}_i\|$, $\bar{a} := \max\{|a| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}_1 \cup \mathcal{S}_2\}$, $\bar{b} := \max\{|b| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}_1 \cup \mathcal{S}_2\}$, $\bar{f} := \max\{|\frac{a}{b}| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}_1 \cup \mathcal{S}_2\}$, and $\bar{g} := \max\{|\frac{1}{b}| : \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathcal{S}_1 \cup \mathcal{S}_2\}$.

Proof of Theorem 1. The proof is a significant extension of that of the main result of [14]. The proof will be given for the case when no noise enters the system, followed by the case with noise.

Fix $\lambda \in (0,1)$. Let $k_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(k_0) \in \mathcal{S}_i$ $(i \in \mathcal{I})$, and $r, w \in \ell_\infty$ be arbitrary. We know that there exists at least one $j \in \mathcal{I}$ so that $\theta^* \in \mathcal{S}_j$; let i^* denote the smallest such j.

First we establish some general bounds to be used throughout the proof. Setting $c_1 := (1+\bar{f})$, $c_2 := \bar{g}$, from the control law in (7) we obtain the general bound

$$\|\phi(k)\| \le c_1|y(k)| + c_2|r(k+1)|; \quad (11)$$

if we define $c_3 := \max\{\bar{a} + \bar{b}\bar{f}, \bar{b}\bar{g}\}$ from the plant equation (1) we have the crude bound

$$|y(k+1)| \le c_3|y(k)| + c_3|r(k+1)| + |w(k)|. \quad (12)$$

Case 1: $w(k) = 0$, for all $k \ge k_0$.

In this part, the proof has several steps. First, we will analyze the behavior for two consecutive instants. Then, we will consider the whole time horizon.

From Proposition 2 we have that for $k \ge k_0 + 1$,

$$\sum_{j=k_0, \phi(j)\ne 0}^{k-1} \frac{|e_{i^*}(j+1)|^2}{\|\phi(j)\|^2} \le 2V_{i^*}(k_0) \le 8\|\mathcal{S}_{i^*}\|^2 \le \underbrace{8\bar{\mathbf{s}}^2}_{=:c_4}. \quad (13)$$

For $j \ge k_0$, define

$$\alpha_j := \begin{cases} \frac{|e_{i^*}(j+1)|}{\|\phi(j)\|} & \text{if } \phi(j) \ne 0 \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

For $\phi(j) \ne 0$, we have

$$|e_{i^*}(j+1)| = \alpha_j\|\phi(j)\|. \quad (15)$$

The final published version can be found in the proceedings of the IEEE 57th Conference on Decision and Control (CDC 2018). DOI: 10.1109/CDC.2018.8619630

For $\phi(j)=0$, we have $y(j)=u(j)=0$; from (1) we conclude that $y(j+1)=0$, and from (2) we conclude that $e_i(j+1)=0$ for all $i\in\mathcal{I}$, which means that (15) holds for this case as well.

Motivated by Lemma 1, now we will analyze the closed-loop behavior on two consecutive instants of time. Let $j\in\mathbb{Z}^+$ be arbitrary; from Lemma 1 we have that either

$$|\varepsilon(k_0+2j+1)|\le|e_{i^*}(k_0+2j+1)| \text{ or}$$
$$|\varepsilon(k_0+2j+2)|\le|e_{i^*}(k_0+2j+2)|.$$

If we define $\bar{\alpha}_{k_0+2j}:=\max\{\alpha_{k_0+2j},\alpha_{k_0+2j+1}\}$ (note from (13) that $\alpha_j\le\sqrt{c_4}$), and then combine the above with (15) and (11) we can have either

$$\begin{aligned}|y(k_0+2j+1)|&\le c_1\bar{\alpha}_{k_0+2j}|y(k_0+2j)|\\&\quad+(1+c_2c_4^{\frac{1}{2}})|r(k_0+2j+1)| \text{ or}\\|y(k_0+2j+2)|&\le c_1\bar{\alpha}_{k_0+2j}|y(k_0+2j+1)|\\&\quad+(1+c_2c_4^{\frac{1}{2}})|r(k_0+2j+2)|;\end{aligned}$$

if we combine the above two cases and use (12), then there exist constants c_5 and c_6 such that for $j\in\mathbb{Z}^+$

$$\begin{aligned}|y(k_0+2j+2)|&\le c_5\bar{\alpha}_{k_0+2j}|y(k_0+2j)|\\&\quad+c_6(|r(k_0+2j+1)|+|r(k_0+2j+2)|).\end{aligned}\tag{16}$$

Now we examine the behavior across the whole time horizon. Observe from (13) that $\sum_{j=0}^{\infty}\bar{\alpha}_{k_0+2j}^2\le c_4$. Now define $\lambda_1=\frac{\lambda^2}{c_5}$. Motivated by Lemma 2, define $c_7:=c_4^{\frac{c_4+1}{2}}(\frac{1}{\lambda_1})^{\frac{c_4}{\lambda_1^2}+1}$. From Lemma 2 we conclude that

$$\prod_{l=0}^{j-1}\bar{\alpha}_{k_0+2l}\le c_7\lambda_1^j,\qquad j\in\mathbb{Z}^+.$$

If we solve the difference inequality (16) recursively and apply the above bound, we can obtain for $j\in\mathbb{Z}^+$

$$|y(k_0+2j)|\le c_7\lambda^{2j}|y(k_0)|+\sum_{l=0}^{j-1}\frac{c_7c_6}{\lambda}\lambda^{2(j-l-1)}|r(k_0+l+1)|.\tag{17}$$

We can use (12) to obtain a bound for the remaining time instants of $y(k)$; if we combine this with (11) then the result (10) follows.

Case 2: $w(k)\ne0$ for some k.

We now analyze the case when there is noise entering the system; this is more complicated since $V_{i^*}(k)$ is no longer monotonically decreasing. Motivated by Case 1, in the following we will be applying Lemma 2(ii) with a larger bound than in (13)–define $\hat{c}_4:=12\bar{\mathbf{s}}^2$. In light of Lemma 2(ii), also define $\lambda_1=\frac{\lambda^2}{c_5}$, $p:=\left\lceil\frac{\frac{\hat{c}_4+1}{2}\ln(\hat{c}_4)+(4\frac{\hat{c}_4}{\lambda_1^2}+1)(\ln(2)-\ln(\lambda_1))}{\ln(2)}\right\rceil$ and $\hat{c}_7:=\hat{c}_4^{\frac{\hat{c}_4+1}{2}}(\frac{2}{\lambda_1})^{\frac{4\hat{c}_4}{\lambda_1^2}+1}$.

Let us turn now to defining two sets in relation to the size of the noise: $K_+=\left\{j\ge k_0:\phi(j)\ne0 \text{ and } \frac{|w(j)|^2}{\|\phi(j)\|^2}<\frac{\bar{\mathbf{s}}^2}{p}\right\}$ and $K_-=\left\{j\ge k_0:\phi(j)=0 \text{ or } \frac{|w(j)|^2}{\|\phi(j)\|^2}\ge\frac{\bar{\mathbf{s}}^2}{p}\right\}$. Now we partition the time index $\{j\in\mathbb{Z}:j\ge k_0\}$ into intervals which oscillate between K_+ and K_-. We can clearly define a (possibly infinite) sequence of intervals of the form $[k_l,k_{l+1})$ which satisfy: 1) k_0 serves as the initial instant of the first interval; 2) $[k_l,k_{l+1})$ either belongs to K_+ or K_-; and 3) if $k_{l+1}\ne\infty$ and $[k_l,k_{l+1})$ belongs to K_+ then $[k_{l+1},k_{l+2})$ belongs to K_-, and vice versa.

Case 2a: $[k_l,k_{l+1})$ belongs to K_-.

Let $j\in[k_l,k_{l+1})$ be arbitrary. So we have $\|\phi(j)\|=0$ or $\frac{|w(j)|^2}{\|\phi(j)\|^2}\ge\frac{\bar{\mathbf{s}}^2}{p}$; from this and the plant model (1), there exist $\hat{c}_8,\hat{c}_9$ so that

$$|y(j)|\le\begin{cases}\hat{c}_8|w(j)| & j=k_l\\ \hat{c}_9|w(j-1)| & j\in[k_l+1,k_{l+1}].\end{cases}\tag{18}$$

Case 2b: $[k_l,k_{l+1})$ belongs to K_+.

Using the same notation as in Case 1, we define $\alpha_j:=\frac{|e_{i^*}(j+1)|}{\|\phi(j)\|}$. For $j\in\mathbb{Z}^+$ so that $k_l+2j+1\le k_{l+1}$, we define $\bar{\alpha}_{k_l+2j}:=\max\{\alpha_{k_l+2j},\alpha_{k_l+2j+1}\}$. At this point, the goal is to apply Lemma 2 to analyze the closed-loop behavior. To this end, from Proposition 2 we see that if $k_{l+1}-k_l\le p$, then

$$\sum_{j=k_l}^{k_{l+1}-1}\alpha_j^2\le8\bar{\mathbf{s}}^2+4(k_{l+1}-k_l)\frac{\bar{\mathbf{s}}^2}{p}\le12\bar{\mathbf{s}}^2=\hat{c}_4,$$

so by Lemma 2(i) we have

$$\prod_{t=0}^{j-1}\bar{\alpha}_{k_l+2t}\le\hat{c}_7\lambda_1^j,\qquad j\in\mathbb{Z}^+\text{ s.t. }k_l+2j+1\le k_{l+1},$$

and if $k_{l+1}-k_l>p$, then by choice of p and Lemma 2(ii) we have that

$$\prod_{t=0}^{j-1}\bar{\alpha}_{k_l+2t}\le\hat{c}_7\lambda_1^j,\qquad j\in\mathbb{Z}^+\text{ s.t. }k_l+2j+1\le k_{l+1},$$

as well.

If we now analyze the closed-loop system as in the noise-free case, we end up with a version of (16) with the noise now included: there exists a constant $\hat{c}_6$ so that

$$\begin{aligned}&|y(k_l+2j+2)|\le c_5\bar{\alpha}_{k_l+2j}|y(k_l+2j)|\\&+\hat{c}_6(|r(k_l+2j+1)|+|r(k_l+2j+2)|+|w(k_l+2j)|\\&+|w(k_l+2j+1)|),j\in\mathbb{Z}^+\text{ s.t. }k_l+2j+1\le k_{l+1}.\end{aligned}\tag{19}$$

Using a similar analysis to that of Case 1, we can solve the above difference inequality and conclude that there exists a constant $\hat{c}_{10}$ so that

$$|y(k)|\le\hat{c}_{10}\lambda^{k-k_l}|y(k_l)|+\sum_{j=k_l}^{k-1}\hat{c}_{10}\lambda^{k-j-1}(|r(j+1)|+|w(j)|),$$
$$k=k_l,k_l+1,\dots,k_{l+1}.\tag{20}$$

The final step here is to glue together the bounds on the sequence of intervals in K_+ and K_-; the argument is identical to that of [14]. ■

IV. Exponential Stabilization for $m\ge2$

Now we consider the case of $m>2$ uncertainty sets. As mentioned earlier, it may be beneficial for performance to have more than two sets. Unfortunately, although the rule in (9) is a well-defined rule for all cases, we are unable to prove that it will work in case of $m>2$. In particular, a potential problem is that the algorithm could oscillate between two bad choices, and never (or rarely) choose the correct one; it is not clear that Lemma 1 would hold for this case. Instead, we propose a modified version of (9). At each point in time k we have an admissible set $\mathcal{I}_k$: we initialize $\mathcal{I}_{k_0}=\mathcal{I}$, and we obtain $\mathcal{I}_k$ from $\mathcal{I}_{k-1}$ by removing all $j\in\mathcal{I}_{k-1}$ satisfying $|\varepsilon(k)|\le|e_j(k)|$; clearly $j=\sigma(k-1)$ satisfies this bound, but more j's may as well; if this results in $\mathcal{I}_k$ being empty, then we reset $\mathcal{I}_k$ to be $\mathcal{I}$. This is formulated in the pseudocode in **Algorithm 1**. The following result is similar to that of Lemma 1.

The final published version can be found in the proceedings of the IEEE 57th Conference on Decision and Control (CDC 2018). DOI: 10.1109/CDC.2018.8619630

Lemma 3. *Consider the plant* (1) *for which $m \geq 2$, and suppose that the controller consisting of the estimator* (3) *and* (4)*, the control law* (7)*, and the switching signal defined in Algorithm 1 is applied. Then for every $k_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(k_0) \in \mathcal{S}_i$ $(i \in \mathcal{I})$, $r, w \in \ell_\infty$, if $\underline{k}$ and $\bar{k}$ are two consecutive reset times of the index set $\mathcal{I}$, then for every $j \in \mathcal{I}$ there exists a $\hat{k} \in (\underline{k}, \bar{k}]$ such that:*

$$|\varepsilon(\hat{k})| \leq |e_j(\hat{k})|.$$

Proof. This follows easily from Algorithm 1. ∎

Theorem 2. *Consider the plant in* (1) *with $\mathcal{I} = \{1, 2, \ldots, m\}$ and suppose that the controller consisting of the estimator* (3) *and* (4)*, the control law* (7)*, and the switching signal defined in Algorithm 1 is applied. For every $\lambda \in (0, 1)$, there exists a constant $\gamma > 0$ such that for every $k_0 \in \mathbb{Z}$, $y_0 \in \mathbb{R}$, $\sigma_0 \in \mathcal{I}$, $\theta^* \in \mathcal{S}$, $\hat{\theta}_i(k_0) \in \mathcal{S}_i$ $(i \in \mathcal{I})$ and $r, w \in \ell_\infty$, the closed-loop system satisfies*

$$\|\phi(k)\| \leq \gamma\lambda^{k-k_0}|y_0| + \sum_{j=k_0}^{k} \gamma\lambda^{k-1-j}|r(j+1)| + \sum_{j=k_0}^{k-1} \gamma\lambda^{k-1-j}|w(j)|, \quad k \geq k_0. \tag{21}$$

Proof. We apply an analysis similar to that of the proof of Theorem 1; instead of analyzing two consecutive instants, we analyze intervals between index set resets. We apply Lemma 3 for the case of $j = i^*$; we further utilize the fact that the maximum length between any consecutive resets is not more than m. We omit details due to space limitations. ∎

Remark. In [15] it is proven that we can leverage the fact that a convolution bound holds in the case of a fixed plant parameter to prove that a convolution bound (with larger constants) also holds if we allow time-variation together with occasional jumps. The argument provided there can be used to prove the same thing here.

V. A Simulation Example

A simulation example is provided to illustrate the results of this paper. Consider the time-varying plant:

$$y(k+1) = a(k)y(k) + b(k)u(k) + w(k),$$

with $\theta^*(k) := [a(k) \;\; b(k)]^\top$ belonging to the uncertainty set:

$$\mathcal{S} = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \in [1,2] \cup [-2,-1], b \in [1,2] \cup [-2,-1] \right\},$$

which can be visualized in Figure 1. Hence, every admissible model is unstable, and the sign of the input gain b is unknown. Here the plant parameters are varying as follows:

$$a(k) = \begin{cases} -\frac{3}{2} - \frac{1}{2}\sin(\frac{1}{20}k), & 51 \leq k \leq 100, 151 \leq k \leq 200 \\ \frac{3}{2} + \frac{1}{2}\sin(\frac{1}{20}k), & \text{otherwise,} \end{cases}$$

$$b(k) = \begin{cases} -\frac{3}{2} - \frac{1}{2}\cos(\frac{1}{15}k), & 101 \leq k \leq 150, 151 \leq k \leq 200 \\ \frac{3}{2} + \frac{1}{2}\cos(\frac{1}{15}k), & \text{otherwise.} \end{cases}$$

In the first approach, we define two convex sets by convexifying the 1^{st} and 2^{nd} quadrants and the 3^{rd} and 4^{th} quadrants, respectively, yielding

$$\mathcal{S}_1 := \left\{ \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [-2,2], b \in [1,2] \right\}, \;\; \mathcal{S}_2 := \left\{ \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [-2,2], b \in [-2,-1] \right\}.$$

Algorithm 1: *when $m \geq 2$*

```
Initialize 𝓘_{k0} = 𝓘 and σ(k0) = σ0;
while k > k0 do
    𝓘_k = 𝓘_{k−1} \ {i ∈ 𝓘_{k−1} : |ε(k)| ≤ |e_i(k)|};
    if 𝓘_k = ∅ then
        𝓘_k = 𝓘                (index set reset)
    end
    σ(k) = argmin_{i ∈ 𝓘_k} |e_i(k)|
end
```

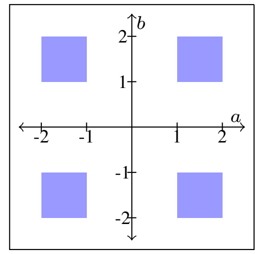


Fig. 1. Uncertainty set $\mathcal{S}$

We will apply the control input (7) using estimates from (3) and (4) and $\sigma(k)$ determined by (9). We set $\hat{\theta}_1(k_0) = [1.5 \;\; 1.5]^\top$, $\hat{\theta}_2(k_0) = [-1.5 \; -1.5]^\top$, $\sigma_0 = 2$, $y_0 = -1$, the reference $r(\cdot)$ to be a unit-amplitude square wave of period 60, and noise $w(\cdot)$ which is a uniformly distributed random signal, with a magnitude of 5% of the reference magnitude. Figures 2(a) display the results. We see that the controller does a reasonable job, even though the switching often chooses the wrong model. Large transient may ensue, but on average the adaptive controller provides good tracking.

As mentioned earlier, it may be beneficial to have more than two convex sets. So in the second approach, we define four convex sets in the following natural way:

$$\mathcal{S}_1 := \left\{ \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [1,2], b \in [1,2] \right\}, \;\; \mathcal{S}_2 := \left\{ \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [1,2], b \in [-2,-1] \right\},$$

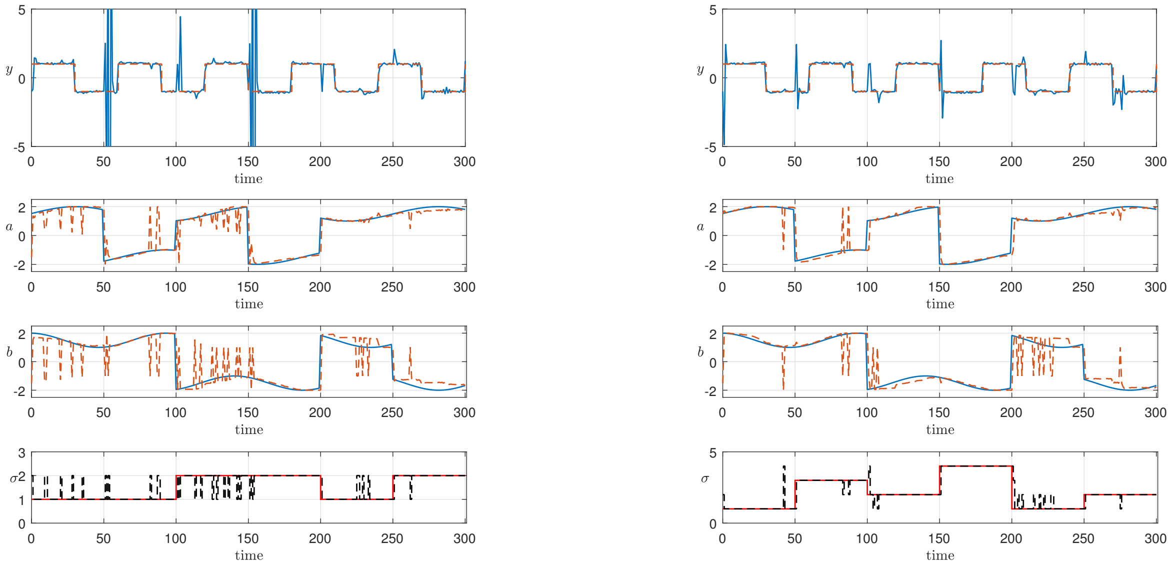
$$\mathcal{S}_3 := \left\{ \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [-2,-1], b \in [1,2] \right\}, \;\; \mathcal{S}_4 := \left\{ \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \in \mathbb{R}^2 : a \in [-2,-1], b \in [-2,-1] \right\}.$$

We will apply the control input (7) using estimates from (3) and (4) and $\sigma(k)$ determined by Algorithm 1. We set $\hat{\theta}_1(k_0) = [1.5 \;\; 1.5]^\top$, $\hat{\theta}_2(k_0) = [1.5 \; -1.5]^\top$, $\hat{\theta}_3(k_0) = [-1.5 \;\; 1.5]^\top$, $\hat{\theta}_4(k_0) = [-1.5 \; -1.5]^\top$, $\sigma_0 = 2$. As above we set $y_0 = -1$, the reference $r(\cdot)$ to be a unit-amplitude square wave of period 60, and noise $w(\cdot)$ which is a uniformly distributed random signal, with a magnitude of 5% of the reference magnitude. Figures 2(b) display the results. We see the controller does a good job of tracking, and with smaller transients than in the first approach. Furthermore, the estimator does a fairly good job of tracking the time-varying parameter. Both examples illustrate that the approach handles time-variation and occasional jumps.

VI. Conclusion and Future Work

In this paper, we have considered the first-order case with unknown plant parameters belonging to a closed and bounded uncertainty set; we designed a one-step-ahead adaptive controller. No assumption on the convexity of the uncertainty set is imposed. A parameter estimation process is run by having multiple parallel estimators with each operating on a compact and convex set. A switching logic is used to determine which parameters are used in the controller. The corresponding one-step-ahead adaptive controller guarantees linear-like convolution bounds on the closed loop behavior, which implies exponential stability and a bounded noise gain.

We would like to extend the proposed switching approach to high order plants, both in the context of one-step-ahead



(a) (b)

Fig. 2. The upper plot shows both the reference (dashed) and the output (solid); the next two plots show the parameter estimates $\hat{\theta}_{\sigma(k)}(k)$ (dashed) and actual parameters $\theta^*(k)$ (solid); the bottom plot shows the switching signal (dashed) and the correct index (solid).

control considered here as well as the pole-placement setup of [15]. While simulation indicate that they work, the proofs remain elusive.

References

[1] S. Baldi, "Multiple Estimation Architecture in Discrete-Time Adaptive Mixing Control," *Machines*, vol. 1, no. 1, pp. 33–49, May 2013.

[2] 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.

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

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

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

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

[7] 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.

[8] 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.

[9] P. Ioannou and K. Tsakalis, "A robust direct adaptive controller," *IEEE Transactions on Automatic Control*, vol. 31, no. 11, pp. 1033–1043, Nov. 1986.

[10] G. Kreisselmeier and B. Anderson, "Robust model reference adaptive control," *IEEE Transactions on Automatic Control*, vol. 31, no. 2, pp. 127–133, Feb. 1986.

[11] M. Kuipers and P. Ioannou, "Multiple Model Adaptive Control With Mixing," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1822–1836, Aug. 2010.

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

[13] D. E. Miller, "Adaptive stabilization using a nonlinear time-varying controller," *IEEE Transactions on Automatic Control*, vol. 39, no. 7, pp. 1347–1359, Jul. 1994.

[14] 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.

[15] 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.

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

[17] A. S. Morse, "Supervisory control of families of linear set-point controllers–Part I. Exact matching," *IEEE Transactions on Automatic Control*, vol. 41, no. 10, pp. 1413–1431, 1996.

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

[19] D. Mudgett and A. S. Morse, "Adaptive stabilization of linear systems with unknown high-frequency gains," *IEEE Transactions on Automatic Control*, vol. 30, no. 6, pp. 549–554, Jun. 1985.

[20] S. Naik, P. Kumar, and B. Ydstie, "Robust continuous-time adaptive control by parameter projection," *IEEE Transactions on Automatic Control*, vol. 37, no. 2, pp. 182–197, 1992.

[21] K. Narendra and Yuan-Hao Lin, "Stable discrete adaptive control," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 456–461, Jun. 1980.

[22] K. Narendra, Yuan-Hao Lin, and L. Valavani, "Stable adaptive controller design, part II: Proof of stability," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 440–448, Jun. 1980.

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

[24] K. Narendra and J. Balakrishnan, "Adaptive control using multiple models," *IEEE Transactions on Automatic Control*, vol. 42, no. 2, pp. 171–187, 1997.

[25] R. Nussbaum, "Some remarks on a conjecture in parameter adaptive control," *Systems & Control Letters*, vol. 3, no. 5, pp. 243–246, Nov. 1983.

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

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

[28] 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.

[29] J. C. Willems and G. I. Byrnes, "Global adaptive stabilization in the absence of information on the sign of the high frequency gain," in *Analysis and Optimization of Systems*. Berlin/Heidelberg: Springer-Verlag, 1984, pp. 49–57.

[30] B. Ydstie, "Transient performance and robustness of direct adaptive control," *IEEE Transactions on Automatic Control*, vol. 37, no. 8, pp. 1091–1105, 1992.