This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

Revisiting Model Reference Adaptive Control: Linear-like Closed-loop Behavior

Mohamad T. Shahab and Daniel E. Miller

***Abstract*—In this paper we examine the model reference adaptive control (MRAC) problem when the commonly used projection algorithm is utilized, subject to several common assumptions on the set of admissible parameters, in particular a compactness constraint as well as knowledge of the sign of the high-frequency gain. It is proven in the literature that for this setup, the closed-loop system is bounded-input bounded-state; since the closed-loop system is not linear time-invariant, this does not imply a bounded gain. Here we prove a much crisper and detailed bound on the closed-loop behavior consisting of three terms: a decaying exponential on the initial condition, a linear-like convolution bound on the exogenous inputs, and a constant scaled by the square root of the constant in the denominator of the estimator update law; we also provide an upper bound on the 2-norm of the tracking error. We then demonstrate that the same kind of bounds hold in the presence of a degree of unmodelled dynamics and plant parameter time-variation.**

I. Introduction

Adaptive control is an approach used to deal with systems with uncertain and/or time-varying parameters. In the classical approach to adaptive control, one combines a linear time-invariant (LTI) compensator together with a tuning mechanism to adjust the compensator parameters to match the plant. The first general proofs came around 1980, e.g. see [6], [9], [23], [26] and [27]. However, the original controllers are typically not robust to unmodelled dynamics, do not tolerate time-variations well, have poor transient behavior and do not handle noise/disturbances well, e.g. see [28]. During the following two decades, a good deal of research was carried out to alleviate these shortcomings; a number of small controller design changes were proposed, such as the use of signal normalization, deadzones and σ-modification, e.g. see [11], [12], [16], [15], and [35]; also, simply using projection onto a convex set of admissible parameters turned out to be powerful, e.g. see [14], [25], [36], [37] and [39]. However, in general these redesigned controllers may provide asymptotic closed-loop behavior but no exponential stability nor bounded gain on the noise are proven[1]; that being said, some of them, especially those using projection, provide a bounded-noise bounded-state property, as well as tolerance of some degree of unmodelled dynamics and/or time-variations.

Recently, for discrete-time LTI plants, in the d-step ahead control setting [18], [21], [22], the model reference adaptive control setting [32], and the pole-placement control setting [19], [20], [33], subject to some standard assumptions a new approach has been proposed which not only provides exponential stability and a bounded gain on the noise, but also a convolution bound on the exogenous inputs; the resulting convolution bound is leveraged to prove tolerance to a degree of time-variations and to a degree of unmodelled dynamics [31]. As far as the authors are aware, such **linear-like convolution bounds have never before been proven in the adaptive setting**. The key idea is to use the original (ideal) projection algorithm in conjunction with a restriction of the parameter estimates to a convex set, although this convexity requirement was relaxed in [29], [20] and [33].

However, in the literature it is very common to adopt a modified version of the ideal projection algorithm in which a constant is added to the denominator of the update equation, which we term the **classical projection algorithm**. This is widely used, and in the disturbance-free situation, stability and tracking is proven under minimal assumptions [9]. In the presence of disturbances, analysis is harder; perhaps the most general result is proven in [36], wherein it is proven that a bounded disturbance yields a bounded state. Other examples where this classical estimation algorithm is utilized include [12], [13], [14], [15], [8], [16], [39], and [24]. Here our objective is to obtain a quantitative bound in terms of the initial condition, the exogenous inputs, and the key estimator parameter; we do this by extending the ideas of our earlier work on the ideal projection algorithm to this setting. This requires some new non-obvious twists in the analysis and yields a modified version of the aforementioned convolution bound. To this end, we impose several standard classical assumptions, including knowledge of the sign of the high-frequency gain; since our objective is to obtain uniform bounds, we impose a natural compactness assumption. Here we prove a bound on the closed-loop system consisting of three terms: a decaying exponential on the initial condition, a linear-like convolution bound on the exogenous inputs, and a constant scaled by the square root of the constant in the denominator of the estimator update law. This bound clearly has richer structure than simply "bounded-disturbance bounded-state", and nicely generalizes our earlier work on the ideal projection algorithm, wherein the last term is missing. Furthermore, in the absence of a disturbance we are able to obtain a crisp bound on the size of the tracking error, rather than the classical one stating simply that it is square summable. We also demonstrate that the same kind of linear-like bounds hold in the presence of a degree of unmodelled dynamics and plant parameter time-variation. This work demonstrates that while such adaptive control laws are inherently nonlinear, surprisingly linear-like bounds on the closed-loop behavior exist. These bounds are clearly non-obvious, and are arguably unexpected, given that the algorithms that we analyze have been around for over 40 years, and have been widely studied, but we are the first to point them out. They provide quantitative bounds on the closed-loop behavior rather than the qualitative bounds which are the only ones available to date.

At this point we briefly put our work into context with respect to our earlier work. All of our work to date has been on the ideal projection algorithm:

(i) we started with the first order case [18];
(ii) we then analyzed the pole placement stability setting [19], [20];
(iii) we extended the pole placement setting to handle tracking and order uncertainty [30], [33];
(iv) we extended the pole placement approach of (ii) to handle the one-step-ahead adaptive control problem [21], [22];
(v) most recently, we have extended (iv) to the more general model reference adaptive control problem in the conference paper [32].

Here, in the this paper, we will build on the results of (ii), (iv) and (v) to analyze the case of the classical estimator. It can be viewed as an extension of [32] discussed in (v). However,

M.T. Shahab was with the Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division at King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia. Email: moh_shahab@hotmail.com, m4shahab@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.

Support for this research was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by King Abdullah University of Science and Technology (KAUST).

[1]An exception is the work of Ydstie [39] where a bounded gain is proven.

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

no proofs are provided there and the classical estimator case is not analyzed at all; here we develop completely new proofs for this case.

Notation. We use standard notation throughout the paper. We denote $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{Z}^+$, $\mathbb{N}$ and $\mathbb{C}$ as the set of real numbers, integers, non-negative integers, natural numbers and complex numbers, respectively. We will denote the Euclidean-norm of a vector and the induced norm of a matrix by the subscript-less default notation $\|\cdot\|$. Let $\mathbb{S}(\mathbb{R}^{p\times q})$ denote the set of $\mathbb{R}^{p\times q}$-valued sequences. Also, $\boldsymbol{\ell}_\infty$ denotes the set of bounded sequences. 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 $\mathrm{Proj}_\Omega\{\cdot\} : \mathbb{R}^p \to \Omega$ denote the projection onto the set Ω in the 2-norm; because the set Ω is closed and convex, the function Proj_Ω is well-defined. If $\Omega \subset \mathbb{R}^p$ is a compact (closed and bounded) set, we define $\|\Omega\| := \max_{x\in\Omega}\|x\|$. Let $\lceil\cdot\rceil$ denote the ceiling function. Let I_p denote the identity matrix of size p. Define the normal vector $\mathbf{e}_j \in \mathbb{R}^p$ of appropriate length p as

$$\mathbf{e}_j := \big[\underbrace{0 \quad \cdots \quad 0}_{j-1 \text{ elements}} \quad 1 \quad 0 \quad \cdots \quad 0\big]^\top.$$

Last of all, for a signal $f \in \mathbb{S}(\mathbb{R})$ which is sufficiently well-behaved to have a z-transform, we let $F(z)$ denote this quantity.

II. The Setup

In this paper we consider the following linear time-invariant (LTI) discrete-time plant:

$$y(t) + \sum_{i=1}^{n} a_i y(t-i) = \sum_{i=0}^{m} b_i u(t-d-i) + w(t), \quad t \in \mathbb{Z}, \tag{1}$$

with $y(t) \in \mathbb{R}$ as the measured output, $u(t) \in \mathbb{R}$ as the control input, and $w(t) \in \mathbb{R}$ as the noise/disturbance. The plant parameters are regularized and the system delay is exactly d, i.e. $b_0 \neq 0$. Associated with this plant are the polynomials $\mathbf{A}(z^{-1}) := 1 + \sum_{i=1}^{n} a_i z^{-i}$ and $\mathbf{B}(z^{-1}) := \sum_{i=0}^{m} b_i z^{-i}$, the transfer function $z^{-d}\frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}$, and the unknown plant parameter vector:

$$\theta := \begin{bmatrix} a_1 & a_2 & \cdots & a_n & b_0 & b_1 & \cdots & b_m \end{bmatrix}^\top;$$

we assume that θ belongs to a known set $\mathcal{S}_{ab} \subset \mathbb{R}^{n+m+1}$. Observe that such a plant can be expressed in the z-transform domain as

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

The control objective is closed-loop stability and asymptotic tracking of a given reference signal $y^*(t) \in \mathbb{R}$ generated as the output of a reference model. More specifically, given pre-designed polynomials $\mathbf{H}(z^{-1}) := \sum_{i=0}^{n'-d} h_i z^{-i}$ and $\mathbf{L}(z^{-1}) := 1+\sum_{i=1}^{n'} l_i z^{-i}$ (with $n' \leq n$), and given a bounded exogenous signal $r(t) \in \mathbb{R}$, we utilize the following reference model expressed in the z-transform form:

$$\mathbf{L}(z^{-1})Y^*(z) = z^{-d}\mathbf{H}(z^{-1})R(z). \tag{3}$$

We assume that the roots of $\mathbf{L}(z^{-1})$ belong to the open unit desk, i.e. the reference model is stable. If we define the **tracking error** ε by

$$\varepsilon(t) := y(t) - y^*(t), \tag{4}$$

then the goal is to drive ε to zero asymptotically while maintaining "closed-loop stability".

Remark 1. *Notice that for the d-step-ahead control problem, the reference model is simply $Y^*(z) = z^{-d}R(z)$.*

We impose the following assumptions on the set of admissible parameters.

Assumption 1. $\mathcal{S}_{ab}$ is closed and bounded (compact), and for each $\theta \in \mathcal{S}_{ab}$, the corresponding $\mathbf{B}(z^{-1})$ has roots in the open unit disk and the sign of b_0 is always the same.

The boundedness requirement on $\mathcal{S}_{ab}$ is reasonable in practical situations; it is used here to prove uniform bounds and decay rates on the closed-loop behavior. The constraint on the roots of $\mathbf{B}(z^{-1})$ is a requirement that the plant be minimum phase; this is necessary to ensure tracking of bounded reference signals [17]. Knowledge of the sign of the high-frequency gain b_0 is common in adaptive control [10].

Remark 2. *It is implicit in the assumptions that we know the system delay d as well as upper bounds on the orders of $\mathbf{A}(z^{-1})$ and $\mathbf{B}(z^{-1})$.*

To proceed, we use a parameter estimator together with an adaptive control law based on the Certainty Equivalence Principle. It is convenient to put the plant into the so-called *predictor form*, e.g. see [10]. To this end, by long division we can find $\mathbf{F}(z^{-1}) = \sum_{i=0}^{d-1} f_i z^{-i}$ and $\boldsymbol{\alpha}(z^{-1}) = \sum_{i=0}^{n-1} \alpha_i z^{-i}$ that satisfy the following:

$$\frac{\mathbf{L}(z^{-1})}{\mathbf{A}(z^{-1})} = \mathbf{F}(z^{-1}) + z^{-d}\frac{\boldsymbol{\alpha}(z^{-1})}{\mathbf{A}(z^{-1})};$$

if we now define $\boldsymbol{\beta}(z^{-1}) = \mathbf{F}(z^{-1})\mathbf{B}(z^{-1}) =: \sum_{i=0}^{m+d-1} \beta_i z^{-i}$, then it is easy to verify that the following is true:

$$z^{-d}\frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})} = \frac{\boldsymbol{\beta}(z^{-1})}{z^d\mathbf{L}(z^{-1}) - \boldsymbol{\alpha}(z^{-1})}. \tag{5}$$

So comparing (5) with the plant equation in (2), we are able to re-write the plant equation as

$$\mathbf{L}(z^{-1})[z^dY(z)] = \boldsymbol{\alpha}(z^{-1})Y(z) + \boldsymbol{\beta}(z^{-1})U(z) + \overline{W}(z), \tag{6}$$

with $\overline{W}(z) := z^d\mathbf{F}(z^{-1})W(z)$. Now define a weighted sum of the system output $\overline{y}$ by

$$\overline{y}(t) := y(t) + \sum_{j=1}^{n'} l_j y(t-j); \tag{7}$$

clearly the z-transform of $\overline{y}(t)$ is $\mathbf{L}(z^{-1})Y(z)$, so with

$$\phi(t) := \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-n+1) \\ u(t) \\ u(t-1) \\ \vdots \\ u(t-m-d+1) \end{bmatrix}, \quad \theta^* := \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \\ \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{m+d-1} \end{bmatrix},$$

the time-domain counterpart of (6) in predictor form is

$$\overline{y}(t+d) = \phi(t)^\top\theta^* + \overline{w}(t). \tag{8}$$

Let $\mathcal{S}_{\alpha\beta} \subset \mathbb{R}^{n+m+d}$ denote the set of admissible θ^* that arise from the original plant parameters which lie in $\mathcal{S}_{ab}$; it is clear that the associated mapping from $\mathcal{S}_{ab}$ to $\mathcal{S}_{\alpha\beta}$ is analytic, so the compactness of $\mathcal{S}_{ab}$ means that $\mathcal{S}_{\alpha\beta}$ is compact as well. Furthermore, it is easy to see that $\beta_0 = b_0$. It is desirable that the set of admissible parameters in the new parameter space be convex and closed; so at this point let

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

$\mathcal{S} \subset \mathbb{R}^{n+m+d}$ be any compact and convex set containing $\mathcal{S}_{\alpha\beta}$ for which the $(n+1)$th element is never zero; the convex hull of $\mathcal{S}_{\alpha\beta}$ will do, although it may be more convenient to use a hyperrectangle (for projection purposes). We will show an example on obtaining such a set in the simulation section.

Now define $\overline{Y}^*(z) := \mathbf{L}(z^{-1})Y^*(z)$; then the model reference control law is the one satisfying

$$\overline{y}^*(t+d) = \phi(t)^\top \theta^*.$$

In the absence of noise, and assuming the controller is applied for all $t \in \mathbb{Z}$, we can show that we have $y(t) = y^*(t)$ for all $t \in \mathbb{Z}$. In our case of unknown parameters, we seek an adaptive version of the control law which is applied after some initial time, i.e. for $t \geq t_0$.

A. Initialization

In most adaptive control results, the goal is to prove asymptotic behavior, so the details of the initial condition are unimportant. On the other hand, here we wish to obtain a crisp bound on the transient behavior, in particular a more quantative bound including exponential decay bound with respect to the initial condition and a convolution sum bound with respect to noise/disturbance and reference signals. So, we must proceed carefully. With the definition (7) in mind (and recalling that $n' \leq n$), observe that if we wish to solve (8) for $y(\cdot)$ starting at time t_0, then it is clear that we need an initial condition of

$$x(t_0) := \begin{bmatrix} y(t_0) & y(t_0-1) & \cdots & y(t_0-n-d+2) \\ u(t_0) & u(t_0-1) & \cdots & u(t_0-m-2d+2) \end{bmatrix}^\top.$$

Observe that this is sufficient information to obtain $\phi(t_0), \phi(t_0-1), \ldots, \phi(t_0-d+1)$.

B. Parameter Estimation

We can re-write the plant equation (8) as

$$\overline{y}(t+1) = \phi(t-d+1)^\top \theta^* + \overline{w}(t-d+1), \quad t \geq t_0. \tag{9}$$

Starting with an initial estimate $\theta_0 \in \mathcal{S}$ at time t_0, given an estimate $\hat{\theta}(t)$ of θ^* at time $t \geq t_0$, we define the prediction error by

$$e(t+1) := \overline{y}(t+1) - \phi(t-d+1)^\top \hat{\theta}(t); \tag{10}$$

this is a measure of the error in $\hat{\theta}(t)$. A common way to obtain a new estimate is from the solution of the optimization problem

$$\hat{\theta}(t+1) = \operatorname*{argmin}_x \left\{ \|x - \hat{\theta}(t)\| : \overline{y}(t+1) = \phi(t-d+1)^\top x \right\},$$

yielding the **ideal projection algorithm**:

$$\hat{\theta}(t+1) = \begin{cases} \hat{\theta}(t) & \phi(t-d+1) = 0 \\ \hat{\theta}(t) + \frac{\phi(t-d+1)}{\|\phi(t-d+1)\|^2} e(t+1) & \text{otherwise;} \end{cases} \tag{11}$$

at this point, we can also constrain it to $\mathcal{S}$ by projection. Of course, if $\|\phi(t-d+1)\|$ is close to zero, numerical problems may occur, so it is the norm in the literature (e.g. [10] and [9]) to add a constant to the denominator:[2] with $g > 0$, consider the **classical estimator** (with associated projection onto $\mathcal{S}$):

$$\check{\theta}(t+1) = \hat{\theta}(t) + \frac{\phi(t-d+1)}{g + \|\phi(t-d+1)\|^2} e(t+1), \tag{12a}$$

$$\hat{\theta}(t+1) = \text{Proj}_{\mathcal{S}} \left\{ \check{\theta}(t+1) \right\}. \tag{12b}$$

[2] An exception is [1] where the ideal algorithm (11) is used and Lyapunov stability is proven, but a convolution bound on the exogenous inputs is not proven, and the high-frequency gain is assumed to be known exactly.

The ideal estimator (11) has been analyzed in great detail in our earlier work on adaptive control including the first-order one-step-ahead setup [18], the high-order d-step-ahead setup [21], [22], the model reference setup [32], the pole-placement stability problem [20], and various extensions including multi-estimators and switching [29], [30], [33]. In all of these cases, quite surprisingly we are able to prove, under suitable assumptions, that the closed-loop system exhibits linear-like behavior. Here we will focus on the more commonly used classical estimator (12) (in the model reference adaptive control setting) and prove that the corresponding closed-loop behavior exhibits linear-like behavior with an offset. This is also quite an unexpected result, even to the authors!

Analyzing the closed-loop system requires a careful examination of the estimation algorithm. First define the parameter error by $\tilde{\theta}(t) := \hat{\theta}(t) - \theta^*$. The following result lists properties of the parameter estimator.

Proposition 1. *For every $t_0 \in \mathbb{Z}$, initial condition $x(t_0)$, $\theta_0 \in \mathcal{S}$, $\theta \in \mathcal{S}_{ab}, w \in \boldsymbol{\ell}_\infty$, and $g > 0$, when the parameter estimator (12) is applied to the plant (1):*
(i) the following inequalities hold:

$$\|\hat{\theta}(t+1) - \hat{\theta}(t)\| \leq \frac{|e(t+1)|}{g + \|\phi(t-d+1)\|}, \quad t \geq t_0,$$

$$\|\tilde{\theta}(t)\|^2 \leq \|\tilde{\theta}(\tau)\|^2 + \sum_{j=\tau}^{t-1} \left[-\frac{1}{2} \frac{e(j+1)^2}{g + \|\phi(j-d+1)\|^2} + \frac{2\overline{w}(j-d+1)^2}{g + \|\phi(j-d+1)\|^2} \right], \quad t > \tau \geq t_0;$$

(ii) on every interval of the form $[\underline{t}, \overline{t}) \subset [t_0, \infty)$ which satisfies

$$\|\phi(t-d+1)\|^2 \geq g, \quad t \in [\underline{t}, \overline{t}),$$

it follows that

$$\|\hat{\theta}(t+1) - \hat{\theta}(t)\| \leq \frac{|e(t+1)|}{\|\phi(t-d+1)\|}, \quad t \in [\underline{t}, \overline{t}),$$

$$\|\tilde{\theta}(t)\|^2 \leq \|\tilde{\theta}(\tau)\|^2 + \sum_{j=\tau}^{t-1} \left[-\frac{1}{4} \frac{e(j+1)^2}{\|\phi(j-d+1)\|^2} + \frac{2\overline{w}(j-d+1)^2}{\|\phi(j-d+1)\|^2} \right], \quad \overline{t} \geq t > \tau \geq \underline{t}.$$

Proof. See the Appendix. ∎

C. The Control Law

With the natural partitioning of

$$\hat{\theta}(t) =: \begin{bmatrix} \hat{\alpha}_0(t) & \cdots & \hat{\alpha}_{n-1}(t) & \hat{\beta}_0(t) & \cdots & \hat{\beta}_{m+d-1}(t) \end{bmatrix}^\top,$$

the **model reference adaptive control law** (based on the Certainty Equivalence Principle) is

$$\overline{y}^*(t+d) = \phi(t)^\top \hat{\theta}(t);$$

solving this for $u(t)$ and using the reference model (3), we have

$$u(t) = \frac{1}{\hat{\beta}_0(t)} \left[-\sum_{i=0}^{n-1} \hat{\alpha}_i(t) y(t-i) - \sum_{i=1}^{m+d-1} \hat{\beta}_i(t) u(t-i) + \sum_{i=0}^{n'-d} h_i r(t-i) \right], \quad t \geq t_0. \tag{13}$$

It is convenient for analysis to define an **auxiliary tracking error**:

$$\overline{\varepsilon}(t) := \overline{y}(t) - \overline{y}^*(t); \tag{14}$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

it is easy to show that

$$\overline{\varepsilon}(t) = -\phi(t-d)^\top \tilde{\theta}(t-d) + \overline{w}(t-d), \quad t \ge t_0 + d, \tag{15}$$

$$e(t) = -\phi(t-d)^\top \tilde{\theta}(t-1) + \overline{w}(t-d), \quad t \ge t_0 + 1, \tag{16}$$

as well as

$$\overline{\varepsilon}(t) = e(t) + \phi(t-d)^\top \left[\hat{\theta}(t-1) - \hat{\theta}(t-d)\right], \quad t \ge t_0 + d. \tag{17}$$

Observe that we can compute $\overline{\varepsilon}(t)$, $t \in \{t_0, t_0+1, \ldots, t_0+d-1\}$, from $x(t_0), w$ and y^*. In the next section we develop several models used in the analysis, after which we state and prove our results.

III. The Analysis

To facilitate our analysis, here we construct three different models describing the evolution of $\phi(\cdot)$: a model that does not use parameter estimates but is driven by the tracking error, a crude model driven by the exogenous signals, $w(\cdot)$ and $y^*(\cdot)$, used to bound the size of growth of $\phi(t)$, and thirdly, combining the above two, a model which is driven by perturbed versions of the present and past values of $\phi(\cdot)$ and capturing the effect of the parameter estimation. This third model is crucial to the analysis of the main result of this paper.

We first obtain an equation describing the evolution of $\phi(\cdot)$ which avoids using parameter estimates, though it is driven by the weighted sum of the tracking error $\overline{\varepsilon}(\cdot)$. Using the definition of ε we obtain a formula for $y(t+1)$, and using the plant equation (1) we obtain a formula for $u(t+1)$; it is easy to see that there exists a matrix $A_g \in \mathbb{R}^{(n+m+d)\times(n+m+d)}$ (which depends implicitly on $\theta \in \mathcal{S}_{ab}$) so that the following holds:

$$\begin{aligned}\phi(t+1) &= A_g\phi(t) + \mathbf{e}_1\varepsilon(t+1) + \\ &\frac{1}{b_0}\mathbf{e}_{n+1}\sum_{i=0}^{d} a_{d-i}\varepsilon(t+1+i) + \mathbf{e}_1 y^*(t+1) + \\ &\frac{1}{b_0}\mathbf{e}_{n+1}\left[\sum_{i=0}^{d} a_{d-i}y^*(t+1+i) - w(t+d+1)\right].\end{aligned} \tag{18}$$

The characteristic polynomial of A_g is $\frac{1}{b_0}z^{n+m+d}\mathbf{B}(z^{-1})$, so all of its roots are in the open unit disk.

The model in (18) is similar to the *good* model obtained in the analysis in the d-step ahead control case in [21] and [22] where it is driven by the tracking error ε. However, in the case considered here, we would like to obtain a model which is, instead, driven by $\overline{\varepsilon}$; this will turn out to be crucial in analyzing the closed-loop behavior. To this end, from (14) and the definitions of $\overline{y}$ and $\overline{y}^*$, it is easy to see that

$$\mathcal{E}(z) = \frac{1}{\mathbf{L}(z^{-1})}\overline{\mathcal{E}}(z); \tag{19}$$

so we can represent $\varepsilon(t)$ as the output of an n'th-order system driven by $\overline{\varepsilon}$ as follows: with $\zeta(t) := \begin{bmatrix}\varepsilon(t) & \varepsilon(t-1) & \cdots & \varepsilon(t-n'+1)\end{bmatrix}^\top$, and $A_l \in \mathbb{R}^{n'\times n'}$ defined by

$$A_l := \begin{bmatrix} -l_1 & -l_2 & \cdots & -l_{n'-1} & -l_{n'} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

we have

$$\zeta(t+1) = A_l\zeta(t) + \mathbf{e}_1\overline{\varepsilon}(t+1) \tag{20a}$$

$$\varepsilon(t) = \mathbf{e}_1^\top\zeta(t). \tag{20b}$$

Note that (18) is driven on the RHS by $d+1$ terms of $\varepsilon(\cdot)$; but from (20b) we have

$$\varepsilon(t+1+j) = \mathbf{e}_1^\top\zeta(t+1+j), \quad j = 0, 1, \ldots, d. \tag{21}$$

With this in mind, we construct the following $(n'(d+1))$th-order system driven by $\overline{\varepsilon}(\cdot)$:

$$\begin{bmatrix}\zeta(t+d+2) \\ \zeta(t+d+1) \\ \vdots \\ \vdots \\ \zeta(t+2)\end{bmatrix} = \underbrace{\begin{bmatrix} A_l & & & & \\ I_{n'} & & & & \\ & I_{n'} & & & \\ & & \ddots & & \\ & & & I_{n'} & 0 \end{bmatrix}}_{=:\widetilde{A}_l} \underbrace{\begin{bmatrix}\zeta(t+d+1) \\ \zeta(t+d) \\ \vdots \\ \vdots \\ \zeta(t+1)\end{bmatrix}}_{=:\overline{\zeta}(t)} + \mathbf{e}_1\overline{\varepsilon}(t+d+2). \tag{22}$$

At this point we can combine the models (18) and (22) together with the linking equation (21) to obtain a model driven by the exogenous inputs and $\overline{\varepsilon}$ (rather than ε): with

$$\eta(t) := \frac{1}{b_0}\mathbf{e}_{n+1}\left[\sum_{i=0}^{d} a_{d-i}y^*(t+1+i) - w(t+d+1)\right] + \mathbf{e}_1 y^*(t+1), \tag{23}$$

it follows that there exists a matrix $\tilde{B} \in \mathbb{R}^{(n+m+d)\times(n'(d+1))}$, which depends continuously on $\theta \in \mathcal{S}_{ab}$, to obtain the following $(n+m+d+n'(d+1))$th-order system:

$$\begin{bmatrix}\phi(t+1) \\ \overline{\zeta}(t+1)\end{bmatrix} = \underbrace{\begin{bmatrix} A_g & \tilde{B} \\ \mathbf{0} & \tilde{A}_l \end{bmatrix}}_{=:\tilde{A}_g} \underbrace{\begin{bmatrix}\phi(t) \\ \overline{\zeta}(t)\end{bmatrix}}_{=:\overline{\phi}(t)} + \eta(t) + \mathbf{e}_{n+m+d+1}\overline{\varepsilon}(t+d+2). \tag{24}$$

Before presenting our main model suitable for analysis, we need to analyze a couple of crude models of the closed-loop behavior that we use to bound the size of the growth of $\phi(t)$ and the size of the growth of $\overline{\phi}(t)$ in terms of the exogenous inputs. We use (1) to describe $y(t+1)$, and use the control law (13) together with the obtained equation for $y(t+1)$ to describe $u(t+1)$; we can appropriately define matrices $A_1(t)$, $B_1(t)$ and $B_2(t)$ in terms of $\theta \in \mathcal{S}_{ab}$ and $\hat{\theta}(t+1) \in \mathcal{S}$ so that we have the following **crude model on the behavior of** $\phi(\cdot)$:

$$\begin{aligned}\phi(t+1) &= A_1(t)\phi(t) + \\ &B_1(t)\overline{y}^*(t+d+1) + B_2(t)w(t+1), \quad t \ge t_0.\end{aligned} \tag{25}$$

Furthermore, we combine (25) and (22) to obtain an equation for $\overline{\phi}(t)$: we can appropriately define matrices $B_3(t), B_4(t)$ so that

$$\begin{aligned}\overline{\phi}(t+1) &= \begin{bmatrix} A_1(t) & \mathbf{0} \\ \mathbf{0} & \tilde{A}_l \end{bmatrix}\overline{\phi}(t) + B_3(t)\overline{y}^*(t+d+1) + \\ &B_4(t)w(t+1) + \mathbf{e}_{n+m+d+1}\overline{\varepsilon}(t+d+2), \quad t \ge t_0.\end{aligned} \tag{26}$$

Now we want to find a representation for $\overline{\varepsilon}(t+d+2)$ in the RHS above in terms of $\phi(t)$: from (15) we have $\overline{\varepsilon}(t+d+2) = -\tilde{\theta}(t+2)^\top\phi(t+2) + \overline{w}(t+2)$, so we use (25) to find a representation of $\phi(t+2)$ in terms of $\phi(t)$ and substitute into (26); then we can appropriately define matrices $A_2(t), B_5(t), B_6(t), B_7(t), B_8(t)$ so that the following **crude model on the behavior of** $\overline{\phi}(\cdot)$ is obtained:

$$\begin{aligned}\overline{\phi}(t+1) &= A_2(t)\overline{\phi}(t) + B_5(t)\overline{y}^*(t+d+1) + \\ &B_6(t)w(t+1) + B_7(t)\overline{y}^*(t+d+2) + \\ &B_8(t)w(t+2) + \mathbf{e}_{n+m+d+1}\overline{w}(t+2), \quad t \ge t_0.\end{aligned} \tag{27}$$

Proposition 2. *There exists a constant* $c_1 \ge 1$ *such that for*

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

every $t_0 \in \mathbb{Z}$, initial condition $x(t_0)$, $\theta_0 \in \mathcal{S}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, and $g > 0$, when the adaptive controller (12) and (13) is applied to the plant (1), the following holds:

$$\begin{aligned}
&\|A_1(t)\| \le c_1, \ \ \|A_2(t)\| \le c_1, \ \ \|B_1(t)\| \le c_1,\\
&\|B_2(t)\| \le c_1, \ \ \|B_5(t)\| \le c_1, \ \ \|B_6(t)\| \le c_1,\\
&\|B_7(t)\| \le c_1, \ \ \|B_8(t)\| \le c_1, \ \ t \ge t_0.
\end{aligned}$$

Proof. This result is straightforward to prove. From the construction of the crude models in (25) and (27) described above, for each fixed t the matrices $A_1(t)$, $A_2(t)$, $B_1(t)$, $B_2(t)$ $B_3(t)$, $B_4(t)$, $B_5(t)$, $B_6(t)$, $B_7(t)$, and $B_8(t)$ are defined either in terms of $\theta \in \mathcal{S}_{ab}$, $\hat{\theta}(\cdot) \in \mathcal{S}$, the coefficients of the reference model polynomial $\mathbf{L}(z^{-1})$ or a combination thereof. So since $\mathcal{S}_{ab}, \mathcal{S}$ (itself related to $\mathcal{S}_{\alpha\beta}$) are compact, then the constant c_1 exists. ∎

Finally, we now present a model which describes the evolution of $\phi(\cdot)$ which we use in analyzing the closed-loop behavior. The *good* closed-loop model (24) is driven by a future value of $\overline{\varepsilon}(\cdot)$. We now combine it with the crude model (27) to obtain a new model which is driven by perturbed version of $\overline{\phi}(t)$, with weights associated with the parameter estimation updates. Before proceeding, motivated by the form of the term in the parameter estimator, and as long as $\|\phi(t-d+1)\| \neq 0$, we define

$$\nu(t) := \frac{\phi(t-d+1)}{\|\phi(t-d+1)\|^2} e(t+1).$$

Also, for ease of notation, let us define

$$\tilde{w}(t) := \sum_{j=1}^{\max\{3,d+1\}} \big(|y^*(t+j)| + |\overline{y}^*(t+d+j)| + |w(t+j)| + |\overline{w}(t+j)|\big).$$

Proposition 3. *There exists a constant c_2 so that for every $t_0 \in \mathbb{Z}$, initial condition $x(t_0)$, $\theta_0 \in \mathcal{S}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, and $g > 0$, when the adaptive controller (12) and (13) is applied to the plant (1), on every interval of the form $[\underline{t}, \overline{t}) \subset [t_0, \infty)$ which satisfies $\overline{t} - \underline{t} > d+1$ and*

$$\|\phi(t-d+1)\|^2 \ge g, \quad t \in [\underline{t}, \overline{t}),$$

it is also true that for every $t \in [\underline{t}, \overline{t}-d-1)$, we have

$$\overline{\phi}(t+1) = [\widetilde{A}_g + \Delta(t)]\overline{\phi}(t) + \overline{\eta}(t), \tag{28}$$

with

$$\|\Delta(t)\| \le c_2 \sum_{j=2}^{d+1} \|\nu(t+j)\| \tag{29}$$

and

$$\|\overline{\eta}(t)\| \le c_2 \left(1 + \sum_{j=2}^{d+1} \|\nu(t+j)\|\right) \tilde{w}(t). \tag{30}$$

Proof. See the Appendix. ∎

Notice that the matrix $\widetilde{A}_g$ is a function of $\theta \in \mathcal{S}_{ab}$ and the coefficients of $\mathbf{L}(z^{-1})$; it lies in a corresponding compact set $\mathcal{A} \subset \mathbb{R}^{(n+m+d+n'(d+1))\times(n+m+d+n'(d+1))}$. Furthermore, the eigenvalues of $\widetilde{A}_g$ are at the origin, the roots of $\mathbf{L}(z^{-1})$, and the roots of $\mathbf{B}(z^{-1})$, so they are all in the open unit disk; so we can use classical arguments to prove that for the desired reference model there exists constants γ and $\sigma \in (0,1)$ so that for all $\theta \in \mathcal{S}_{ab}$, we have

$$\|\widetilde{A}_g^k\| \le \gamma \sigma^k, \quad k \ge 0. \tag{31}$$

Indeed, we can choose any σ larger than

$$\underline{\lambda} := \max_{\theta \in \mathcal{S}_{ab}} \big\{|\lambda| : \lambda \in \mathbb{C}, \mathbf{B}(\lambda^{-1}) = 0 \text{ and } \mathbf{L}(\lambda^{-1}) = 0\big\}.$$

Equations of the form given in (28) appear in classical adaptive control approaches. While we can view (28) as a linear time-varying system, we have to keep in mind that $\Delta(t)$ and $\overline{\eta}(t)$ are implicit nonlinear functions of θ, θ_0, $x(t_0)$, r and w. However, this linear time-varying interpretation is very convenient for analysis; to this end, let $\mathbf{\Phi}_A$ denote the state transition matrix of a general time-varying square matrix A. The following result of Kreisselmeier's is useful in analyzing our closed-loop system.

Proposition 4 ([13]). *With $\sigma \in (\underline{\lambda}, 1)$, suppose that $\gamma \ge 1$ is such that (31) is satisfied for every $\widetilde{A}_g \in \mathcal{A}$. For every $\mu \in (\sigma, 1)$, $g_0 \ge 0$, $g_1 \ge 0$, and $g_2 \in \big[0, \frac{\mu-\sigma}{\gamma}\big)$, there exists a constant $\bar{\gamma} \ge 1$ so that for every $\widetilde{A}_g \in \mathcal{A}$ and $\Delta \in \mathbb{S}\big(\mathbb{R}^{(n+m+d+n'(d+1))\times(n+m+d+n'(d+1))}\big)$ satisfying*

$$\sum_{j=\tau}^{t-1} \|\Delta(j)\| \le g_0 + g_1(t-\tau)^{\frac{1}{2}} + g_2(t-\tau), \ \ \overline{t} \ge t > \tau \ge \underline{t},$$

we have $\|\mathbf{\Phi}_{\widetilde{A}_g+\Delta}(t,\tau)\| \le \bar{\gamma}\mu^{t-\tau}, \quad \overline{t} \ge t > \tau \ge \underline{t}$.

Next, we present the main result proving that the closed-loop system enjoys very desirable linear-like behavior.

IV. The Main Result

We now present the main result of this paper. We first show a desirable bound on the closed-loop behavior in Theorem 1. Afterwards, a tracking result is shown in Theorem 2.

Theorem 1. *For every $\lambda \in (\underline{\lambda}, 1)$, there exists a constant $\gamma > 0$ so that for every $t_0 \in \mathbb{Z}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, $\theta_0 \in \mathcal{S}$, plant initial condition $x(t_0)$, and $g > 0$: when the adaptive controller (12) and (13) is applied to the plant (1), the following bound holds:*

$$\|\phi(t)\| \le \gamma\lambda^{t-t_0}\|\phi(t_0)\| + \gamma\sqrt{g} + \sum_{j=t_0}^{t-1} \gamma\lambda^{t-j-1}(|r(j)| + |w(j+1)|) + \gamma|r(t)|, \quad t \ge t_0. \tag{32}$$

Remark 3. *The above result shows that the closed-loop system experiences linear-like behavior. In particular, we see from the above theorem that the bound on ϕ includes 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*

$$\|\phi(t)\| \le \tfrac{\gamma}{1-\lambda}\left(\lambda^{t-t_0}\|x(t_0)\| + \|w\|_\infty + \|r\|_\infty + \sqrt{g}\right), \quad t \ge t_0.$$

In comparison, in the literature it is proven only that ϕ is bounded, e.g. see [36], [37]. Furthermore, this provides a clear bound on $\|\phi(t)\|$ in terms of the parameter g—the smaller that it is, the smaller that this bound is. Hence, we might expect that a smaller g may yield a smaller amount of bursting; this is consistent with the view that a smaller g makes the estimator more responsive to non-zero prediction errors.

Remark 4. *In our earlier work on the MRAC problem [32] which uses the ideal projection algorithm (11), we prove a bound of the same form as in (32) but with the '$\gamma\sqrt{g}$' term removed. On the one hand, this makes intuitive sense in that the ideal projection algorithm can be interpreted as the limiting case of the classical estimator (12a) as $g \to 0$.*

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

On the other hand, the closed-loop system is highly nonlinear, so it is not at all obvious that this should be the case.

To prove Theorem 1 we first prove a result on the system behavior when $\|\phi(\cdot)\|^2 \geq g$. The proof of this is a significant extension of the proof of the main theorem of our conference paper [32] which deals with the use of the **ideal estimator** for analysis on $[t_0, \infty)$.

Proposition 5. *For every $\lambda \in (\underline{\lambda}, 1)$, there exists a constant c so that for every $t_0 \in \mathbb{Z}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, $\theta_0 \in \mathcal{S}$, plant initial condition $x(t_0)$, and estimator constant $g > 0$, when the adaptive controller (12) and (13) is applied to the plant (1), for every interval of the form $[\underline{t}, \bar{t}] \subset [t_0, \infty)$ which satisfies*

$$\|\phi(t-d+1)\|^2 \geq g, \quad t \in [\underline{t}, \bar{t}),$$

it is also the case that

$$\|\phi(t)\| \leq c\lambda^{t-\underline{t}}\|\phi(\underline{t})\| + c\sum_{j=\underline{t}}^{t-1} \lambda^{t-j-1}\tilde{w}(j) \tag{33}$$

for $t = \underline{t}, \underline{t}+1, \ldots, \bar{t}-1, \bar{t}$.

Proof. See the Appendix. ■

Proof of Theorem 1. Fix $\lambda \in (\underline{\lambda}, 1)$. Let $t_0 \in \mathbb{Z}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, $\theta_0 \in \mathcal{S}$, $x(t_0) \in \mathbb{R}^{n+m+3d-2}$ and $g > 0$ be arbitrary.

First, define the region where $\phi(\cdot)$ is large:

$$S_{\mathsf{L}} := \{t \geq t_0 : \|\phi(t)\|^2 \geq g\},$$

and the region where $\phi(\cdot)$ is small:

$$S_{\mathsf{S}} := \{t \geq t_0 : \|\phi(t)\|^2 < g\}.$$

Observe that this partition clearly depends on θ_0, θ^*, $x(t_0)$, r, and w. We will apply Proposition 5 to analyze the closed-loop behavior on sub-intervals of S_{L}; we will analyze the closed-loop behavior on S_{S} in a direct manner. Before doing so, we we partition the timeline into intervals which oscillate between S_{L} and S_{S}. To this end, it is easy to see that we can define a (possibly infinite) sequence of intervals of the form $[k_l, k_{l+1})$ satisfying: (i) $k_0 = t_0$; (ii) $[k_l, k_{l+1})$ either belongs to S_{L} or S_{S}; and (iii) if $k_{l+1} \neq \infty$ and $[k_l, k_{l+1})$ belongs to S_{L} (respectively, S_{S}), then the interval $[k_{l+1}, k_{l+2})$ must belong to S_{S} (respectively, S_{L}).

Step 1: Behavior on S_{S}.

Let $[k_l, k_{l+1}) \subset S_{\mathsf{S}}$ be arbitrary; then we clearly have

$$\|\phi(t)\| \leq \sqrt{g}, \quad t \in [k_l, k_{l+1});$$

by using (25) and applying Proposition 2, we have

$$\begin{aligned}\|\phi(k_{l+1})\| &\leq c_1\big(\|\phi(k_{l+1}-1)\| + |\overline{y}^*(k_{l+1}+d)| + |w(k_{l+1})|\big) \\ &\leq c_1\big(\sqrt{g} + |\overline{y}^*(k_{l+1}+d)| + |w(k_{l+1})|\big).\end{aligned}$$

Summarizing the above, we obtain:

$$\|\phi(t)\| \leq \begin{cases} \sqrt{g}, & t \in [k_l, k_{l+1}) \\ c_1\big(\sqrt{g} + \tilde{w}(t-1)\big), & t = k_{l+1}. \end{cases} \tag{34}$$

Step 2: Behavior on S_{L}.

Let $[k_l, k_{l+1}) \subset S_{\mathsf{L}}$ be arbitrary. First, if $k_{l+1} - k_l < d+1$, then using the crude model on ϕ in (25) and applying Proposition 2, if we define $\bar{\gamma}_1 := \left(\frac{c_1}{\lambda}\right)^{d+1}$, then we have

$$\begin{aligned}&\|\phi(t)\| \leq \bar{\gamma}_1\lambda^{t-k_l}\|\phi(k_l)\| + \\ &\sum_{j=\underline{t}}^{t-1} \bar{\gamma}_1\lambda^{t-j-1}(|\overline{y}^*(j+d+1)| + |w(j+1)|), \quad t \in [k_l, k_{l+1}].\end{aligned} \tag{35}$$

Now suppose that $k_{l+1} - k_l \geq d+1$. By Proposition 5, we know that there exists a constant c so that for every interval $[\underline{t}, \bar{t}] \subset [t_0, \infty)$ for which

$$\|\phi(t-d+1)\|^2 \geq g, \quad t \in [\underline{t}, \bar{t}), \tag{36}$$

then we have

$$\|\phi(t)\| \leq c\lambda^{t-\underline{t}}\|\phi(\underline{t})\| + \sum_{j=\underline{t}}^{t-1} c\lambda^{t-j-1}\tilde{w}(t), \quad t \in [\underline{t}, \bar{t}]. \tag{37}$$

If we set $\underline{t} = k_l + d - 1$ and $\bar{t} = k_{l+1} + d - 2$, then it follows from (36) and (37) that

$$\begin{aligned}\|\phi(t)\| \leq c\lambda^{t-k_l}\|\phi(k_l)\| &+ \sum_{j=k_l}^{t-1} c\lambda^{t-j-1}\tilde{w}(j), \\ t \in [k_l + d - 1, k_{l+1} &+ d - 2].\end{aligned} \tag{38}$$

Notice that we need to find a bound for the rest of the interval, namely for $t \in [k_l, k_l + d - 1) \cup (k_{l+1} + d - 2, k_{l+1}]$. To this end, we use the crude model (25), apply Proposition 2 to each interval, and use the above bound: we conclude that there exists a constant $\bar{\gamma}_2$ such that

$$\begin{aligned}\|\phi(t)\| \leq \bar{\gamma}_2\lambda^{t-k_l}\|\phi(k_l)\| &+ \bar{\gamma}_2\sum_{j=k_l}^{t-1} \lambda^{t-j-1}\tilde{w}(j), \\ t \in [k_l, k_{l+1}].&\end{aligned} \tag{39}$$

Step 3: Behavior on the whole timeline.

Claim 1. *There exists a constant $\bar{\gamma}$ so that the following bound holds:*

$$\|\phi(t)\| \leq \bar{\gamma}\lambda^{t-t_0}\|\phi(t_0)\| + \bar{\gamma}\sum_{j=t_0}^{t-1} \lambda^{t-j-1}\tilde{w}(j) + \bar{\gamma}\sqrt{g}, \quad t \geq t_0. \tag{40}$$

Proof of Claim 1. If $[k_0, k_1) = [t_0, k_1) \subset S_{\mathsf{L}}$, then (40) holds for $[t_0, k_1]$ by (39) as long as $\bar{\gamma} \geq \bar{\gamma}_2$.

If $[k_0, k_1) = [t_0, k_1) \subset S_{\mathsf{S}}$, then from (34) we see that (40) holds as long as $\bar{\gamma} \geq c_1$.

We now use induction; suppose that (40) holds for $[k_0, k_l]$; we need to prove that it holds for $t \in (k_l, k_{l+1}]$ as well. If $[k_l, k_{l+1}) \subset S_{\mathsf{S}}$, then from (34) we see that (40) holds on $(k_l, k_{l+1}]$ also as long as $\bar{\gamma} \geq c_1$. On the other hand, if $[k_l, k_{l+1}) \subset S_{\mathsf{L}}$, then $k_l - 1 \in S_{\mathsf{S}}$; from (34) we have that

$$\|\phi(k_l)\| \leq c_1\big(\sqrt{g} + \tilde{w}(k_l - 1)\big);$$

combining this with (39), we have

$$\begin{aligned}\|\phi(t)\| &\leq \bar{\gamma}_2\lambda^{t-k_l}\|\phi(k_l)\| + \bar{\gamma}_2\sum_{j=k_l}^{t-1} \lambda^{t-j-1}\tilde{w}(j) \\ &\leq \bar{\gamma}_2 c_1\lambda^{t-k_l}\big(\sqrt{g} + \tilde{w}(k_l - 1)\big) + \bar{\gamma}_2\sum_{j=k_l}^{t-1} \lambda^{t-j-1}\tilde{w}(j) \\ &\leq c_1(\bar{\gamma}_2 + 1)\sum_{j=k_l-1}^{t-1} \lambda^{t-j-1}\tilde{w}(j) + \bar{\gamma}_2 c_1\sqrt{g}, \\ &\quad t \in [k_l, k_{l+1}].\end{aligned}$$

So the bound (40) holds as long as $\bar{\gamma} \geq c_1(\bar{\gamma}_2 + 1)$. □

The last step is to convert the bound proven in Claim 1 to one of the desired form, i.e. we need to replace $\tilde{w}$ with w and r. Observe that from its definition and the reference model, $\tilde{w}(t)$ is a weighted sum of $\{|w(t+1)|, |w(t+2)|, \ldots, |w(t+$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

$d + \max\{3, d+1\})|, |r(t-n'+d+1)|, |r(t-n'+d+2)|, \ldots, |r(t)|, |r(t+1)|, \ldots, |r(t+d+\max\{3, d+1\})|\}$. Using a causality argument similar to that used in the proof of Theorem 1 of [22], we simplify (40) and remove extraneous terms to end up with the desired bound. ∎

Next, we present a result on tracking.

Theorem 2. *For every $\lambda \in (\underline{\lambda}, 1)$, there exists a constant $\bar{c} > 0$ so that for every $t_0 \in \mathbb{Z}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, $\theta_0 \in \mathcal{S}$, $x(t_0)$, and $g > 0$, when the adaptive controller (12) and (13) is applied to the plant (1): (a) if $w = 0$, then the following holds:*

$$\sum_{k=t_0+d}^{\infty} \varepsilon(k)^2 \le \bar{c}(\|x(t_0)\|^2 + \|r\|_\infty^2 + g); \tag{41}$$

(b) if w is non-zero, then the following holds:

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} \varepsilon(j)^2 \le \limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} w(j)^2 \times \bar{c} \max\left\{1, \frac{\limsup_{k\to\infty} |w(k)|^2 + \limsup_{k\to\infty} |r(k)|^2}{g}\right\}. \tag{42}$$

Remark 5. *In the absence of noise, most adaptive controllers guarantee only that the tracking error is square summable, e.g. see [10]. Here we prove a stronger result, namely, an upper bound on the 2-norm in terms of the size of $x(t_0)$ and r.*

Remark 6. *When noise is entering the system, we prove that the average power of the tracking error is bounded by the average power of the disturbance with a coefficient depending on the value of the estimator constant g. The bound in Theorem 2 is reminiscent of the bounds proven in [25]. In particular, Theorem 2 says that if $\|r\|_\infty + \|w\|_\infty \le 1$, then*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} \varepsilon(j)^2 \le \bar{c}(1+\tfrac{1}{g}) \limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} w(j)^2.$$

Proof of Theorem 2. Fix $\lambda \in (\underline{\lambda}, 1)$. Let $t_0 \in \mathbb{Z}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, $\theta_0 \in \mathcal{S}$, $g > 0$ and $x(t_0) \in \mathbb{R}^{n+m+3d-2}$ be arbitrary.

From (17), the first property in part (i) of Proposition 1, and the Cauchy-Schwarz inequality, we can easily obtain

$$\frac{\bar{\varepsilon}(t)^2}{g + \|\phi(t-d)\|^2} \le d \sum_{j=0}^{d-1} \frac{e(t-j)^2}{g + \|\phi(t-j-d)\|^2}, \quad t \ge t_0 + d;$$

so by the second property of part (i) of Proposition 1 we see that for $\tau \ge t_0 + d$ we obtain

$$\begin{aligned} \sum_{j=\tau}^{t} \frac{\bar{\varepsilon}(j)^2}{g + \|\phi(j-d)\|^2} &\le d^2 \sum_{j=\tau-d+1}^{t} \frac{e(j)^2}{g + \|\phi(j-d)\|^2} \\ &\le 8d^2\|\mathcal{S}\|^2 + 4d^2 \sum_{j=\tau-d+1}^{t} \frac{\overline{w}(j-d)^2}{g + \|\phi(j-d)\|^2} \\ &\le 8d^2\|\mathcal{S}\|^2 + \frac{4d^2}{g} \sum_{j=\tau-d+1}^{t} \overline{w}(j-d)^2. \end{aligned} \tag{43}$$

Then we can obtain

$$\sum_{j=\tau}^{t} \bar{\varepsilon}(j)^2 \le (g + \sup_{q\in[\tau,t]} \|\phi(q-d)\|^2) \times \left(8d^2\|\mathcal{S}\|^2 + \frac{4d^2}{g} \sum_{j=\tau-d+1}^{t} \overline{w}(j-d)^2\right).$$

But ε and $\bar{\varepsilon}$ are related by a stable transfer function (see (19)). If we apply Parseval's Theorem to (20) and use the fact that A_l is stable, we see that there exists a constant $\bar{c}_0$ so that

$$\sum_{j=\tau}^{t} \varepsilon(j)^2 \le \bar{c}_0\left(\|\zeta(\tau)\|^2 + \sum_{j=\tau+1}^{t+1} \bar{\varepsilon}(j)^2\right), \quad \tau \ge t_0 + d. \tag{44}$$

Using that fact that $\zeta(\tau) = \begin{bmatrix}\varepsilon(\tau) & \varepsilon(\tau-1) & \cdots & \varepsilon(\tau-n'+1)\end{bmatrix}^\top = \begin{bmatrix}y(\tau)-y^*(\tau) & y(\tau-1)-y^*(\tau-1) & \cdots & y(\tau-n'+1)-y^*(\tau-n'+1)\end{bmatrix}^\top$, the definition of $\phi(\cdot)$, and that fact that $n' \le n$ by assumption, we see that

$$\|\zeta(\tau)\|^2 \le 2\|\phi(\tau)\|^2 + 2n\|y^*\|_\infty^2, \quad \tau \ge t_0 + d.$$

Combining this with (44), we see that for $\tau \ge t_0 + d$:

$$\sum_{j=\tau}^{t} \varepsilon(j)^2 \le 2n\bar{c}_0\left(\|\phi(\tau)\|^2 + \|y^*\|_\infty^2 + \sum_{j=\tau+1}^{t+1} \bar{\varepsilon}(j)^2\right). \tag{45}$$

If we now combine this with the bound on $\sum_{j=\tau}^{t} \bar{\varepsilon}(j)^2$ given above (suitably delayed by one step), and since y^* and r are related by a stable transfer function (the reference model (3)), we see that there exists a constant $\bar{c}_1$ so that for $\tau \ge t_0 + d$:

$$\sum_{j=\tau}^{t} \varepsilon(j)^2 \le \bar{c}_1\Bigg(\|\phi(\tau)\|^2 + \|r\|_\infty^2 + (g + \sup_{q\in[\tau+1,t+1]} \|\phi(q-d)\|^2) \times \left(8d^2\|\mathcal{S}\|^2 + \frac{4d^2}{g} \sum_{j=\tau-d+2}^{t+1} \overline{w}(j-d)^2\right)\Bigg). \tag{46}$$

By Theorem 1 there exists a constants $\bar{c}_2 \ge 1$ such that

$$\|\phi(t)\| \le \bar{c}_2\lambda^{t-t_0}\|\phi(t_0)\| + \bar{c}_2\sqrt{g} + \bar{c}_2 \sum_{j=t_0}^{t-1} \lambda^{t-j-1}(|w(j+1)| + |r(j)|) + \bar{c}_2|r(t)|, \quad t \ge t_0. \tag{47}$$

If $w = 0$, then it is easy to see that there exists a constant $\bar{c}_3$ so that

$$\sum_{t=t_0+d}^{\infty} \varepsilon(t)^2 \le \bar{c}_3(\|x(t_0)\|^2 + \|r\|_\infty^2 + g),$$

which yields the desired bound (41).

Next we prove the bound (42) for the case when w is not necessarily zero at all times. By (47) we can choose $\tilde{t} \ge t_0 + d$, implicitly depending on $r, w, x(t_0), \theta_0, \theta^*$, so that

$$\|\phi(t-d)\| \le \frac{2\bar{c}_2}{1-\lambda}(\sqrt{g} + \limsup_{k\to\infty} |w(k+1)| + \limsup_{k\to\infty} |r(k)|), \quad t \ge \tilde{t}. \tag{48}$$

Incorporating this into (46), there is exists a constant $\bar{c}_4$ so that

$$\sum_{j=\tilde{t}}^{t} \varepsilon(j)^2 \le \bar{c}_4\Bigg(\|\phi(\tilde{t})\|^2 + \|r\|_\infty^2 + [\limsup_{k\to\infty} |w(k+1)|^2 + \limsup_{k\to\infty} |r(k)|^2 + g] \times$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

$$\left(8d^2\|\mathcal{S}\|^2 + \frac{4d^2}{g}\sum_{j=\tilde{t}-d+2}^{t+1}\overline{w}(j-d)^2\right)\Bigg).$$

This means that

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=\tilde{t}}^{\tilde{t}+T-1}\varepsilon(j)^2 \le \limsup_{T\to\infty}\frac{1}{T}\sum_{j=\tilde{t}-d+1}^{\tilde{t}+T-1}\overline{w}(j-d)^2\times$$
$$4d^2\bar{c}_4\left(1+\frac{\limsup_{k\to\infty}|w(k)|^2+\limsup_{k\to\infty}|r(k)|^2}{g}\right).$$

But $\overline{w}(t)$ is a weighted sum of $\{w(t+1),\ldots,w(t+d)\}$, and the boundedness of all variables makes the starting point of the average sums irrelevant, so after simplification the desired bound follows. ■

V. Robustness Results

It turns out that the convolution bounds proven in Theorem 1 will guarantee robustness to a degree of time-variations and unmodelled dynamics.

A. Tolerance to Time-variation

The linear-like bound proven in Theorem 1 can be leveraged to prove similar behavior in the presence of slow time-variations and/or occasional jumps. To proceed, we first rewrite the the original plant model in (1) to incorporate the vector $\phi(t)$ by padding θ with zeros in the obvious spots and labeling it $\overline{\theta}$. The time-varying version can now be written as

$$y(t+1)=\phi(t)^\top\overline{\theta}(t)+w(t+1). \tag{49}$$

We define $\overline{\mathcal{S}_{ab}}\subset\mathbb{R}^{m+n+d}$ to represent the padded elements of $\mathcal{S}_{ab}$, which is clearly compact. We adopt a common model of time-variations used in adaptive control, e.g. see [13].

Definition 1. For $c_0\ge 0$ and $\epsilon>0$, let $\mathfrak{s}(\overline{\mathcal{S}_{ab}},c_0,\epsilon)$ denote the subset of $\boldsymbol{\ell}(\mathbb{R}^{n+m+d})$ whose elements $\overline{\theta}$ satisfy $\overline{\theta}(t)\in\overline{\mathcal{S}_{ab}}$ for every $t\in\mathbb{Z}$ and

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

We now show that for every $c_0\ge 0$, the presented MRAC approach tolerates time-varying parameters if ϵ is small enough.

Theorem 3. *For every $\lambda_1\in(\underline{\lambda},1)$ and $c_0\ge 0$, there exists a $\bar{\gamma}_1>0$ and $\bar{\epsilon}>0$ so that for every $t_0\in\mathbb{Z}$, $\theta_0\in\mathcal{S}$, $\epsilon\in[0,\bar{\epsilon}]$, $\overline{\theta}\in\mathfrak{s}(\overline{\mathcal{S}_{ab}},c_0,\epsilon)$, initial condition $x(t_0)$, $r,w\in\boldsymbol{\ell}_\infty$, and $g>0$, when the adaptive controller (12) and (13) is applied to the time-varying plant (49), the following holds:*

(i) $\|\phi(t)\|\le\bar{\gamma}_1\lambda_1^{t-t_0}\|\phi(t_0)\|+\bar{\gamma}_1\sqrt{g}+$
$$\sum_{j=t_0}^{t-1}\bar{\gamma}_1\lambda_1^{t-j-1}(|r(j)|+|w(j+1)|)+\bar{\gamma}_1|r(t)|,\quad t\ge t_0,$$

(ii) $|\varepsilon(t)|\le\bar{\gamma}_1\lambda_1^{t-t_0}\|\phi(t_0)\|+\bar{\gamma}_1\sqrt{g}+$
$$\sum_{j=t_0}^{t}\bar{\gamma}_1\lambda_1^{t-j}(|r(j)|+|w(j)|),\qquad t\ge t_0+1.$$

(iii) If $c_0=0$ and $w=0$, then

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=t_0}^{t_0+T-1}\varepsilon(j)^2\le$$
$$\bar{\gamma}_1\sqrt{\epsilon}\left(g+\limsup_{t\to\infty}|r(t)|^2+\tfrac{1}{g}\limsup_{t\to\infty}|r(t)|^4\right).$$

(iv) If $c_0=0$, $w=0$, and $d=1$, then

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{j=t_0}^{t_0+T-1}\varepsilon(j)^2\le\bar{\gamma}_1\sqrt{\epsilon}\left(g+\limsup_{t\to\infty}|r(t)|^2\right).$$

Remark 7. *Theorem 3(ii) provides a weak bound on the size of the tracking error. Theorem 3(iii) and 3(iv) provide an asymptotic bound on the average error in the absence of a disturbance and where there are no large jumps in the plant parameter variation; observe that if $d=1$ then the bound is especially nice.*

Proof of Theorem 3. Fix $\lambda_1\in(\underline{\lambda},1)$, $\lambda\in(\underline{\lambda},\lambda_1)$ and $c_0\ge 0$; define $\bar{c}_0:=\max\{1,c_0\}$. Let $t_0\in\mathbb{Z}$, $\theta_0\in\mathcal{S}$, x_0, and $r,w\in\boldsymbol{\ell}_\infty$ be arbitrary. With $\varpi\in\mathbb{N}$, we will consider $\phi(t)$ on intervals of the form $[t_0+i\varpi,t_0+(i+1)\varpi]$; we will be analyzing these intervals in groups of ϖ (to be chosen shortly); we set $\bar{\epsilon}=\frac{\bar{c}_0}{\varpi^2}$, and let $\epsilon\in[0,\bar{\epsilon}]$ and $\overline{\theta}\in\mathfrak{s}(\overline{\mathcal{S}_{ab}},c_0,\epsilon)$ be arbitrary.

First of all, for $i\in\mathbb{Z}^+$ we can rewrite the plant equation (49) as

$$y(t+1)=\phi(t)^\top\overline{\theta}(t_0+i\varpi)+ w(t+1)+\underbrace{\phi(t)^\top[\overline{\theta}(t)-\overline{\theta}(t_0+i\varpi)]}_{=:\tilde{n}_i(t)},\quad t\in\mathbb{Z}. \tag{50}$$

This describes an LTI plant; since $\overline{\theta}(t_0+i\varpi)\in\overline{\mathcal{S}_{ab}}$, by Theorem 1 there exists a constant $c>0$, independent of i, so that

$$\|\phi(t)\|\le c\lambda^{t-t_0-i\varpi}\|\phi(t_0+i\varpi)\|+c\sqrt{g}+\sum_{j=t_0+i\varpi}^{t-1}c\lambda^{t-j-1}(|r(j)|+|w(j+1)|+|\tilde{n}_i(j)|)+c|r(t)|,\quad t\in\mathbb{Z}. \tag{51}$$

We will now proceed to use this bound on $\phi(t)$ for $t\in[t_0+i\varpi,t_0+(i+1)\varpi]$.

The above is a difference inequality associated with a first order system; using this observation together with the fact that $c\ge 1$, we see that if we define

$$\psi(t+1)=\lambda\psi(t)+|r(t)|+|w(t+1)|+|\tilde{n}_i(t)|,\quad t\in[t_0+i\varpi,t_0+(i+1)\varpi),$$

with $\psi(t_0+i\varpi)=\|\phi(t_0+i\varpi)\|$, then

$$\|\phi(t)\|\le c\psi(t)+c(|r(t)|+\sqrt{g}),\quad t\in[t_0+i\varpi,t_0+(i+1)\varpi]. \tag{52}$$

Now we analyze this equation for $i=0,1,\ldots,\varpi-1$.

Case 1: $|\tilde{n}_i(t)|\le\frac{1}{2c}(\lambda_1-\lambda)\|\phi(t)\|$ for all $t\in[t_0+i\varpi,t_0+(i+1)\varpi]$.

In this case

$$\begin{aligned}\psi(t+1)&\le\lambda\psi(t)+|r(t)|+|w(t+1)|+|\tilde{n}_i(t)|\\&\le\lambda\psi(t)+|r(t)|+|w(t+1)|+\tfrac{1}{2c}(\lambda_1-\lambda)\|\phi(t)\|\\&\le\lambda\psi(t)+|r(t)|+|w(t+1)|+\\&\quad\tfrac{1}{2c}(\lambda_1-\lambda)\,c[\psi(t)+|r(t)|+\sqrt{g}]\\&\le(\tfrac{\lambda+\lambda_1}{2})\psi(t)+|w(t+1)|+2|r(t)|+(\tfrac{\lambda_1-\lambda}{2})\sqrt{g},\\&\qquad t\in[t_0+i\varpi,t_0+(i+1)\varpi],\end{aligned}$$

which means that

$$\psi(t)\le\left(\frac{\lambda+\lambda_1}{2}\right)^{t-t_0-i\varpi}\psi(t_0+i\varpi)+$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

$$\sum_{j=t_0+i\varpi}^{t-1} \left(\frac{\lambda+\lambda_1}{2}\right)^{t-1-j} (2|r(j)| + |w(j+1)| + (\tfrac{\lambda_1-\lambda}{2})\sqrt{g}),$$
$$t = t_0+i\varpi, \ldots, t_0+(i+1)\varpi.$$

This, in turn, implies that

$$\|\phi(t_0+(i+1)\varpi)\| \le c\left(\frac{\lambda+\lambda_1}{2}\right)^{\varpi}\|\phi(t_0+i\varpi)\| +$$
$$\sum_{j=t_0+i\varpi}^{t_0+(i+1)\varpi-1} c\left(\frac{\lambda+\lambda_1}{2}\right)^{t_0+(i+1)\varpi-1-j}(2|r(j)|+|w(j+1)|)+$$
$$+c\left[|r(t_0+(i+1)\varpi)| + \underbrace{\left(1+\left(\frac{1}{1-\frac{\lambda_1+\lambda}{2}}\right)\left(\tfrac{\lambda_1-\lambda}{2}\right)\right)}_{=:\overline{\gamma}}\sqrt{g}\right]. \tag{53}$$

Case 2: $|\tilde{n}_i(t)| > \frac{1}{2c}(\lambda_1-\lambda)\|\phi(t)\|$ for some $t \in [t_0+i\varpi, t_0+(i+1)\varpi]$.

Since $\overline{\theta}(t) \in \overline{\mathcal{S}_{ab}}$ for $t \ge t_0$, we see

$$|\tilde{n}_i(t)| \le 2\left\|\overline{\mathcal{S}_{ab}}\right\|\|\phi(t)\|, \quad t \in [t_0+i\varpi, t_0+(i+1)\varpi].$$

This means that

$$\begin{aligned}\psi(t+1) &\le \lambda\psi(t) + |r(t)| + |w(t+1)| + |\tilde{n}_i(t)| \\ &\le \lambda\psi(t) + |r(t)| + |w(t+1)| + \\ &\quad 2\|\overline{\mathcal{S}_{ab}}\|c(\psi(t)+|r(t)|+\sqrt{g}) \\ &\le \underbrace{2\left(1+c\left\|\overline{\mathcal{S}_{ab}}\right\|\right)}_{=:\gamma_1}[\psi(t)+|r(t)|+|w(t+1)|+\sqrt{g}], \\ &\qquad t \in [t_0+i\varpi, t_0+(i+1)\varpi],\end{aligned}$$

which means that

$$\psi(t) \le \gamma_1^{t-t_0-i\varpi}\psi(t_0+i\varpi) +$$
$$\sum_{j=t_0+i\varpi}^{t-1}\gamma_1^{t-j-1}(|r(j)|+|w(j+1)|+\sqrt{g}),$$
$$t = t_0+i\varpi, \ldots, t_0+(i+1)\varpi.$$

This, in turn, implies that

$$\begin{aligned}&\|\phi(t_0+(i+1)\varpi)\| \le c\gamma_1^{\varpi}\|\phi(t_0+i\varpi)\| + \\ &c\sum_{j=t_0+i\varpi}^{t_0+(i+1)\varpi-1}\gamma_1^{t_0+(i+1)\varpi-j-1}(|r(j)|+|w(j+1)|+\sqrt{g})+ \\ &c(|r(t_0+i\varpi)|+\sqrt{g}) \\ &\le c\gamma_1^{\varpi}\|\phi(t_0+i\varpi)\| + c\left(\tfrac{2\gamma_1}{\lambda+\lambda_1}\right)^{\varpi}\times \\ &\sum_{j=t_0+i\varpi}^{t_0+(i+1)\varpi-1}\left(\tfrac{\lambda+\lambda_1}{2}\right)^{t_0+(i+1)\varpi-j-1}(|r(j)|+|w(j+1)|)+ \\ &c|r(t_0+i\varpi)| + c\underbrace{(1+\varpi\gamma_1^{\varpi})}_{=:\overline{\gamma}_2(\varpi)}\sqrt{g}.\end{aligned} \tag{54}$$

On the interval $[t_0, t_0+\varpi^2]$ there are ϖ sub-intervals of length ϖ; furthermore, because of the choice of $\overline{\epsilon}$ we have that

$$\sum_{j=t_0}^{t_0+\varpi^2-1}\|\overline{\theta}(j+1)-\overline{\theta}(j)\| \le c_0+\epsilon\varpi^2 \le c_0+\overline{\epsilon}\varpi^2 \le 2\overline{c}_0.$$

Indeed, with a simple calculation, at most $N_1 := \frac{4c_0c}{\lambda_1-\lambda}$ sub-intervals fall into the category of *Case 2*, with the remaining number falling into the category of *Case 1*. Henceforth we assume that $\varpi > N_1$. Using (53) and (54) to analyze the behavior of the closed-loop system on the interval $[t_0, t_0+\varpi^2]$, we obtain the crude bound of

$$\begin{aligned}&\|\phi(t_0+\varpi^2)\| \le c^{\varpi}\gamma_1^{N_1\varpi}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{\varpi(\varpi-N_1)}\|\phi(t_0)\| + \\ &2\varpi\left(\tfrac{2\gamma_1}{\lambda+\lambda_1}\right)^{\varpi}(c\gamma_1^{\varpi+1})^{\varpi}\left(\tfrac{2}{\lambda+\lambda_1}\right)^{(\varpi+1)\varpi}\times \\ &\sum_{j=t_0}^{t_0+\varpi^2-1}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{t_0+\varpi^2-j-1}(|r(j)|+|w(j+1)|)+ \\ &\varpi(c\gamma_1^{\varpi})^{\varpi}[|r(t_0+\varpi^2)|+(\overline{\gamma}+\overline{\gamma}_2(\varpi))\sqrt{g}].\end{aligned} \tag{55}$$

At this point we would like to choose ϖ so that

$$c^{\varpi}\gamma_1^{N_1\varpi}\left(\tfrac{\lambda_1+\lambda}{2}\right)^{\varpi^2-\varpi N_1} \le \lambda_1^{\varpi^2} \Leftrightarrow$$
$$c^{\varpi}\gamma_1^{N_1\varpi}\left(\tfrac{2}{\lambda+\lambda_1}\right)^{\varpi N_1} \le \left(\tfrac{2\lambda_1}{\lambda_1+\lambda}\right)^{\varpi^2};$$

notice that $\frac{2\lambda_1}{\lambda_1+\lambda} > 1$, so if we take the log of both sides, we see that we need

$$\varpi\ln(c) + N_1\varpi\ln(\gamma_1) + N_1\varpi\ln\left(\tfrac{2}{\lambda+\lambda_1}\right) \le \varpi^2\ln\left(\tfrac{2\lambda_1}{\lambda_1+\lambda}\right),$$

which will clearly be the case for large enough ϖ, so at this point we choose such an ϖ. So it follows from (55) that there exists a constant γ_2 such that

$$\|\phi(t_0+\varpi^2)\| \le \lambda_1^{\varpi^2}\|\phi(t_0)\| + \gamma_2\sqrt{g}+$$
$$\gamma_2\sum_{j=t_0}^{t_0+\varpi^2-1}\lambda_1^{t_0+\varpi^2-j-1}(|r(j)|+|w(j+1)|)+\gamma_2|r(t_0+\varpi^2)|.$$

Indeed, by time-invariance of the closed-loop system we see that

$$\|\phi(\bar{t}+\varpi^2)\| \le \lambda_1^{\varpi^2}\|\phi(\bar{t})\| + \gamma_2\sqrt{g}+$$
$$\gamma_2\sum_{j=\bar{t}}^{\bar{t}+\varpi^2-1}\lambda_1^{\bar{t}+\varpi^2-j-1}(|r(j)|+|w(j+1)|)+\gamma_2|r(\bar{t}+\varpi^2)|,\ \bar{t} \ge t_0.$$

Solving iteratively and simplifying yield

$$\|\phi(t_0+i\varpi^2)\| \le \lambda_1^{i\varpi^2}\|\phi(t_0)\| + \tfrac{\gamma_2}{1-\lambda_1^{\varpi^2}}\sqrt{g}+$$
$$\gamma_2\sum_{j=t_0}^{t_0+i\varpi^2-1}\lambda_1^{t_0+i\varpi^2-j-1}(|r(j)|+|w(j+1)|)+$$
$$\tfrac{\gamma_2}{1-\lambda_1^{\varpi^2}}|r(t_0+i\varpi^2)|, \quad i \in \mathbb{Z}^+. \tag{56}$$

We now combine this bound with the bound (52), and the bounds (53) and (54) which hold on the intervals of Case 1 and the intervals of Case 2, respectively; we conclude that there exists a constant γ_3 so that

$$\|\phi(t)\| \le \gamma_3\lambda_1^{t-t_0}\|\phi(t_0)\| + \gamma_3\sqrt{g}+$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

$$\gamma_3 \sum_{j=t_0}^{t-1} \lambda_1^{t-j-1}(|r(j)| + |w(j+1)|) + \gamma_3 |r(t)|, \; t \geq t_0, \quad (57)$$

as desired. This finishes the proof of the stability result in Part (i) of the Theorem.

To prove Part (ii), we first observe that

$$\varepsilon(t+1) = \phi(t)^\top \overline{\theta}(t) + w(t+1) - y^*(t+1), \quad t \geq t_0.$$

Using the facts that $\overline{\theta}$ is uniformly bounded and y^* is a filtered version of r, together with the bound on $\phi(t)$ in Part (i), the bound on $\varepsilon(t)$ follows.

We now turn to proving the tracking result in Part (iii). We consider the special case of a zero disturbance ($w = 0$) and when there are no jumps in the plant parameters ($c_0 = 0$). With $\overline{\epsilon}$ as chosen above, define $\widetilde{\epsilon} := \min\{\overline{\epsilon}, \frac{1}{4d^2}\}$, and let $\epsilon \in (0, \widetilde{\epsilon}]$ be arbitrary. We will analyze the closed-loop system on intervals of length

$$N_\epsilon := \left\lceil \tfrac{1}{\sqrt{\epsilon}} \right\rceil;$$

observe that[3]

$$\frac{1}{N_\epsilon} \leq 2\sqrt{\epsilon}. \quad (58)$$

First of all, in this case it follows from Part (i) that there exists a γ_4 such that

$$\limsup_{t\to\infty} \|\phi(t)\| \leq \gamma_4 \sqrt{g} + \gamma_4 \tfrac{2-\lambda_1}{1-\lambda_1} \limsup_{t\to\infty} |r(t)|.$$

Hence,

$$\limsup_{t\to\infty} \|\phi(t)\|^2 \leq \underbrace{2\gamma_4^2 g + 2\gamma_4^2 \left(\tfrac{2-\lambda_1}{1-\lambda_1}\right)^2 \limsup_{t\to\infty} |r(t)|^2}_{=:\phi_\infty(g,r)}. \quad (59)$$

Second of all, if we argue as in the first part of the proof of Part (i), we see that for all $i \in \mathbb{N}$:

$$y(t+1) = \phi(t)^\top \overline{\theta}(t_0 + iN_\epsilon) + \underbrace{\phi(t)^\top [\overline{\theta}(t) - \overline{\theta}(t_0 + iN_\epsilon)]}_{=:\Psi_i(t)}, \quad t \in \mathbb{Z}; \quad (60)$$

hence,

$$|\Psi_i(t)| \leq \epsilon N_\epsilon \|\phi(t)\|, \quad t \in [t_0 + (i-1)N_\epsilon, t_0 + (i+1)N_\epsilon]. \quad (61)$$

We would like to use the above to obtain a bound on the weighted error $\overline{\varepsilon}(\cdot)$. To proceed, we apply Proposition 1 to (60). We need to be careful since the disturbance entering the plant (60) is now $\Psi_i(t)$ rather than $w(t+1)$; to this end, recall that the definition of the polynomial $\mathbf{F}(\cdot)$ from Section II is of the form $\mathbf{F}_i(z^{-1}) = f_{i,0} + f_{i,1}z^{-1} + \cdots + f_{i,d-1}z^{-(d-1)}$ corresponding to plant parameters $\overline{\theta}(t_0 + iN_\epsilon)$, and define

$$\overline{\Psi}_i(t) := f_{i,0}\Psi_i(t+d) + f_{i,1}\Psi_i(t+d-1) + \cdots + f_{i,d-1}\Psi_i(t+1);$$

since $\overline{\theta}$ lies in a compact set, the $f_{i,j}$'s also lie in a compact set as well, which means that there exists a constant γ_5, independent of i, so that

$$\overline{\Psi}_i(t)^2 \leq \gamma_5 \sum_{j=1}^{d} \Psi_i(t+j)^2. \quad (62)$$

[3] $N_\epsilon \geq \frac{1}{\sqrt{\epsilon}} - 1 \Rightarrow N_\epsilon\sqrt{\epsilon} \geq 1 - \sqrt{\epsilon} \geq 1 - \sqrt{\widetilde{\epsilon}} \geq 1 - \frac{1}{2} = \frac{1}{2}$.

Arguing as in the first part of the proof of Theorem 2, we see that

$$\sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \frac{\overline{\varepsilon}(j)^2}{g + \|\phi(j-d)\|^2} \leq 8d^2\|\mathcal{S}\|^2 + 4d^2 \sum_{j=t_0+iN_\epsilon-d+1}^{t_0+(i+1)N_\epsilon-1} \frac{\overline{\Psi}_i(j-d-1)^2}{g + \|\phi(j-d)\|^2}$$
$$\leq 8d^2\|\mathcal{S}\|^2 + 4d^2 \sum_{j=t_0+iN_\epsilon-d+1}^{t_0+(i+1)N_\epsilon-1} \frac{1}{g}\overline{\Psi}_i(j-d-1)^2, \quad i \geq 1. \quad (63)$$

At this point, we would like to obtain a bound on the sum on the RHS above. Using the facts that $N_\epsilon \leq \frac{2}{\sqrt{\epsilon}}$ and $2d - 1 \leq N_\epsilon$ and utilizing (59) and (62), there exists a γ_6 so that

$$\limsup_{i\to\infty} \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon-d+1}^{t_0+(i+1)N_\epsilon-1} \frac{1}{g}\overline{\Psi}_i(j-d-1)^2 \leq \tfrac{\gamma_5}{g}\epsilon^2 N_\epsilon^2 \times \limsup_{i\to\infty} \frac{d}{N_\epsilon} \sum_{j=t_0+iN_\epsilon-2d+1}^{t_0+(i+1)N_\epsilon-1} \|\phi(j)\|^2$$
$$\leq \tfrac{4d\gamma_5}{g} \tfrac{N_\epsilon+2d-1}{N_\epsilon}\epsilon \times \phi_\infty(g,r) \leq \tfrac{\gamma_6}{g}\epsilon \times \phi_\infty(g,r).$$

If we now incorporate this into (63) and use (58):

$$\limsup_{i\to\infty} \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \frac{\overline{\varepsilon}(j)^2}{g + \|\phi(j-d)\|^2} \leq 16d^2\|\mathcal{S}\|^2\sqrt{\epsilon} + \tfrac{4d^2\gamma_6}{g}\sqrt{\epsilon} \times \phi_\infty(g,r). \quad (64)$$

But

$$\limsup_{i\to\infty} \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \frac{\overline{\varepsilon}(j)^2}{g + \|\phi(j-d)\|^2} \geq \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \overline{\varepsilon}(j)^2 \frac{1}{g + \phi_\infty(g,r)},$$

so using (64), we end up with

$$\limsup_{i\to\infty} \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \overline{\varepsilon}(j)^2 \leq [16d^2\|\mathcal{S}\|^2\sqrt{\epsilon} + \tfrac{4d^2\gamma_6}{g}\sqrt{\epsilon} \times \phi_\infty(g,r)] \times [g + \phi_\infty(g,r)].$$

Now if we substitute the definition of $\phi_\infty(g,r)$ from (59) into the above and simplify, then we see that there exists γ_7 so that

$$\limsup_{i\to\infty} \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \overline{\varepsilon}(j)^2 \leq \gamma_7\sqrt{\epsilon}(g + \limsup_{t\to\infty} |r(t)|^2 + \tfrac{1}{g} \limsup_{t\to\infty} |r(t)|^4).$$

It is easy to prove that this implies that

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} \overline{\varepsilon}(j)^2 \leq \gamma_7\sqrt{\epsilon}(g + \limsup_{t\to\infty} |r(t)|^2 + \tfrac{1}{g} \limsup_{t\to\infty} |r(t)|^4).$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

From (19), ε and $\bar{\varepsilon}$ are related by a stable transfer function $\frac{1}{\mathbf{L}(z^{-1})} =: \mathbf{P}(z^{-1})$, and using $\|\mathbf{P}\|_\infty$ to denote the maximum of the transfer function $\mathbf{P}(z^{-1})$ on the unit circle, it follows from the above and Parseval's Theorem that

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} \varepsilon(j)^2 \le \gamma_7 \|\mathbf{P}\|_\infty^2 \sqrt{\epsilon}(g + \limsup_{t\to\infty} |r(t)|^2 + \tfrac{1}{g} \limsup_{t\to\infty} |r(t)|^4),$$

which is the desired result.

We now turn to the proof of Part (iv). It is a simplified version of that of the proof of Part (iii), so we simply mention the key difference. Basically, since $\bar{\Psi}_i(t) = \Psi_i(t+1)$ in this case, the corresponding version of (63) is much simpler:

$$\sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \frac{\bar{\varepsilon}(j)^2}{g + \|\phi(j-d)\|^2} \le 8d^2\|\mathcal{S}\|^2 + 4d^2\epsilon^2 N_\epsilon^2.$$

Then simplification carries forward, yielding

$$\limsup_{i\to\infty} \frac{1}{N_\epsilon} \sum_{j=t_0+iN_\epsilon}^{t_0+(i+1)N_\epsilon-1} \bar{\varepsilon}(j)^2 \le 8d^2(\|\mathcal{S}\|^2 + 2)\sqrt{\epsilon} \times [g + \phi_\infty(g, r)].$$

After simplification, we obtain the desired bound. ■

B. Tolerance to Unmodelled Dynamics

We now consider the time-varying plant (49) with the term $d_\Delta(t) \in \mathbb{R}$ added to represent the unmodelled dynamics:

$$y(t+1) = \bar{\theta}(t)^\top \phi(t) + w(t+1) + d_\Delta(t), \quad t \in \mathbb{Z}. \quad (65)$$

We also adopt a common model of unmodelled dynamics used in adaptive control [12]: with $\rho \in (0,1)$ and $\mu > 0$,

$$\mathfrak{w}(t+1) = \rho\mathfrak{w}(t) + \rho\|\phi(t)\|, \ \mathfrak{w}(t_0) = \mathfrak{w}_0 \quad (66a)$$
$$|d_\Delta(t)| \le \mu\mathfrak{w}(t) + \mu\|\phi(t)\|. \quad (66b)$$

It turns out that this model subsumes classical additive uncertainty, multiplicative uncertainty, and uncertainty in a coprime factorization subject to a strict causality constraint; see [20] for a more detailed explanation about this model.

Theorem 4. *For every $\rho \in (0,1)$, $\lambda_2 \in (\max\{\underline{\lambda}, \rho\}, 1)$ and $c_0 \ge 0$, there exist $\bar{\mu} > 0$, $\bar{\epsilon} > 0$ and $c_2 > 0$ so that for every $t_0 \in \mathbb{Z}$, initial condition $x(t_0)$, $\mathfrak{w}_0 \in \mathbb{R}$, $\theta_0 \in \mathcal{S}$, $r, w \in \boldsymbol{\ell}_\infty$, $\epsilon \in [0, \bar{\epsilon}]$, $\bar{\theta} \in \mathfrak{s}(\mathcal{S}_{ab}, c_0, \epsilon)$, $\mu \in (0, \bar{\mu})$, and $g > 0$, when the adaptive controller (12) and (13) is applied to the time-varying plant (65), with d_Δ satisfying (66), then the following holds:*

(i) $$\|\phi(t)\| \le c_2\lambda_2^{t-t_0}(\|x(t_0)\| + |\mathfrak{w}_0|) + c_2\sqrt{g} + \sum_{j=t_0}^{t-1} c_2\lambda_2^{t-j-1}(|w(j+1)| + |r(j)|) + c_2|r(t)|, \quad t \ge t_0;$$

(ii) $$|\varepsilon(t)| \le c_2\lambda_2^{t-t_0}(\|x(t_0)\| + |\mathfrak{w}_0|) + c_2\sqrt{g} + \sum_{j=t_0}^{t} c_2\lambda_2^{t-j}(|w(j)| + |r(j)|), \quad t \ge t_0 + 1.$$

(iii) If $w = 0$ and $c_0 = 0$, then

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} \varepsilon(j)^2 \le c_2(\sqrt{\epsilon} + \mu^2) \times \left(g + \limsup_{t\to\infty} |r(t)|^2 + \tfrac{1}{g} \limsup_{t\to\infty} |r(t)|^4\right).$$

Proof. Fix $\rho \in (0,1)$, $\lambda_1 \in (\underline{\lambda}, 1)$, $\lambda_2 \in (\max\{\lambda_1, \rho\}, 1)$ and $c_0 \ge 0$. By Theorem 3, there exists $\bar{\epsilon} > 0$ and $\bar{\gamma}_1$ so that for every $\epsilon \in [0, \bar{\epsilon}]$, $\bar{\theta} \in \mathfrak{s}(\mathcal{S}_{ab}, c_0, \epsilon)$,

$$\|\phi(t)\| \le \bar{\gamma}_1\lambda_1^{t-t_0}\|\phi(t_0)\| + \bar{\gamma}_1\sqrt{g} + \sum_{j=t_0}^{t-1} \bar{\gamma}_1\lambda_1^{t-j-1}(|r(j)| + |w(j+1)| + |d_\Delta(j)|) + \bar{\gamma}_1|r(t)|, \quad t \ge t_0. \quad (67)$$

First, we convert this inequality to an equality, so consider the associated difference equation

$$\tilde{\phi}(t+1) = \lambda_1\tilde{\phi}(t) + |r(t)| + |w(t+1)| + \mu\tilde{\mathfrak{w}}(t) + \mu\tilde{\phi}(t), \quad \tilde{\phi}(t_0) = \|\phi(t_0)\|,$$

together with the difference equation based on (66):

$$\tilde{\mathfrak{w}}(t+1) = \rho\tilde{\mathfrak{w}}(t) + \rho\tilde{\phi}(t), \quad \tilde{\mathfrak{w}}(t_0) = |\mathfrak{w}_0|.$$

Using induction together with (67) and (66), we can prove that

$$\|\phi(t)\| \le \bar{\gamma}_1\tilde{\phi}(t) + \bar{\gamma}_1|r(t)| + \bar{\gamma}_1\sqrt{g}, \quad (68a)$$
$$|\mathfrak{w}(t)| \le \tilde{\mathfrak{w}}(t), \quad t \ge t_0. \quad (68b)$$

If we combine the difference equations for $\tilde{\phi}(t)$ and $\tilde{\mathfrak{w}}(t)$, we obtain

$$\begin{bmatrix} \tilde{\phi}(t+1) \\ \tilde{\mathfrak{w}}(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda_1 + \mu & \mu \\ \rho & \rho \end{bmatrix}}_{=:A_{\text{cl}}(\mu)} \begin{bmatrix} \tilde{\phi}(t) \\ \tilde{\mathfrak{w}}(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}(|r(t)| + |w(t+1)|), \quad t \ge t_0. \quad (69)$$

Now we see that $A_{\text{cl}}(\mu) \to \begin{bmatrix} \lambda_1 & 0 \\ \rho & \rho \end{bmatrix}$ as $\mu \to 0$, and this matrix has eigenvalues of $\{\lambda_1, \rho\}$ which are both less than $\lambda_2 < 1$. Now we choose $\bar{\mu} > 0$ so that all eigenvalues are less than $(\frac{\lambda_2}{2} + \frac{1}{2}\max\{\lambda_1, \rho\})$ in magnitude for $\mu \in (0, \bar{\mu}]$. By defining $\kappa := \frac{\lambda_2}{2} - \frac{1}{2}\max\{\lambda_1, \rho\}$, we use the proof technique of Desoer in [3] to conclude that for $\mu \in (0, \bar{\mu}]$, we have

$$\|A_{cl}(\mu)^k\| \le \underbrace{\left(\frac{3 + 2\rho + 2\bar{\mu}}{\kappa^2}\right)}_{=:\gamma_2} \lambda_2^k, \quad k \ge 0;$$

if we use this in solving (69) and then apply the bounds in (68), it follows that

$$\left\|\begin{bmatrix} \phi(t) \\ \mathfrak{w}(t) \end{bmatrix}\right\| \le \bar{\gamma}_1\gamma_2\lambda_2^{t-t_0}\left\|\begin{bmatrix} \phi(t_0) \\ \mathfrak{w}_0 \end{bmatrix}\right\| + \bar{\gamma}_1\sqrt{g} + \sum_{j=t_0}^{t-1} \bar{\gamma}_1\gamma_2\lambda_2^{t-j-1}(|r(j)| + |w(j+1)|) + \bar{\gamma}_1|r(t)|, \quad t \ge t_0 \quad (70)$$

as desired. This concludes the proof of Part (i).

To prove Part (ii), we first observe that

$$\varepsilon(t+1) = \phi(t)^\top\bar{\theta}(t) + w(t+1) + d_\Delta(t) - y^*(t+1), \quad t \ge t_0.$$

Using the facts that $\bar{\theta}$ is uniformly bounded, y^* is a filtered version of r, and d_Δ is bounded above by a filtered version of ϕ, together with the bound on ϕ given in Part (i), the bound on $\varepsilon(t)$ follows.

Now we turn to prove the tracking result in Part (iii) when

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

there is no noise entering the system, i.e. $w = 0$, and no large jumps in the plant parameters, i.e. $c_0 = 0$. First of all, from the definition in (66) it follows that

$$\limsup_{t\to\infty} |d_\Delta(t)| \le \frac{\mu}{1-\rho} \limsup_{k\to\infty} \|\phi(k)\|. \tag{71}$$

Second of all, observe that the plant model (65)-(66) is the same as that of the time-varying setup of (49) with the addition of the disturbance d_Δ. If we repeat the proof of Theorem 3(iii) and incorporate the bound on $d_\Delta(t)$ given in (71), then after a straightforward modification we end up with the desired bound. ■

Remark 8. *The main result in Theorems 1 and 2 and the subsequent robustness results of Theorems 3 and 4 should be extendable to the multi-input-multi-output (MIMO) case, but it may not be straightforward. In the analysis of the present case, we mainly rely on constructing the model in (18) based on the original plant model which would bring more complexity in the case of MIMO; this is due to the interactor matrix associated with the MIMO plant (see [9], [5], [7], [34], [4], [2] and [38]). We anticipate that our analytical approach can be extended in a straightforward way to certain classes of MIMO plants, namely ones which have their associated interactor matrix to be diagonal and with equal diagonal elements, i.e. equal delay across all channels; accordingly we could construct a similar model to that of (18) for the MIMO case, then the rest of the analysis would follow in a manner similar to what's there in the rest of our present paper. An extension to the more general case would be more difficult. Due to space limitations, we leave extensions to the MIMO case to future work.*

VI. A Simulation Example

Here we provide a simulation example to illustrate the results of this paper. We consider the following mass-friction system with a time-varying mass:

$$M(t)\ddot{x}(t) + b\dot{x}(t) = F(t),$$

where M is the mass and b is the viscous friction coefficient; the output is the position x and the input is the applied force F. With h as the sampling period, if we define $y(t) := x(th)$ and $u(t) := F(th)$ we can obtain the corresponding discrete-time plant with added noise/disturbance as follows:

$$y(t+1) = -a_1(t)y(t) - a_2(t)y(t-1) + \\ b_0(t)u(t) + b_1(t)u(t-1) + w(t+1);$$

with $K = 1/b$ and $\tau(t) = M(t)/b$, we have $a_1(t) = -\big(1 + e^{-h/\tau(t)}\big)$, $a_2(t) = e^{-h/\tau(t)}$, $b_0(t) = K\big(h - \tau(t) + \tau(t)e^{-h/\tau(t)}\big)$, and $b_1(t) = K\big(\tau(t) - (h+\tau(t))e^{-h/\tau(t)}\big)$. Here we consider $M(t) \in [0.1, 2]$ kg and $b = 10$ N·s/m^2. Note here that the delay d equals one. We want to apply an adaptive controller such that the closed-loop system follows the behavior of a reference model (3) with $n' = 2$; so following the discussion at the beginning of Section II transforming the plant into the predictor form by way of long division, we see that $\alpha_0(t) = l_1 - a_1(t)$, $\alpha_1(t) = l_2 - a_2(t)$, $\beta_0(t) = b_0(t)$, and $\beta_1(t) = b_1(t)$. We choose a reference model represented by

$$\mathbf{L}(z^{-1}) := 1 - z^{-1} + \tfrac{3}{16}z^{-2}, \quad \text{and} \quad \mathbf{H}(z^{-1}) := \tfrac{3}{16},$$

which has poles in the open unit disk as required; then if we choose $h = 0.05$ s, we can set

$$\mathcal{S} := \left\{ \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \beta_0 \\ \beta_1 \end{bmatrix} \in \mathbb{R}^4 : \alpha_0 \in [0,1], \alpha_1 \in [\tfrac{-13}{16}, \tfrac{3}{16}], \\ \beta_0 \in [\tfrac{1}{2}, 5] \times 10^{-3}, \beta_1 \in [0,2] \times 10^{-3} \right\} \supset \mathcal{S}_{\alpha\beta}.$$

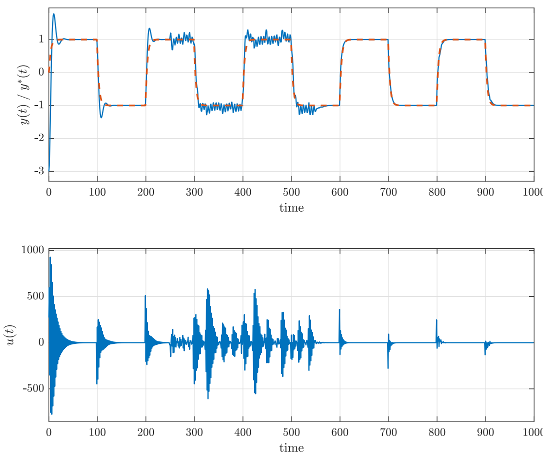

Fig. 1. The first plot shows both the output $\boldsymbol{y(t)}$ (solid) and the reference $\boldsymbol{y^*(t)}$ (dashed); the second plot shows the control input $\boldsymbol{u(t)}$.

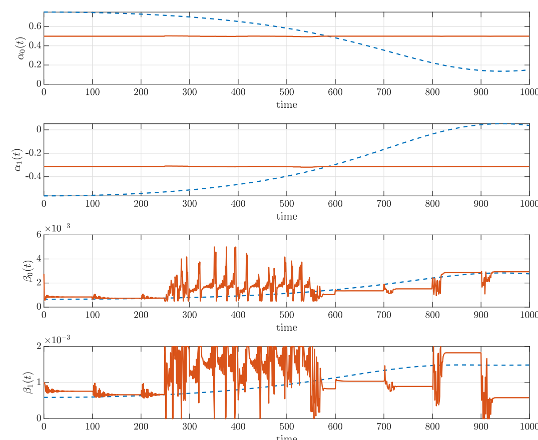

Fig. 2. The plots show the parameter estimates $\hat{\boldsymbol{\theta}}\boldsymbol{(t)}$ (solid) as well as the actual parameters $\boldsymbol{\theta^*}$ (dashed).

We apply the adaptive controller (12) and (13) with $g = \frac{1}{2}$ to this plant with the mass given by

$$M(t) = 1 + \tfrac{3}{4}\cos(\tfrac{h}{15}t),$$

and the disturbance given by:

$$w(t) = \begin{cases} \frac{1}{10}\cos(20t), & 250 < t \le 550 \\ 0, & \text{otherwise.} \end{cases}$$

We set $r(t)$ to be a unit square wave with period of 200 steps. We set $y(-1) = y(0) = -3, u(-1) = 0$, and the initial parameter estimates to the midpoint of the respective intervals. You can see the simulation results in Figures 1 and 2. The controller does an overall good job of tracking. The tracking degrades when the disturbance enters the system but tracking performance improves when the disturbance returns to zero.

VII. Summary and Conclusions

In this paper we examine the model reference adaptive control (MRAC) problem when the commonly used projection algorithm is utilized, subject to several common assumptions on the set of admissible parameters, in particular a compactness constraint as well as knowledge of the sign of the high-frequency gain. While in the literature it is proven that the nonlinear closed-loop system is bounded-input bounded-state, here we prove a quantitative bound consisting of three terms: 1) a decaying exponential on the initial condition, 2)

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

a linear-like convolution bound on the exogenous inputs, and 3) a constant scaled by the square root of the constant in the denominator of the estimator update law. This bound is leveraged to prove tolerance to a degree of unmodelled dynamics and plant parameter time-variation as well. In the noise free case, in the literature it is proven that a bounded reference signal guarantees a square summable tracking error; here we prove a quantitative bound—an explicit upper bound on this sum in terms of the plant initial condition, the size of the reference signals, and the size of the constant in the denominator of the estimator update law. We also provide bounds on the tracking error when there is a non-zero disturbance and/or there are time-variations in the plant parameters and unmodelled dynamics.

We would like to extend these desirable results to the cases when the sign of the high-frequency gain and/or the delay in the system are unknown; an adaptive controller using multiple estimators and switching, along the lines of [29] and [33], may be needed. We would like also to extend our results to the case of MIMO plants.

Appendix

Proof of Proposition 1. Let $t_0 \in \mathbb{Z}$, $\bar{t} > \underline{t} \geq t_0$, $x(t_0) \in \mathbb{R}^{n+m+3d-2}$, $\theta_0 \in \mathcal{S}, \theta \in \mathcal{S}_{ab}, w \in \boldsymbol{\ell}_\infty$, and $g > 0$ be arbitrary.

First we prove part (i). From (12) it is easy to see that

$$\begin{aligned}\|\hat{\theta}(t+1) - \hat{\theta}(t)\| &\leq \|\check{\theta}(t+1) - \hat{\theta}(t)\| \\ &= \left\| \frac{\phi(t-d+1)e(t+1)}{g + \|\phi(t-d+1)\|^2} \right\| \\ &\leq \frac{|e(t+1)|}{g + \|\phi(t-d+1)\|}, \quad t \geq t_0, \end{aligned} \tag{72}$$

as desired for the first property of part (i).

We now prove the second property of part (i). First define $\tilde{\check{\theta}}(t) := \check{\theta}(t) - \theta^*$; from (12a) we obtain

$$\tilde{\check{\theta}}(t+1) = \tilde{\theta}(t) + \frac{\phi(t-d+1)e(t+1)}{g + \|\phi(t-d+1)\|^2}$$

$$\begin{aligned}\Rightarrow \|\tilde{\check{\theta}}(t+1)\|^2 &= \|\tilde{\theta}(t)\|^2 + \frac{\|\phi(t-d+1)\|^2 e(t+1)^2}{[g + \|\phi(t-d+1)\|^2]^2} + \\ &\quad 2\frac{\tilde{\theta}(t)^\top \phi(t-d+1)e(t+1)}{g + \|\phi(t-d+1)\|^2} \\ &\leq \|\tilde{\theta}(t)\|^2 + \frac{e(t+1)^2}{g + \|\phi(t-d+1)\|^2} + \\ &\quad 2\frac{\tilde{\theta}(t)^\top \phi(t-d+1)e(t+1)}{g + \|\phi(t-d+1)\|^2}, \quad t \geq t_0. \end{aligned} \tag{73}$$

From (9) and (10), we obtain

$$e(t+1) = -\tilde{\theta}(t)^\top \phi(t-d+1) + \overline{w}(t-d+1);$$

using this to find a representation for $\tilde{\theta}(t)^\top \phi(t-d+1)$ in (73) we obtain

$$\begin{aligned}\|\tilde{\check{\theta}}(t+1)\|^2 &\leq \|\tilde{\theta}(t)\|^2 + \frac{e(t+1)^2}{g + \|\phi(t-d+1)\|^2} + \\ &\quad 2\frac{\overline{w}(t-d+1)e(t+1) - e(t+1)^2}{g + \|\phi(t-d+1)\|^2} \\ &= \|\tilde{\theta}(t)\|^2 - \frac{e(t+1)^2}{g + \|\phi(t-d+1)\|^2} + \\ &\quad 2\frac{\overline{w}(t-d+1)e(t+1)}{g + \|\phi(t-d+1)\|^2} \\ &\leq \|\tilde{\theta}(t)\|^2 - \frac{1}{2}\frac{e(t+1)^2}{g + \|\phi(t-d+1)\|^2} + \\ &\quad 2\frac{\overline{w}(t-d+1)^2}{g + \|\phi(t-d+1)\|^2}, \quad t \geq t_0. \end{aligned} \tag{74}$$

Since projection does not make the parameter estimate worse, we have $\|\tilde{\theta}(t+1)\| \leq \|\tilde{\check{\theta}}(t+1)\|$; combining this with (74) and iterating, we obtain the second property part of (i).

Next we prove part (ii) for when

$$\|\phi(t-d+1)\|^2 \geq g > 0, \quad t \in [\underline{t}, \bar{t})$$

is satisfied. We see that for $t \in [\underline{t}, \bar{t})$:

$$\frac{\phi(t-d+1)}{g + \|\phi(t-d+1)\|^2} e(t+1) = \underbrace{\frac{\|\phi(t-d+1)\|^2}{g + \|\phi(t-d+1)\|^2}}_{=:z(t)} \frac{\phi(t-d+1)}{\|\phi(t-d+1)\|^2} e(t+1). \tag{75}$$

It is clear that $z(t) \leq 1$; so from (12) and (75), it is easy to see that

$$\begin{aligned}\|\hat{\theta}(t+1) - \hat{\theta}(t)\| &\leq \|\check{\theta}(t+1) - \hat{\theta}(t)\| \\ &= \left\| z(t)\frac{\phi(t-d+1)e(t+1)}{\|\phi(t-d+1)\|^2} \right\| \\ &\leq \frac{|e(t+1)|}{\|\phi(t-d+1)\|}, \quad t \in [\underline{t}, \bar{t}), \end{aligned} \tag{76}$$

which is the first property of part (ii).

We now prove the second property of part (ii). First define $\tilde{\check{\theta}}(t) := \check{\theta}(t) - \theta^*$; from (12a) and (75), we obtain

$$\tilde{\check{\theta}}(t+1) = \tilde{\theta}(t) + z(t)\frac{\phi(t-d+1)e(t+1)}{\|\phi(t-d+1)\|^2}$$

$$\begin{aligned}\Rightarrow \|\tilde{\check{\theta}}(t+1)\|^2 &= \|\tilde{\theta}(t)\|^2 + z(t)^2\frac{e(t+1)^2}{\|\phi(t-d+1)\|^2} + \\ &\quad 2z(t)\frac{\tilde{\theta}(t)^\top \phi(t-d+1)e(t+1)}{\|\phi(t-d+1)\|^2}. \end{aligned} \tag{77}$$

From (9) and (10), we obtain

$$e(t+1) = -\tilde{\theta}(t)^\top \phi(t-d+1) + \overline{w}(t-d+1);$$

using this to find a representation for $\tilde{\theta}(t)^\top \phi(t-d+1)$ in (77) we obtain

$$\begin{aligned}\|\tilde{\check{\theta}}(t+1)\|^2 &\leq \|\tilde{\theta}(t)\|^2 + z(t)^2\frac{e(t+1)^2}{\|\phi(t-d+1)\|^2} + \\ &\quad 2z(t)\frac{\overline{w}(t-d+1)e(t+1) - e(t+1)^2}{\|\phi(t-d+1)\|^2}. \end{aligned} \tag{78}$$

Since $z(t) \in [0,1]$, it follows that $z(t)^2 \leq z(t)$, so from (78) and using similar analysis to that used to derive (74), we have

$$\begin{aligned}\|\tilde{\check{\theta}}(t+1)\|^2 &\leq \|\tilde{\theta}(t)\|^2 + z(t)\Bigg[\frac{e(t+1)^2}{\|\phi(t-d+1)\|^2} + \\ &\quad 2\frac{\overline{w}(t-d+1)e(t+1) - e(t+1)^2}{\|\phi(t-d+1)\|^2}\Bigg] \\ &\leq \|\tilde{\theta}(t)\|^2 + z(t)\Bigg[-\frac{1}{2}\frac{e(t+1)^2}{\|\phi(t-d+1)\|^2} + \\ &\quad 2\frac{\overline{w}(t-d+1)^2}{\|\phi(t-d+1)\|^2}\Bigg]. \end{aligned}$$

Since projection does not make the parameter estimate worse, we have

$$\|\tilde{\theta}(t+1)\|^2 \le \|\tilde{\theta}(t)\|^2 + z(t)\bigg[-\frac{1}{2}\frac{e(t+1)^2}{\|\phi(t-d+1)\|^2} + 2\frac{\overline{w}(t-d+1)^2}{\|\phi(t-d+1)\|^2}\bigg]. \tag{79}$$

By iterating this and rearranging we obtain for $\bar{t} \ge t > \tau \ge \underline{t}$:

$$\begin{aligned}\sum_{t=\tau}^{t-1} \frac{z(j)e(j+1)^2}{2\|\phi(j-d+1)\|^2} &\le \|\tilde{\theta}(\tau)\|^2 - \|\tilde{\theta}(t)\|^2 + \sum_{j=\tau}^{t-1}\frac{2z(j)\overline{w}(j-d+1)^2}{\|\phi(j-d+1)\|^2} \\ &\le \|\tilde{\theta}(\tau)\|^2 - \|\tilde{\theta}(t)\|^2 + \sum_{j=\tau}^{t-1}\frac{2\overline{w}(j-d+1)^2}{\|\phi(j-d+1)\|^2}.\end{aligned} \tag{80}$$

As $\|\phi(t-d+1)\|^2 \ge g$ for $t \in [\underline{t}, \bar{t})$, we have

$$\frac{1}{2\|\phi(t-d+1)\|^2} \le \frac{1}{g+\|\phi(t-d+1)\|^2},$$

so we obtain from the definition of $z(t)$:

$$\frac{z(t)e(t+1)^2}{\|\phi(t-d+1)\|^2} = \frac{e(t+1)^2}{g+\|\phi(t-d+1)\|^2} \ge \frac{e(t+1)^2}{2\|\phi(t-d+1)\|^2};$$

using the above in (80) yields

$$\sum_{t=\tau}^{t-1}\frac{e(j+1)^2}{4\|\phi(j-d+1)\|^2} \le \|\tilde{\theta}(\tau)\|^2 - \|\tilde{\theta}(t)\|^2 + \sum_{j=\tau}^{t-1}\frac{2\overline{w}(j-d+1)^2}{\|\phi(j-d+1)\|^2}, \quad \bar{t} \ge t > \tau \ge \underline{t}; \tag{81}$$

after rearranging, we obtain the second bound of part (ii). ■

Proof of Proposition 3. First, define a square matrix $P := \begin{bmatrix} I_{n+m+d} & \\ & \mathbf{0}\end{bmatrix}$ of size $(n+m+d+n'(d+1))$; then we have

$$\overline{\phi}(t)^\top P = \begin{bmatrix}\phi(t)^\top & \mathbf{0}\end{bmatrix}, \tag{82}$$

so $\overline{\phi}(t)^\top P \overline{\phi}(t) = \phi(t)^\top\phi(t) = \|\phi(t)\|^2$.

As long as $\|\phi(t-d)\|^2 > 0$, define

$$\Delta_1(t) := \frac{\overline{\varepsilon}(t)}{\|\phi(t-d)\|^2}\mathbf{e}_{n+m+d+1}\overline{\phi}(t-d)^\top P; \tag{83}$$

we can represent the term containing $\overline{\varepsilon}(t+d+2)$ in the RHS of (24) as:

$$\mathbf{e}_{n+m+d+1}\overline{\varepsilon}(t+d+2) = \Delta_1(t+d+2)\overline{\phi}(t+2). \tag{84}$$

We now use (27) (which is valid for $t \ge t_0$) to represent $\overline{\phi}(t+2)$ in the RHS of (84) in terms $\overline{\phi}(t)$; if we do this and incorporate the result into (24) then we obtain an equation of the desired form (28) with $\Delta(t)$ and $\overline{\eta}(t)$ defined by

$$\Delta(t) := \Delta_1(t+d+2)A_2(t+1)A_2(t), \tag{85}$$

and

$$\begin{aligned}\overline{\eta}(t) :=\ & \eta(t) + \\ & \Delta_1(t+d+2)\Big[A_2(t+1)B_5(t)\overline{y}^*(t+d+1) + \\ & A_2(t+1)B_6(t)w(t+1) + \\ & \big(A_2(t+1)B_7(t)+B_5(t+1)\big)\overline{y}^*(t+d+2) + \\ & \big(A_2(t+1)B_8(t)+B_6(t+1)\big)w(t+2) + \\ & A_2(t+1)\mathbf{e}_{n+m+d+1}\overline{w}(t+2) + B_7(t+1)\overline{y}^*(t+d+3) + \\ & B_8(t+1)w(t+3) + \mathbf{e}_{n+m+d+1}\overline{w}(t+3)\Big].\end{aligned} \tag{86}$$

Next we prove the desired bound on $\Delta(t)$. From (82) and (83) we see that $\|\Delta_1(t)\| \le \frac{|\overline{\varepsilon}(t)|}{\|\phi(t-d)\|}$. From (17) and part (ii) of Proposition 1, we obtain

$$\begin{aligned}\|\Delta_1(t+d+2)\| &\le \frac{|\overline{\varepsilon}(t+d+2)|}{\|\phi(t+2)\|} \\ &\le \frac{|e(t+d+2)|}{\|\phi(t+2)\|} + \|\hat{\theta}(t+d+1) - \hat{\theta}(t+2)\| \\ &\le \sum_{j=0}^{d-1}\frac{|e(t+d+2-j)|}{\|\phi(t+2-j)\|} = \sum_{j=1}^{d}\|\nu(t+d+2-j)\|, \\ & t \in [\underline{t}, \bar{t}-d-1).\end{aligned} \tag{87}$$

Notice that (87) is only valid on $[\underline{t}, \bar{t}-d-1)$ and not on the whole interval $[\underline{t}, \bar{t})$. Then from (85), using the bound in (87) and Proposition 2, we can easily show that there exist a constant so that (29) holds.

Lastly, using the definition of $\eta(t)$ in (23) along with the bound in (87), it is easy to see that there exists a constant so that we obtain the desired bound (30) on $\overline{\eta}$. ■

Proof of Proposition 5. Fix $\lambda \in (\underline{\lambda}, 1)$. Let $t_0 \in \mathbb{Z}$, $\theta \in \mathcal{S}_{ab}$, $r, w \in \boldsymbol{\ell}_\infty$, $\theta_0 \in \mathcal{S}$, $g > 0$ and $x(t_0) \in \mathbb{R}^{n+m+3d-2}$ be arbitrary; as well, let $[\underline{t}, \bar{t}] \subset [t_0, \infty)$ be an arbitrary interval which satisfies

$$\|\phi(t-d+1)\|^2 \ge g, \quad t \in [\underline{t}, \bar{t}).$$

Now choose $\sigma \in (\underline{\lambda}, \lambda)$.

We will analyze (28) of Proposition 3 to obtain a bound on $\overline{\phi}(t)$. Before proceeding, as pointed out in (31) there exists a constant γ_1 so that for every $\widetilde{A}_g \in \mathcal{A}$, $\|\widetilde{A}_g^k\| \le \gamma_1\sigma^k, k \ge 0$. Also, we need to compute a bound on the sum of several $\|\Delta(\cdot)\|$ terms; since there are d terms on the RHS of (29), by the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\sum_{j=\tau}^{t-1}\|\Delta(j)\| &\le dc_2\sum_{j=\tau+1}^{t+d-1}\|\nu(j+1)\| \\ &\le d^2c_2\left[\sum_{j=\tau+1}^{t+d-1}\|\nu(j+1)\|^2\right]^{\frac{1}{2}}(t-\tau+d-1)^{\frac{1}{2}}, \\ & \bar{t}-d-1 \ge t > \tau \ge \underline{t};\end{aligned} \tag{88}$$

but $(t_2-t_1+d-1)^{\frac{1}{2}} \le d(t_2-t_1)^{\frac{1}{2}}, t_2 > t_1$, so incorporating this and the definition of $\nu(\cdot)$ we have

$$\begin{aligned}\sum_{j=\tau}^{t-1}\|\Delta(j)\| &\le d^3c_2\left[\sum_{j=\tau+1}^{t+d-1}\|\nu(j+1)\|^2\right]^{\frac{1}{2}}(t-\tau)^{\frac{1}{2}} \\ &= d^3c_2\left[\sum_{j=\tau+2}^{t+d}\frac{|e(j+1)|^2}{\|\phi(j-d+1)\|^2}\right]^{\frac{1}{2}}(t-\tau)^{\frac{1}{2}}, \\ & \bar{t}-d-1 \ge t > \tau \ge \underline{t}.\end{aligned} \tag{89}$$

Now we consider the closed-loop system behavior on $[\underline{t}, \bar{t}]$.

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3462555

To proceed, we partition the interval into two parts: one in which the noise $\overline{w}(\cdot)$ is small versus $\phi(\cdot)$ and one where it is not. To this end, with $\mathfrak{v} > 0$ to be chosen shortly, define

$$S_{\text{good}} := \left\{ j \in [\underline{t}, \overline{t}] : \frac{|\overline{w}(j)|^2}{\|\phi(j)\|^2} < \mathfrak{v} \right\},$$

$$S_{\text{bad}} := \left\{ j \in [\underline{t}, \overline{t}] : \frac{|\overline{w}(j)|^2}{\|\phi(j)\|^2} \geq \mathfrak{v} \right\};$$

clearly $[\underline{t}, \overline{t}] = S_{\text{good}} \cup S_{\text{bad}}$.[4] Observe that this partition implicitly depends on $\theta \in \mathcal{S}_{ab}$, as well as the initial conditions. We will easily obtain bounds on the closed-loop system behavior on S_{bad}; we will apply Proposition 4 to analyze the behavior on S_{good}. Before proceeding, we partition the timeline into intervals which oscillate between S_{good} and S_{bad}. To this end, it is easy to see that we can define a (possibly infinite) sequence of intervals of the form $[k_i, k_{i+1})$ satisfying: (i) $k_0 = \underline{t}$; (ii) $[k_i, k_{i+1})$ either belongs to S_{good} or S_{bad}; and (iii) if $k_{i+1} \neq \overline{t}$ and $[k_i, k_{i+1})$ belongs to S_{good} (respectively, S_{bad}), then the interval $[k_{i+1}, k_{i+2})$ must belong to S_{bad} (respectively, S_{good}).

Now we analyze the closed-loop behavior on each interval.

Case 1: The behavior on S_{bad}.

Let $j \in [k_i, k_{i+1}) \subset S_{\text{bad}}$ be arbitrary. In this case, we have $\frac{|\overline{w}(j)|^2}{\|\phi(j)\|^2} \geq \mathfrak{v}$, so,

$$\|\phi(j)\| \leq \frac{1}{\sqrt{\mathfrak{v}}} |\overline{w}(j)|, \quad j \in [k_i, k_{i+1}); \tag{90}$$

then from the crude model (25) and Proposition 2, we have

$$\|\phi(j+1)\| \leq \frac{c_1}{\sqrt{\mathfrak{v}}} |\overline{w}(j)| + c_1 |\overline{y}^*(j+d+1)| + c_1 |w(j+1)|, \quad j \in [k_i, k_{i+1});$$

combining this with (90) yields:

$$\|\phi(j)\| \leq \begin{cases} \frac{1}{\sqrt{\mathfrak{v}}} |\overline{w}(j)|, & j = k_i \\ c_1 \left(\frac{1}{\sqrt{\mathfrak{v}}} + 1 \right) [|\overline{w}(j-1)| + |\overline{y}^*(j+d)| + |w(j)|], & j = k_i + 1, \dots, k_{i+1}. \end{cases} \tag{91}$$

Case 2: The behavior on S_{good}.

Suppose that $[k_i, k_{i+1})$ lies in S_{good}; notice that the bound on $\|\Delta(t)\|$ in (89) occasionally extends outside S_{good}; so we handle the some time steps at the beginning and at the end of the interval $[k_i, k_{i+1})$ separately. To this end, and for purposes which will become apparent in the rest of the proof, define $\overline{d} := \max\{d, n'\} + 1$.

First suppose that $k_{i+1} - k_i \leq 4\overline{d}$; then using the crude model on ϕ in (25) and Proposition 2, it is easy to show that if we define $\gamma_2 := \left(\frac{c_1}{\lambda}\right)^{4\overline{d}}$, then we have

$$\|\phi(t)\| \leq \gamma_2 \lambda^{t-k_i} \|\phi(k_i)\| + \sum_{j=k_i}^{t-1} \gamma_2 \lambda^{t-j-1} (|\overline{y}^*(j+d+1)| + |w(j+1)|), \quad t \in [k_i, k_{i+1}]. \tag{92}$$

Now suppose that $k_{i+1} - k_i > 4\overline{d}$. Define $\overline{k}_i := k_i + \overline{d}$ and $\underline{k}_{i+1} := k_{i+1} - \overline{d}$. By part (ii) of Proposition 1 and using the facts that $\|\tilde{\theta}(t)\| \leq 2\|\mathcal{S}\|$, $\overline{d} \geq d+1$, and that $\frac{|\overline{w}(j)|^2}{\|\phi(j)\|^2} < \mathfrak{v}$ for $j \in [k_i, k_{i+1})$, from (89) we obtain:

$$\sum_{j=\tau}^{t-1} \|\Delta(j)\| \leq d^3 c_2 \left[16\|\mathcal{S}\|^2 + 8\mathfrak{v}(t-\tau+d-1)\right]^{\frac{1}{2}} (t-\tau)^{\frac{1}{2}},$$

$$\underline{k}_{i+1} \geq t > \tau \geq \overline{k}_i.$$

If we restrict $\mathfrak{v} \leq 1$, and define $\gamma_3 := d^3 c_2 \big([16\|\mathcal{S}\|^2 + 8(d-1)]^{\frac{1}{2}} + 2 \big)$, then we obtain

$$\sum_{j=\tau}^{t-1} \|\Delta(j)\| \leq \gamma_3 (t-\tau)^{\frac{1}{2}} + \gamma_3 \mathfrak{v}^{\frac{1}{2}} (t-\tau), \quad \underline{k}_{i+1} \geq t > \tau \geq \overline{k}_i.$$

We now apply Proposition 4: set $g_0 = 0, g_1 = \gamma_3, g_2 = \gamma_3 \mathfrak{v}^{\frac{1}{2}}, \mu = \lambda, \gamma = \gamma_1$; we need $g_2 = \gamma_3 \mathfrak{v}^{\frac{1}{2}} < \frac{\lambda - \sigma}{\gamma_1}$, so if we set $\mathfrak{v} := \min \left\{ 1, \frac{1}{2} \left(\frac{\lambda - \sigma}{\gamma_3 \gamma_1} \right)^2 \right\}$, then from Proposition 4 we see that there exists a constant γ_4 so that the state transition matrix $\mathbf{\Phi}_{\widetilde{A}_g + \Delta}(t, \tau)$ satisfies

$$\|\mathbf{\Phi}_{\widetilde{A}_g + \Delta}(t, \tau)\| \leq \gamma_4 \lambda^{t-\tau}, \quad \underline{k}_{i+1} \geq t > \tau \geq \overline{k}_i. \tag{93}$$

Before solving (28), we obtain a bound on $\bar{\eta}(t)$; from part (ii) of Proposition 1, we see that $\|\nu(t)\| \leq \sqrt{16\|\mathcal{S}\|^2 + 8\mathfrak{v}}, \underline{k}_{i+1} \geq t \geq \overline{k}_i$, so there exists a constant γ_5 so that $\|\bar{\eta}(t)\| \leq \gamma_5 \tilde{w}(t)$, $\underline{k}_{i+1} \geq t \geq \overline{k}_i$. Then, using the bound in (93) to solve (28) we see that there exists a constant γ_6 so that

$$\|\overline{\phi}(t)\| \leq \gamma_6 \lambda^{t-k_i} \|\overline{\phi}(\overline{k}_i)\| + \sum_{j=k_i}^{t-1} \gamma_6 \lambda^{t-j-1} \tilde{w}(j), \quad t \in [\overline{k}_i, \underline{k}_{i+1}]. \tag{94}$$

We want to have a bound on the whole interval $[k_i, k_{i+1})$, and we would like it to be in terms of ϕ instead of $\overline{\phi}$. First, we obtain a bound in terms of ϕ. It is obvious that $\|\phi(t)\| \leq \|\overline{\phi}(t)\|, t \in [\overline{k}_i, \underline{k}_{i+1}]$. Second, from the definitions of $\overline{\phi}(\cdot), \overline{\zeta}(\cdot)$ and $\zeta(\cdot)$, it is easy to see that there exists a constant c_3 such that

$$\|\overline{\phi}(\overline{k}_i)\| \leq c_3 \sum_{j=1}^{2(d+1)} \|\phi(\overline{k}_i + j)\| + c_3 \sum_{j=-n'+2}^{2(d+1)} |y^*(\overline{k}_i + j)|, \quad t \geq t_0;$$

we use the crude model on $\phi(\cdot)$ in (25) and Proposition 2 to obtain bounds on $\|\phi(\overline{k}_i + j)\|, j = 1, 2, \dots, 2(d+1)$, in terms of $\|\phi(k_i)\|$. Incorporating this into (94) and the fact that $\overline{d} \geq n' - 2$, after simplification we see that there exists a constant γ_7 so that

$$\|\phi(t)\| \leq \gamma_7 \lambda^{t-k_i} \|\phi(k_i)\| + \gamma_7 \sum_{j=k_i}^{t-1} \lambda^{t-j-1} \tilde{w}(j), \quad t \in [\overline{k}_i, \underline{k}_{i+1}]. \tag{95}$$

Next, we use the crude model on $\phi(\cdot)$ in (25) and Proposition 2 to find bounds on $\|\phi(t)\|$ for $t \in [k_i, \overline{k}_i)$ and for $t \in (\underline{k}_{i+1}, k_{i+1}]$ and combine them with (95) to conclude that there exists a constant γ_8 so that

$$\|\phi(t)\| \leq \gamma_8 \lambda^{t-k_i} \|\phi(k_i)\| + \sum_{j=k_i}^{t-1} \gamma_8 \lambda^{t-j-1} \tilde{w}(j), \quad t \in [k_i, k_{i+1}], \tag{96}$$

which we combine with (92) to conclude Case 2.

We now glue together the bounds on S_{good} and S_{bad} to obtain a bound which holds on all of $[\underline{t}, \overline{t}]$ using an identical argument used in gluing together similar bounds in Step 3 of the proof of Theorem 1 to end up with the desired bound (33). ■

[4] If the noise is zero, the S_{good} may be the whole interval $[\underline{t}, \overline{t}]$.

References

[1] S. Akhtar and D. S. Bernstein, "Lyapunov-stable discrete-time model reference adaptive control," in *Proc. Amer. Control Conf.*, Jun. 2005, pp. 3174–3179.

[2] X. Chen, "Adaptive sliding mode control for discrete-time multi-input multi-output systems," *Automatica*, vol. 42, no. 3, pp. 427–435, 2006.

[3] C. A. Desoer, "Slowly varying discrete system $x_{i+1} = A_i x_i$," *Electronics Letters*, vol. 6, no. 11, pp. 339–340, May 1970.

[4] L. Dugard and J. M. Dion, "Direct adaptive control for linear multivariable systems," *International Journal of Control*, vol. 42, no. 5, pp. 1251–1281, 1985.

[5] L. Dugard, G. C. Goodwin, and C. de Souza, "Prior knowledge in model reference adaptive control of multiinput multioutput systems," *IEEE Trans. Autom. Control*, vol. 29, no. 8, pp. 761–764, 1984.

[6] 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.

[7] G. C. Goodwin and R. Long, "Generalization of results on multivariable adaptive control," *IEEE Trans. Autom. Control*, vol. 25, no. 6, pp. 1241–1245, 1980.

[8] G. C. Goodwin and D. Q. Mayne, "A parameter estimation perspective of continuous time model reference adaptive control," *Automatica*, vol. 23, no. 1, pp. 57–70, 1987.

[9] 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.

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

[11] 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.

[12] 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.

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

[14] 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.

[15] 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.

[16] 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.

[17] D. E. Miller, "On necessary assumptions in discrete-time model reference adaptive control," *Int. J. Adaptive Control Signal Process.*, vol. 10, no. 6, pp. 589–602, 1996.

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

[19] D. E. Miller, "Classical discrete-time adaptive control revisited: Exponential stabilization," in *Proc. IEEE Conf. Control Technol. Appl.* IEEE, Aug. 2017, pp. 1975–1980.

[20] 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.

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

[22] D. E. Miller and M. T. Shahab, "Adaptive tracking with exponential stability and convolution bounds using vigilant estimation," *Math. Control Signals Syst.*, vol. 32, pp. 241–291, 2020.

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

[24] A. S. Morse, D. Q. Mayne, and G. C. Goodwin, "Applications of hysteresis switching in parameter adaptive control," *IEEE Trans. Autom. Control*, vol. 37, no. 9, pp. 1343–1354, 1992.

[25] 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.

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

[27] K. S. 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.

[28] 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.

[29] M. T. Shahab and D. E. Miller, "Multi-estimator based adaptive control which provides exponential stability: The first-order case," in *Proc. IEEE Conf. Decis. Control*, Dec. 2018, pp. 2223–2228.

[30] M. T. Shahab and D. E. Miller, "Adaptive Set-Point Regulation using Multiple Estimators," in *Proc. IEEE Conf. Decis. Control*, Dec. 2019, pp. 84–89.

[31] M. T. Shahab and D. E. Miller, "The Inherent Robustness of a New Approach to Adaptive Control," in *Proc. IEEE Conf. Control Technol. Appl.*, Aug. 2020, pp. 510–515.

[32] M. T. Shahab and D. E. Miller, "Model Reference Adaptive Control with Linear-like Closed-loop Behavior," in *Proc. IEEE Conf. Decis. Control*, Dec. 2021, pp. 1069–1074.

[33] M. T. Shahab and D. E. Miller, "Asymptotic tracking and linear-like behavior using multi-model adaptive control," *IEEE Trans. Autom. Control*, vol. 67, no. 1, pp. 203–219, 2022.

[34] G. Tao and P. A. Ioannou, "Robust stability and performance improvement of discrete-time multivariable adaptive control systems," *International Journal of Control*, vol. 50, no. 5, pp. 1835–1855, 1989.

[35] K. S. Tsakalis and P. A. Ioannou, "Adaptive control of linear time-varying plants: a new model reference controller structure," *IEEE Trans. Autom. Control*, vol. 34, no. 10, pp. 1038–1046, 1989.

[36] 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.

[37] 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.

[38] C. Yang, L. Zhai, S. S. Ge, T. Chai, and T. H. Lee, "Adaptive model reference control of a class of mimo discrete-time systems with compensation of nonparametric uncertainty," in *Proc. Amer. Control Conf.*, 2008, pp. 4111–4116.

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

Mohamad T. Shahab received the Ph.D. degree in electrical and computer engineering from the University of Waterloo in 2020, and the M.Sc. degree in systems engineering and the B.Sc. degree in control and instrumentation systems engineering, both from KFUPM, in 2015 and 2007, respectively. After earning his Ph.D. degree he joined, as a postdoctoral fellow, King Abdullah University of Science & Technology (KAUST). He was also a research engineer with Baker Hughes during 2012–2015. Current research interests include adaptive control, optimization & control, and applications to robotics & autonomous systems.

Daniel E. Miller 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 Professor *Emeritus*. 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.