



Adaptive tracking with exponential stability and convolution bounds using vigilant estimation

Daniel E. Miller¹  · Mohamad T. Shahab¹

Received: 5 October 2019 / Accepted: 24 March 2020 / Published online: 6 April 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

Classical discrete-time adaptive controllers provide asymptotic stabilization and tracking; neither exponential stabilization nor a bounded noise gain is typically proven. In our recent work, it is shown, in both the pole placement stability setting and the first-order one-step-ahead tracking setting, that if the original, ideal, projection algorithm is used (subject to the common assumption that the plant parameters lie in a convex, compact set and that the parameter estimates are restricted to that set) as part of the adaptive controller, then a linear-like convolution bound on the closed-loop behaviour can be proven; this immediately confers exponential stability and a bounded noise gain, and it can be leveraged to provide tolerance to unmodelled dynamics and plant parameter variation. In this paper, we solve the much harder problem of adaptive tracking; under classical assumptions on the set of unmodelled parameters, including the requirement that the plant be minimum phase, we are able to prove not only the linear-like properties discussed above, but also very desirable bounds on the tracking performance. We achieve this by using a modified version of the ideal projection algorithm, termed as **vigilant estimator**: it is equally alert when the plant state is large or small and is turned off when it is clear that the disturbance is overwhelming the estimation process.

Keywords Adaptive control · Projection algorithm · Exponential stability · Bounded gain · Convolution bound

This research was supported by a grant from the Natural Sciences Research Council of Canada.

✉ Daniel E. Miller
miller@uwaterloo.ca

¹ Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

1 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. While adaptive control has been studied as far back as the 1950s, the first general proofs that parameter adaptive controllers work came around 1980, e.g. see [3,5,27,32,33]. However, the original controllers are typically not robust to unmodelled dynamics, do not tolerate time-variations well, have poor transient behaviour and do not handle noise (or disturbances) well, e.g. see [34]. During the following 2 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 [9,11,14,15,37]; arguably the simplest is that of using projection onto a convex set of admissible parameters, e.g. see [13,31,41–44]. However, in general these redesigned controllers provide asymptotic stability and not exponential stability, with no bounded gain on the noise; that being said, some of them, especially those which use projection, provide a bounded-noise bounded-state property, as well as tolerance of some degree of unmodelled dynamics and/or time-variations.

The goal of this paper is to design adaptive controllers for which the closed-loop system exhibits highly desirable LTI-like system properties, such as exponential stability, a bounded gain on the noise and ideally a convolution bound on the input-output behaviour;¹ in addition, we wish to obtain “good tracking” of a reference signal. As far as the authors are aware, in the classical approach to adaptive control a bounded gain on the noise is proven only in [44]; however, neither a “crisp” exponential bound on the effect of the initial condition nor a convolution bound on the closed-loop behaviour is proven. While it is possible to prove a form of exponential stability if the reference input is sufficiently persistently exciting, e.g. see [2], this places a stringent requirement on an exogenous input, which we would like to avoid.

There are several non-classical approaches to adaptive control which provide some of the LTI-like system properties. First of all, in [4,23] a logic-based switching approach was used to sequence through a predefined list of candidate controllers; while exponential stability is proven, the transient behaviour can be quite poor and a bounded gain on the noise is not proven. In a related approach, a high-gain controller is used to provide excellent transient and steady-state tracking for minimum phase systems [22]; in this case as well, a bounded gain on the noise is not proven. A more sophisticated logic-based approach, labelled supervisory control, was proposed by Morse—see [7,8,28,29,39]; here a supervisor switches in an efficient way between candidate controllers, and in certain circumstances a bounded gain on the noise can be proven—see [30,40], and the Concluding Remarks section of [29]. A related approach, called localization-based switching adaptive control, uses a falsification approach to prove exponential stability as well as a degree of tolerance of disturbances, e.g. see

¹ It is proven in [36] that for a large class of plant/controller combinations, a convolution bound on the input-output behaviour can be leveraged to prove tolerance to slow variations in the plant parameters as well as to a small amount of unmodelled dynamics.

[45], though a bounded gain on the noise is not proven. In none of the above cases is a convolution bound on the closed-loop behaviour proven.

Another non-classical approach to adaptive control, proposed by the first author, is based on periodic probing, estimation and control: rather than estimate the plant or controller parameters, the goal is to estimate what the control signal would be if the plant parameters and plant state were known and the ‘ideal controller’ were applied. Under suitable assumptions on the plant uncertainty, exponential stability and a bounded gain on the noise are achieved, and a degree of unmodelled dynamics and slow time-variations are allowed; for non-minimum phase systems, near optimal transient performance is also provided—see [17,18,38], while for minimum-phase systems, near exact tracking is provided, even in the presence of rapid time-variations—see [16,24]. While a convolution bound is not proven, the biggest drawback is that while a bounded gain on the noise is always achieved, it tends to increase dramatically the closer that one gets to optimality. Furthermore, because of the nature of the approach, it only works in the continuous-time domain.

This brings us to the proposed approach of the paper, wherein we show how to achieve our objectives in the discrete-time setting under some classical assumptions on the set of plant parameters. We adopt a common approach to classical adaptive control—the use of a projection algorithm-based estimator together with a tuneable compensator whose parameters are chosen via the certainty equivalence principle. However, while in the literature it is very common to use a modified version of the ideal projection algorithm in order to avoid division by zero,² here, we adopt another approach to alleviate this numerical concern. We label the resulting estimator a “vigilant estimator”: in the absence of noise it is equally alert to small signals as large signals (unlike the modified version), and in the noisy case we turn off the estimator update if it is clear that the noise is overwhelming the data. Indeed, in earlier work by the authors on the first-order setting [19] and in the pole placement setting of [20,25], versions of this estimator are used as a building block of adaptive controllers which provide exponential stability, a bounded gain on the noise, and linear-like convolution bounds on the closed-loop behaviour; as far as the authors are aware, **such LTI-like bounds have never before been proven in the adaptive setting**. The objective of the present paper is to use the general approach of [19,20,25] to analyse the much harder adaptive tracking control problem for high-order systems.

We initially expected the analysis to follow in a straight-forward manner from either the first-order tracking approach of [19] or the pole placement approach of [20,25]; this has not proven to be the situation. It turns out that the importance of the system delay in this setting creates significant additional complexity not found in either case, and the tracking objective is not present in the latter case. That being said, we have adopted ideas from the pole placement setting of [20,25] as a starting point, but we have carried out several innovative modifications in order to allow us to prove not

² An exception is the work of Ydstie [43,44], who considers the ideal projection algorithm as a special case; however, a crisp bound on the effect of the initial condition and a convolution bound on the effect of the exogenous inputs are not proven. Another notable exception is the work of Akhtar and Bernstein [1], where they are able to prove Lyapunov stability; however, they do not prove a convolution bound on the effect of the exogenous inputs either, and they assume that the high-frequency gain is known.

only the same highly desirable linear-like properties, but also several highly desirable tracking results:

- (i) if there is no noise, we prove an explicit 2-norm bound on the size of the tracking error in terms of the size of the initial condition and the reference signal (in the literature on classical adaptive control it is typically proven only that the tracking error is square summable);
- (ii) if there is no noise but there are slow time-variations, then we prove that we can bound the size of the average tracking error by the size of the time-variation (in the literature this situation is rarely considered);
- (iii) if there is noise, then under some technical conditions, we prove a bound on the average tracking error in terms of the average disturbance times another complicated quantity (in the literature it is usually proven only that the tracking error is bounded).

The proofs of the results contained herein are significantly more involved than the proofs of the classical results in the literature, such as the seminal work of Goodwin et al. [5]. The main reason is that here we are interested in obtaining crisp quantitative bounds on the closed-loop response, rather than the typical qualitative bounds provided in most of the literature.³ Last of all, we would like to point out that an early version of this work appears in [26].

At this point, we provide an outline of the paper. In Sect. 2, we introduce the plant, rewrite it into a form more amenable to analysis and introduce the parameter estimator as well as the adaptive control law. In Sect. 3, we introduce three high-level models for use in system analysis; we provide several technical results which will be useful in the closed-loop analysis. In Sect. 4, we use the three models of the previous section to prove that highly desirable convolution bounds on the system state hold. In Sect. 5, we show that convolution bounds still hold even in the presence of a degree of unmodelled dynamics and plant parameter variations. In Sect. 6, we derive bounds on the tracking error in a variety of situations: the noise-free case, the noise-free case with time-variation and the time-invariant noisy case. In Sect. 7, we will provide several illustrative simulation examples. Last of all, we wrap up with a Summary and Conclusions in Sect. 8.

Before proceeding, we present some mathematical preliminaries. Let \mathbf{Z} denote the set of integers, \mathbf{Z}^+ the set of non-negative integers, \mathbf{N} the set of natural numbers, \mathbf{R} the set of real numbers, and \mathbf{R}^+ the set of non-negative real numbers. We use the Euclidean 2-norm for vectors and the corresponding induced norm for matrices and denote the norm of a vector or matrix by $\|\cdot\|$. We let $s(\mathbf{R}^{n \times m})$ denote the set of all $\mathbf{R}^{n \times m}$ -valued sequences, and we let $l_\infty(\mathbf{R}^{n \times m})$ denote the subset of bounded sequences; we define the norm of $u \in l_\infty(\mathbf{R}^{n \times m})$ by $\|u\|_\infty := \sup_{k \in \mathbf{Z}} \|u(k)\|$.

If $\mathcal{S} \subset \mathbf{R}^p$ is a convex and compact set, we define $\|\mathcal{S}\| := \max_{x \in \mathcal{S}} \|x\|$ and the function $\pi_{\mathcal{S}} : \mathbf{R}^p \rightarrow \mathcal{S}$ denotes the projection onto \mathcal{S} ; it is well-known that $\pi_{\mathcal{S}}$ is well-defined.

³ An exception is the work of [12]: they provide crisp bounds on the size of the tracking error; however, they do not prove the nice linear-like properties which we desire.

2 The setup

In this paper, we start with a linear time-invariant discrete-time plant described by

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

with

- $y(t) \in \mathbf{R}$ the measured output,
- $u(t) \in \mathbf{R}$ the control input,
- $w(t) \in \mathbf{R}$ the disturbance (or noise) input;
- the parameters regularized so that $a_0 = 1$, and
- the system delay $d \geq 1$ being known (so $b_0 \neq 0$).

Associated with this plant model are the polynomials $A(z^{-1}) := \sum_{i=0}^n a_i z^{-i}$ and $B(z^{-1}) := \sum_{i=0}^m b_i z^{-i}$, the transfer function $z^{-d} \frac{B(z^{-1})}{A(z^{-1})}$, and the list of plant parameters:

$$\theta_{ab}^* := [a_1 \cdots a_n \ b_0 \cdots b_m]^T.$$

Remark 1 It is straight-forward to verify that if the system has a disturbance at both the input and output, then it can be converted to a system of the above form.

The goal is closed-loop stability and asymptotic tracking of an **exogenous reference input** $y^*(t)$. We impose several assumptions on the set of admissible parameters. The first set is standard in the literature, e.g. [5,6]:

Assumption 1

- (i) n is known;
- (ii) m is known;
- (iii) the system delay d is known;
- (iv) $\text{sgn}(b_0)$ is known;
- (v) the polynomial $B(z^{-1})$ has all of its zeros in the open unit disk.

Remark 2 Since we do not require $a_n \neq 0$, Assumption 1 (i) can be interpreted as assuming that an upper bound on n is known; similarly, Assumption 1 (ii) can be interpreted as assuming that an upper bound on m is known. The constraint on the zeros of $B(z^{-1})$ in Assumption 1 (v) is a requirement that the plant be minimum phase; this is necessary to ensure tracking of an arbitrary bounded reference signal [21].

The second assumption is less standard:

Assumption 2 The set of admissible parameters, which we label $\mathcal{S}_{ab} \subset \mathbf{R}^{n+m+1}$, is compact.

Remark 3 The boundedness requirement on \mathcal{S}_{ab} is quite reasonable in practical situations; it is used here to prove uniform bounds and decay rates on the closed-loop behaviour.

There are many ways to pose and solve an adaptive tracking problem, with the d -step-ahead approach and the more general model reference control approach being the standard ones—see [5,6]. To minimize complexity, we choose the first one, since it is the simpler of the two. To proceed we use a parameter estimator together with an adaptive d -step-ahead control law. To design the estimator, it is convenient to put the plant into the so-called predictor form. To this end, following [6], we carry out long division by dividing $A(z^{-1})$ into one, and define $F(z^{-1}) = \sum_{i=0}^{d-1} f_i z^{-i}$ and $G(z^{-1}) = \sum_{i=0}^{n-1} g_i z^{-i}$ satisfying

$$\frac{1}{A(z^{-1})} = F(z^{-1}) + z^{-d} \frac{G(z^{-1})}{A(z^{-1})}.$$

Hence, if we define

$$\begin{aligned} \beta(z^{-1}) &:= \sum_{i=0}^{m+d-1} \beta_i z^{-i} := F(z^{-1})B(z^{-1}), \\ \alpha(z^{-1}) &:= \sum_{i=0}^{n-1} \alpha_i z^{-i} := G(z^{-1}), \\ \bar{w}(t) &:= f_0 w(t+d) + \dots + f_{d-1} w(t+1), \end{aligned}$$

then we can rewrite the plant model as

$$\begin{aligned} y(t+d) &= \sum_{i=0}^{n-1} \alpha_i y(t-i) + \sum_{i=0}^{m+d-1} \beta_i u(t-i) + \bar{w}(t) \\ &= \underbrace{\begin{bmatrix} y(t) \\ \vdots \\ y(t-n+1) \\ u(t) \\ \vdots \\ u(t-m-d+1) \end{bmatrix}}_{=: \phi(t)^T} \underbrace{\begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{n-1} \\ \beta_0 \\ \vdots \\ \beta_{m+d-1} \end{bmatrix}}_{=: \theta^*} + \bar{w}(t), \quad t \in \mathbf{Z}. \end{aligned} \tag{2}$$

Let $\mathcal{S}_{\alpha\beta}$ denote the set of admissible θ^* which arise from the original plant parameters which lie in \mathcal{S}_{ab} ; since the associated mapping is analytic, it is clear that the compactness of \mathcal{S}_{ab} means that $\mathcal{S}_{\alpha\beta}$ is compact as well. Furthermore, it is easy to see that $f_0 = 1$, so $\beta_0 = b_0$, which means that the sign of β_0 is always the same. It is convenient that the set of admissible parameters in the new parameter space be

convex and closed; so at this point let $\mathcal{S} \subset \mathbf{R}^{n+m+d}$ be any **compact and convex set** containing $\mathcal{S}_{\alpha\beta}$ for which the $n + 1^{th}$ element (the one which corresponds to β_0) is never zero, e.g. the convex hull of $\mathcal{S}_{\alpha\beta}$ would do.

The d -step-ahead control law is the one given by

$$y^*(t + d) = \sum_{i=0}^{n-1} \alpha_i y(t - i) + \sum_{i=0}^{m+d-1} \beta_i u(t - i),$$

or equivalently

$$u(t) = \frac{1}{\beta_0} \left[y^*(t + d) - \sum_{i=0}^{n-1} \alpha_i y(t - i) - \sum_{i=1}^{m+d-1} \beta_i u(t - i) \right];$$

in the absence of a disturbance, and assuming that this controller is applied for all $t \in \mathbf{Z}$, we have $y(t) = y^*(t)$ for all $t \in \mathbf{Z}$. Of course, if the plant parameters are unknown, we need to use estimates; also, the adaptive version of the d -step-ahead control law is only applied after some initial time, i.e. for $t \geq t_0$.

2.1 Initialization

In most adaptive control results, the goal is to prove asymptotic behaviour, so the details of the initial condition are unimportant. Here, however, we wish to get a bound on the transient behaviour so we must proceed carefully. In the pole placement setting of [20], this was relatively straight-forward: the delay plays no role, the controller is strictly causal, and we start the plant estimator off at time t_0 , with an initial “plant state” of $\phi(t_0) = \phi_0$. Here we have a more complicated situation, even in the case of $d = 1$, since the proposed controller is not strictly causal.

To proceed, observe that if we wish to solve (2) for $y(t)$ starting at time t_0 , then it is clear that we need an initial condition of

$$x_0 := \left[y(t_0 - 1) \quad \cdots \quad y(t_0 - n - d + 1) \quad u(t_0 - 1) \quad \cdots \quad u(t_0 - m - 2d + 1) \right]^T \in \mathbf{R}^{n+m+3d-2}. \tag{3}$$

Observe that this is sufficient information to obtain $\{\phi(t_0 - 1), \dots, \phi(t_0 - d)\}$.

2.2 Parameter estimation

We can rewrite the plant (2) as

$$y(t) = \phi(t - d)^T \theta^* + \bar{w}(t - d), \quad t \geq t_0, \tag{4}$$

with an initial condition of x_0 . Given an estimate $\hat{\theta}(t)$ of θ^* at time t , we define the **prediction error** by

$$e(t + 1) := y(t + 1) - \phi(t - d + 1)^T \hat{\theta}(t); \tag{5}$$

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

$$\operatorname{argmin}_{\theta} \{ \|\theta - \hat{\theta}(t)\| : y(t + 1) = \phi(t - d + 1)^T \theta \},$$

yielding the **ideal (projection) algorithm**

$$\hat{\theta}(t + 1) = \begin{cases} \hat{\theta}(t) & \text{if } \phi(t - d + 1) = 0 \\ \hat{\theta}(t) + \frac{\phi(t-d+1)}{\|\phi(t-d+1)\|^2} e(t + 1) & \text{otherwise;} \end{cases} \tag{6}$$

at this point, we can also constrain the estimate to \mathcal{S} by projection. Of course, if $\phi(t - d + 1)$ is close to zero, numerical problems can occur, so it is very common in the literature (e.g. [5,6]) to add a constant to the denominator and possibly another gain in the numerator: with $0 < \bar{\alpha} < 2$ and $\bar{\beta} > 0$, consider the **classical algorithm**⁴

$$\hat{\theta}(t + 1) = \hat{\theta}(t) + \frac{\bar{\alpha}\phi(t)}{\bar{\beta} + \phi(t)^T \phi(t)} e(t + 1). \tag{7}$$

However, as pointed out in [19,20,25], this can lead to the loss of exponential stability and a loss of a bounded gain on the noise.

We propose a middle ground: as proposed in [20,25], we turn off the estimation if it is clear that the disturbance signal $\bar{w}(t)$ is swamping the estimation error.⁵ More specifically, if we examine the update law in (6) when $\phi(t - d + 1) \neq 0$, we see that

$$\|\hat{\theta}(t + 1) - \hat{\theta}(t)\| = \frac{|e(t + 1)|}{\|\phi(t - d + 1)\|}.$$

Suppose that $\theta_0 \in \mathcal{S}$ and $\hat{\theta}(t) \in \mathcal{S}$; if this update quantity is large but less than $2\|\mathcal{S}\|$, then it could very well be that $\hat{\theta}(t + 1) \in \mathcal{S}$ and it could be that the large update is due to $\hat{\theta}(t)$ being very inaccurate; on the other hand, if this quantity is larger than $2\|\mathcal{S}\|$, then it is clear that $\hat{\theta}(t + 1) \notin \mathcal{S}$ and probably quite inaccurate—in this case the disturbance may be fairly large relative to the other signals. To this end, with $\delta \in (0, \infty]$, we turn off the estimator if the update is larger than $2\|\mathcal{S}\| + \delta$ in magnitude; so define $\rho_{\delta} : \mathbf{R}^{n+m+d} \times \mathbf{R} \rightarrow \{0, 1\}$ by

$$\rho_{\delta}(\phi(t - d + 1), e(t + 1)) := \begin{cases} 1 & \text{if } |e(t + 1)| < (2\|\mathcal{S}\| + \delta)\|\phi(t - d + 1)\| \\ 0 & \text{otherwise;} \end{cases}$$

⁴ We can make this even more general by letting $\bar{\alpha}$ be time-varying.

⁵ It is common in the literature to turn off the estimator if the prediction error is small, e.g. see [6]; this is a completely different idea.

given $\hat{\theta}(t_0 - 1) = \theta_0$, for $t \geq t_0 - 1$ we define⁶

$$\check{\theta}(t + 1) = \hat{\theta}(t) + \rho_\delta(\phi(t - d + 1), e(t + 1)) \times \frac{\phi(t - d + 1)}{\|\phi(t - d + 1)\|^2} e(t + 1), \tag{8}$$

which we then project onto \mathcal{S} :

$$\hat{\theta}(t + 1) := \pi_{\mathcal{S}}(\check{\theta}(t + 1)). \tag{9}$$

Remark 4 We label this approach **vigilant estimation** for two reasons.

- (i) First of all, suppose that the disturbance is zero, i.e. $w(t) = 0$, and $\phi(t - d + 1) \neq 0$. Then a careful examination of the update quantity reveals that $\rho_\delta(\cdot, \cdot) = 1$ and

$$\rho_\delta(\cdot, \cdot) \frac{\phi(t - d + 1)}{\|\phi(t - d + 1)\|^2} e(t + 1) = - \frac{\phi(t - d + 1)\phi(t - d + 1)^T}{\|\phi(t - d + 1)\|^2} [\hat{\theta}(t) - \theta^*];$$

so we see that the gain on the parameter estimate error $\hat{\theta}(t) - \theta^*$ is exactly

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

which is **scale invariant**—if we replace $\phi(t - d + 1)$ by $c\phi(t - d + 1)$ for any non-zero c , then the quantity is the same. Hence, the estimator is as alert when ϕ is small as when it is large; this differs from the classical algorithm, wherein the gain gets small when ϕ is small.

- (ii) Second of all, as discussed above, the update algorithm turns off if it is clear that the disturbance is over-whelming the estimation process.

2.3 Properties of the estimation algorithm

Analysing the closed-loop system will require a careful analysis of the estimation algorithm. We define the parameter estimation error by $\tilde{\theta}(t) := \hat{\theta}(t) - \theta^*$ and the corresponding Lyapunov function associated with $\tilde{\theta}(t)$, namely $V(t) := \tilde{\theta}(t)^T \tilde{\theta}(t)$. In the following result, we list a property of $V(t)$; it is a straight-forward generalization of what holds in the pole placement setup of [20,25].

Proposition 1 *For every $t_0 \in \mathbf{Z}$, $x_0 \in \mathbf{R}^{n+m+3d-2}$, $\theta_0 \in \mathcal{S}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, y^* , $w \in l_\infty$, and $\delta \in (0, \infty]$, when the estimator (8) and (9) is applied to the plant (1), the following holds:*

⁶ If $\delta = \infty$, then we adopt the understanding that $\infty \times 0 = 0$, in which case this formula collapses into the original version (6).

$$\begin{aligned} \|\hat{\theta}(t+1) - \hat{\theta}(t)\| &\leq \rho_\delta(\phi(t-d+1), e(t+1)) \times \frac{|e(t+1)|}{\|\phi(t-d+1)\|}, \quad t \geq t_0 - 1, \\ V(t) &\leq V(t) + \sum_{j=t}^{t-1} \rho_\delta(\phi(j-d+1), e(j+1)) \\ &\quad \times \left[-\frac{1}{2} \frac{[e(j+1)]^2}{\|\phi(j-d+1)\|^2} + 2 \frac{[\bar{w}(j-d+1)]^2}{\|\phi(j-d+1)\|^2} \right], \quad t \geq t \geq t_0 - 1. \end{aligned} \tag{10}$$

2.4 The control law

The elements of $\hat{\theta}(t)$ are partitioned in a natural way as

$$\left[\hat{\alpha}_0(t) \cdots \hat{\alpha}_{n-1}(t) \hat{\beta}_0(t) \cdots \hat{\beta}_{m+d-1}(t) \right]^T.$$

The *d*-step-ahead adaptive control law is that of

$$y^*(t+d) = \hat{\theta}(t)^T \phi(t), \quad t \geq t_0,$$

or equivalently

$$\sum_{i=0}^{m+d-1} \hat{\beta}_i(t) u(t-i) = y^*(t+d) - \sum_{i=0}^{n-1} \hat{\alpha}_i(t) y(t-i), \quad t \geq t_0. \tag{11}$$

Hence, as is common in this setup, we assume that the controller has access to the reference signal $y^*(t)$ exactly *d* time units in advance.

Remark 5 With this choice of control law, it is easy to prove that the prediction error $e(t)$ and the tracking error

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

are different if $d \neq 1$. Indeed, it is easy to see that

$$\varepsilon(t) = -\phi(t-d)^T \tilde{\theta}(t-d) + \bar{w}(t-d), \quad t \geq t_0 + d, \tag{12}$$

$$e(t) = -\phi(t-d)^T \tilde{\theta}(t-1) + \bar{w}(t-d), \quad t \geq t_0. \tag{13}$$

Notice, in particular, that (12) provides a nice closed-form expression for the tracking error $\varepsilon(t)$ only for $t \geq t_0 + d$; the reason is that the tracking error for $t = t_0, \dots, t_0 + d - 1$ is determined by x_0, w and y^* .

The goal of this paper is to prove that the adaptive controller consisting of the estimator (8), (9) together with the control equation (11) yields highly desirable linear-like convolution bounds on the closed-loop behaviour as well as provides good tracking

of y^* . In the next section, we develop several models used in the development, after which we state and prove the main result.

Remark 6 While the proposed adaptive controller which consists of the estimator (8), (9) together with the controller (11) is nonlinear, when it is applied to the plant it turns out that the closed-loop system enjoys the homogeneity property. While it does not enjoy the additivity property needed for linearity, we will soon see that we are still able to prove linear-like convolution bounds on the closed-loop behaviour.

3 Preliminary analysis

In the pole-placement adaptive control setup of our earlier work [20,25], a key closed-loop model consists of an update equation for $\phi(t)$, with the state matrix consisting of controller and plant estimates; this was effective—the characteristic polynomial of this matrix is time-invariant and has all roots in the open unit disk. If we were to mimic this in the d -step-ahead setup, the characteristic polynomial would have roots which are time-varying, with some at zero and the rest at the zeros of the naturally defined polynomial $\hat{\beta}(t, z^{-1})$, which is time-varying and may not have roots in the open unit disk. Hence, at this point we make an important deviation from the approach of [20,25]: we construct three different models for use in the analysis.

3.1 A good closed-loop model

In the first model, we obtain an update equation for $\phi(t)$ which avoids the use of plant parameter estimates, but which is driven by the tracking error. Only two elements of ϕ have a complicated description:

$$\phi_1(t + 1) = y(t + 1) = \varepsilon(t + 1) + y^*(t + 1), \quad t \geq t_0 - 1,$$

and the $u(t + 1)$ term, for which we use the original plant model to write

$$\begin{aligned} \phi_{n+1}(t + 1) &= u(t + 1) \\ &= \frac{1}{b_0} \left[\sum_{i=0}^d a_i (\varepsilon(t + d + 1 - i) + y^*(t + d + 1 - i)) \right. \\ &\quad \left. + \sum_{i=d+1}^n a_i y(t + d + 1 - i) - \sum_{i=0}^{m-1} b_{i+1} u(t - i) - w(t + d + 1) \right]. \end{aligned}$$

With $e_i \in \mathbf{R}^{n+m+d}$ the i^{th} normal vector, if we now define

$$B_1 := e_1, \quad B_2 := e_{n+1}, \tag{14}$$

then it is easy to see that there exists a matrix $A_g \in \mathbf{R}^{(n+m+d) \times (n+m+d)}$ (which depends implicitly on $\theta_{ab}^* \in \mathcal{S}_{ab}$) so that the following equation holds:

$$\begin{aligned} \phi(t + 1) = & A_g \phi(t) + B_1 \varepsilon(t + 1) + B_2 \sum_{j=0}^d \left[\frac{a_{d-j}}{b_0} \varepsilon(t + 1 + j) + \frac{a_{d-j}}{b_0} y^*(t + 1 + j) \right] \\ & + B_1 y^*(t + 1) - \frac{1}{b_0} B_2 w(t + d + 1), \quad t \geq t_0 - 1. \end{aligned} \tag{15}$$

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

3.2 A crude closed-loop model

At times, we will need to use a crude model to bound the size of the growth of $\phi(t)$ in terms of the exogenous inputs. Once again, only two elements of $\phi(t)$ have a complicated description: to describe $y(t + 1)$ we use the plant model (1):

$$\begin{aligned} \phi_1(t + 1) &= y(t + 1) \\ &= - \sum_{i=1}^n a_i y(t + 1 - i) + \sum_{i=0}^m b_i u(t + 1 - d - i) + w(t + 1) \\ &=: \bar{\theta}_{ab}^* \phi(t) + w(t + 1), \end{aligned}$$

and to describe $u(t + 1)$ we use the control law:

$$\begin{aligned} y^*(t + d) &= \hat{\theta}(t)^T \phi(t) \\ \Rightarrow y^*(t + d + 1) &= \hat{\theta}(t + 1)^T \phi(t + 1), \quad t \geq t_0 - 1; \end{aligned}$$

it is easy to define $\bar{\theta}_{\alpha\beta}(t)$ in terms of the elements of $\hat{\theta}(t + 1)$ so that

$$y^*(t + d + 1) = \bar{\theta}_{\alpha\beta}(t)^T \phi(t) + \hat{\alpha}_0(t + 1)y(t + 1) + \hat{\beta}_0(t + 1)u(t + 1), \quad t \geq t_0 - 1.$$

If we combine this with the formula for $y(t + 1)$ above, we end up with

$$\begin{aligned} u(t + 1) = & \frac{1}{\hat{\beta}_0(t + 1)} [-\bar{\theta}_{\alpha\beta}(t) - \hat{\alpha}_0(t + 1)\bar{\theta}_{ab}^*] \phi(t) \\ & + \frac{1}{\hat{\beta}_0(t + 1)} y^*(t + d + 1) - \frac{\hat{\alpha}_0(t + 1)}{\hat{\beta}_0(t + 1)} w(t + 1), \quad t \geq t_0 - 1. \end{aligned}$$

Hence, we can define matrices $A_b(t)$, $B_3(t)$ and $B_4(t)$ so that

$$\phi(t + 1) = A_b(t)\phi(t) + B_3(t)y^*(t + d + 1) + B_4(t)w(t + 1), \quad t \geq t_0 - 1; \tag{16}$$

due to the compactness of \mathcal{S}_{ab} , $\mathcal{S}_{\alpha\beta}$ and \mathcal{S} , the following is immediate:

Proposition 2 *There exists a constant c_1 so that for every $t_0 \in \mathbf{Z}$, $x_0 \in \mathbf{R}^{n+m+3d-2}$, $\theta_0 \in \mathcal{S}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, y^* , $w \in l_\infty$, and $\delta \in (0, \infty]$, when the adaptive controller (8), (9) and (11) is applied to the plant (1), the following holds:*

$$\|A_b(t)\| \leq c_1, \quad \|B_3(t)\| \leq c_1, \quad \|B_4(t)\| \leq c_1, \quad t \geq t_0 - 1.$$

3.3 A better closed-loop model

The good closed-loop model (15) is driven by future tracking error signals. We can now combine this with the crude closed-loop model (16) to create a new model which is driven by perturbed versions of the present and past values of ϕ , with the weights associated with the parameter update law. Motivated by the form of the term in the parameter estimator, we first define

$$v(t - 1) := \rho_\delta(\phi(t - d), e(t)) \times \frac{\phi(t - d)}{\|\phi(t - d)\|^2} e(t), \quad t \geq t_0. \tag{17}$$

The following result plays a pivotal role in the analysis of the closed-loop system.

Proposition 3 *There exists a constant c_2 so that for every $t_0 \in \mathbf{Z}$, $x_0 \in \mathbf{R}^{n+m+3d-2}$, $\theta_0 \in \mathcal{S}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, y^* , $w \in l_\infty$, and $\delta \in (0, \infty]$, when the adaptive controller (8), (9) and (11) is applied to the plant (1), the following holds:*

$$\phi(t + 1) = A_g \phi(t) + \sum_{j=0}^{d-1} \Delta_j(t) \phi(t - j) + \eta(t), \quad t \geq t_0 + d - 1,$$

with

$$\|\eta(t)\| \leq c_2(1 + \|v(t + 2)\| + \dots + \|v(t + d + 1)\|) \times \left[\sum_{j=1}^{d+1} |y^*(t + j)| + \sum_{j=1}^{d+1} (|w(t + j)| + |\bar{w}(t + 2 - j)|) \right]$$

and

$$\|\Delta_j(t)\| \leq c_2(\|v(t - d + 1)\| + \dots + \|v(t + d)\|), \quad j = 0, \dots, d - 1.$$

Proof See Appendix. □

To make the model of Proposition 3 amenable to analysis, we define a new extended state variable and associated matrices:

$$\bar{\phi}(t) := \begin{bmatrix} \phi(t) \\ \phi(t - 1) \\ \vdots \\ \phi(t - d + 1) \end{bmatrix}, \quad \bar{A}_g := \begin{bmatrix} A_g & & & & \\ I & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & I & 0 \end{bmatrix}, \tag{18}$$

and

$$\bar{B}_1 := \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \Delta(t) = \begin{bmatrix} \Delta_0(t) & \Delta_1(t) & \cdots & \Delta_{d-1}(t) \\ 0 & \cdots & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{19}$$

which gives rise to a state-space model which will play a key role in our analysis:

$$\bar{\phi}(t + 1) = [\bar{A}_g + \Delta(t)]\bar{\phi}(t) + \bar{B}_1\eta(t), \quad t \geq t_0 + 2d - 2. \tag{20}$$

Now A_g arises from $\theta_{ab}^* \in \mathcal{S}_{ab}$ and lies in a corresponding compact set \mathcal{A} ; furthermore, its eigenvalues are at zero and at the zeros of $B(z^{-1})$, which has all of its roots in the open unit disk, so we can use classical arguments to prove that there exist constants γ and $\sigma \in (0, 1)$ so that for all $\theta_{ab}^* \in \mathcal{S}_{ab}$, we have

$$\|\bar{A}_g^i\| \leq \gamma\sigma^i, \quad i \geq 0; \tag{21}$$

indeed, for every σ larger than

$$\underline{\lambda} := \max_{\theta_{ab}^* \in \mathcal{S}_{ab}} \{|\lambda| : \lambda \in \mathbf{C} \text{ and } B(\lambda^{-1}) = 0\},$$

we can choose γ so that (21) holds.

Equations of the form given in (20) arise in classical adaptive control approaches. While we can view (20) as a linear time-varying system, we have to bear in mind that both $\Delta(t)$ and $\eta(t)$ are (nonlinear) implicit functions of all of the data: θ_{ab}^* , θ_0 , x_0 , y^* and w . That being said, the linear time-varying interpretation is very convenient for analysis; to this end, we let Φ_A denote the state transition matrix of a general time-varying square matrix A . The following result will prove useful in analysing our closed-loop systems on sub-intervals for which the constraints hold.

Proposition 4 *With $\sigma \in (\underline{\lambda}, 1)$, suppose that $\gamma \geq 1$ is such that (21) is satisfied for every $A_g \in \mathcal{A}$. For every $\mu \in (\sigma, 1)$, $\beta_0 \geq 0$, $\beta_1 \geq 0$, and*

$$\beta_2 \in \left[0, \frac{1}{\gamma}(\mu - \sigma)\right),$$

there exists a constant $\bar{\gamma} \geq 1$ so that for every $A_g \in \mathcal{A}$ and $\Delta \in \mathcal{S}(\mathbf{R}^{(n+m+d)d \times (n+m+d)d})$ satisfying

$$\sum_{i=\underline{t}}^{\bar{t}-1} \|\Delta(i)\| \leq \beta_0 + \beta_1(t - \tau)^{1/2} + \beta_2(t - \tau), \quad \underline{t} \leq \tau \leq t - 1 \leq \bar{t}, \tag{22}$$

we have

$$\|\Phi_{\bar{A}_g + \Delta}(t, \tau)\| \leq \bar{\gamma}\mu^{t-\tau}, \quad \underline{t} \leq \tau < t \leq \bar{t}. \tag{23}$$

Proof Fix $\sigma \in (\underline{\lambda}, 1)$ and $\gamma \geq 1$ so that (21) is satisfied for every $A_g \in \mathcal{A}$. For every $\mu \in (\sigma, 1)$, $\beta_0 \geq 0$, $\beta_1 \geq 0$, and

$$\beta_2 \in \left[0, \frac{1}{\gamma}(\mu - \sigma) \right),$$

it follows from the lemma of Kreisselmeier [10] that there exists a constant $\bar{\gamma}$, which is independent of $A_g \in \mathcal{A}$ (though dependant on β_2, γ, μ , and σ), so that if (22) holds then (23) holds as well. □

In the next section, we prove that the closed-loop system is exponentially stable and that there is a convolution bound on the closed-loop behaviour. Following that, we analyse robustness and address the tracking problem.

4 Closed-loop stability

Theorem 1 *For every $\delta \in (0, \infty]$ and $\lambda \in (\underline{\lambda}, 1)$, there exists a constant $c > 0$ so that for every $t_0 \in \mathbf{Z}$, plant parameter $\theta_{ab}^* \in \mathcal{S}_{ab}$, exogenous signals $y^*, w \in \ell_\infty$, estimator initial condition $\theta_0 \in \mathcal{S}$, and plant initial condition*

$$x_0 = [y(t_0 - 1) \cdots y(t_0 - n - d + 1) \ u(t_0 - 1) \cdots u(t_0 - m - 2d + 1)]^T,$$

when the adaptive controller (8), (9) and (11) is applied to the plant (1), the following bound holds:

$$\|\phi(k)\| \leq c\lambda^{k-t_0} \|x_0\| + \sum_{j=t_0}^k c\lambda^{k-j} (|y^*(j+d)| + |w(j)|), \quad k \geq t_0. \tag{24}$$

Remark 7 Theorem 1 implies that the system has a bounded gain (from y^* and w to y) in every p -norm.

Remark 8 Most adaptive controllers are proven to yield a weak form of stability, such as boundedness (in the presence of a non-zero disturbance) or asymptotic stability (in the case of a zero disturbance), which means that details surrounding initial conditions can be ignored. Here the goal is to prove a stronger, linear-like, convolution bound as well as exponential stability, so it requires a much more detailed analysis.

Proof Fix $\delta \in (0, \infty]$ and $\lambda \in (\underline{\lambda}, 1)$, and let $t_0 \in \mathbf{Z}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, $y^*, w \in \ell_\infty$, $\theta_0 \in \mathcal{S}$, and $x_0 \in \mathbf{R}^{n+m+3d-2}$ be arbitrary. Now choose $\sigma \in (\underline{\lambda}, \lambda)$. Observe that x_0 gives rise to $\phi(t_0 - 1), \dots, \phi(t_0 - d + 1)$, and therefore $\bar{\phi}(t_0 - 1)$, which we label $\bar{\phi}_0$; it is clear that $\|\bar{\phi}_0\| \leq d\|x_0\|$ and

$$\|\phi(t_0 - j)\| \leq \|x_0\|, \quad j = 1, \dots, d - 1.$$

Step 1 Preamble and preliminary results

To proceed we will analyse (20), namely

$$\bar{\phi}(t + 1) = [\bar{A}_g + \Delta(t)]\bar{\phi}(t) + \bar{B}_1\eta(t), \quad t \geq t_0 + 2d - 2, \tag{25}$$

and obtain a bound on $\bar{\phi}(t)$ in terms of $\eta(t)$, $\bar{w}(t)$, and $y^*(t)$, which we will then convert to the desired form. First of all, we see from Proposition 3 that there exists a constant γ_1 so that⁷

$$\begin{aligned} \|\eta(t)\| &\leq \gamma_1(1 + \|v(t + 2)\| + \dots + \|v(t + d + 1)\|) \\ &\quad \times \left[\sum_{j=1}^{d+1} |y^*(t + j)| + \sum_{j=1}^{d+1} (|w(t + j)| + |\bar{w}(t + 2 - j)|) \right] \\ &\leq \gamma_1(1 + \|v(t + 2)\| + \dots + \|v(t + d + 1)\|) \\ &\quad \times \underbrace{\left[\sum_{j=1}^{d+1} |y^*(t + j)| + \sum_{j=1-d}^{d+1} (|w(t + j)| + |\bar{w}(t + j)|) \right]}_{=:\tilde{w}(t)}, \end{aligned} \tag{26}$$

and

$$\|\Delta(t)\| \leq \gamma_1(\|v(t - d + 1)\| + \dots + \|v(t + d)\|), \quad t \geq t_0 + d - 1. \tag{27}$$

Second of all, following Sect. 3.3 we see that there exists a constant γ_2 so that for every $A_g \in \mathcal{A}$, we have that the corresponding matrix \bar{A}_g satisfies

$$\|\bar{A}_g^i\| \leq \gamma_2\sigma^i, \quad i \geq 0.$$

Before proceeding, we provide several preliminary results. The first shows that we can always obtain a nice bound on the closed-loop behaviour on a short interval.

Claim 1 There exists a constant γ_3 so that for every $\underline{t} \geq t_0 - 1$, we have

$$\|\phi(t)\| \leq \gamma_3\lambda^{t-\underline{t}}\|\phi(\underline{t})\| + \gamma_3 \sum_{j=\underline{t}}^{t-1} \lambda^{t-1-j}|\tilde{w}(j)|, \quad t = \underline{t}, \underline{t} + 1, \dots, \underline{t} + 4d.$$

Proof of Claim 1 Using the crude model given in (16) together with Proposition 2, we see that there exists a constant $c_1 \geq 1$ so that

$$\|\phi(t + 1)\| \leq c_1\|\phi(t)\| + c_1|y^*(t + d + 1)| + c_1|w(t + 1)|, \quad t \geq t_0 - 1.$$

Using the definition of $\tilde{w}(t)$, this immediately implies that

$$\|\phi(t + 1)\| \leq c_1\|\phi(t)\| + c_1|\tilde{w}(t)|, \quad t \geq t_0 - 1. \tag{28}$$

⁷ We've chosen to define $\tilde{w}(t)$ in this way to simplify the analysis later on in the proof.

Now let $t \geq t_0 - 1$ be arbitrary. By solving (28) from $t = t$, we have

$$\begin{aligned} \|\phi(t + 1)\| &\leq c_1^{t-t} \|\phi(t)\| + \sum_{j=t}^{t-1} c_1^{t-1-j} |\tilde{w}(j)| \\ &\leq \left(\frac{c_1}{\lambda}\right)^{t-t} [\lambda^{t-t} \|\phi(t)\| + \sum_{j=t}^{t-1} \lambda^{t-1-j} |\tilde{w}(j)|], \quad t \geq t. \end{aligned}$$

If we set $\gamma_3 := \left(\frac{c_1}{\lambda}\right)^{4d}$, then the result follows. □

In order to apply Proposition 4, we need to compute a bound on the sum of $\|\Delta(t)\|$ terms; the following result provides an avenue.

Claim 2 There exists a constant γ_4 so that

$$\sum_{j=t_1}^{t_2-1} \|\Delta(j)\| \leq \gamma_4 \left[\sum_{j=t_1-d+1}^{t_2+d-1} \|v(j)\|^2 \right]^{1/2} (t_2 - t_1)^{1/2}, \quad t_2 > t_1 \geq t_0 + d - 1.$$

Proof of Claim 2 Since there are $2d$ terms in the RHS of (27), it is easy to see that

$$\sum_{j=t_1}^{t_2-1} \|\Delta(j)\| \leq 2d \sum_{j=t_1-d+1}^{t_2+d-1} \|v(j)\|.$$

If we apply the Cauchy–Schwarz inequality, then we have

$$\sum_{j=t_1-d+1}^{t_2+d-1} \|v(j)\| \leq \left[\sum_{j=t_1-d+1}^{t_2+d-1} \|v(j)\|^2 \right]^{1/2} (t_2 - t_1 + 2d - 1)^{1/2}.$$

But

$$(t_2 - t_1 + 2d - 1)^{1/2} \leq (2d)^{1/2} (t_2 - t_1)^{1/2}, \quad t_2 > t_1,$$

so the result follows. □

Step 2 Partition the time-line.

Now we consider the closed-loop system behaviour. To proceed, we partition the timeline into two parts: one in which the noise is small and one where it is not; the idea is that when the noise is small, the closed-loop behaviour should be similar to that when the noise is zero. Of course, here we have to choose the notion of “small” carefully: we do so by using the scaled version of the noise, namely $\frac{|\tilde{w}(t)|}{\|\phi(t)\|}$, which appears in the bound on V given in Proposition 1. To this end, with $\xi > 0$ to be chosen shortly, partition $\{j \in \mathbf{Z} : j \geq t_0\}$ into

$$S_{\text{good}} := \left\{ j \geq t_0 : \phi(j) \neq 0 \text{ and } \frac{[\bar{w}(j)]^2}{\|\phi(j)\|^2} < \xi \right\},$$

$$S_{\text{bad}} := \left\{ j \geq t_0 : \phi(j) = 0 \text{ or } \frac{[\bar{w}(j)]^2}{\|\phi(j)\|^2} \geq \xi \right\};$$

clearly $\{j \in \mathbf{Z} : j \geq t_0\} = S_{\text{good}} \cup S_{\text{bad}}$.⁸ Observe that this partition implicitly depends on the system parameters $\theta_{ab}^* \in \mathcal{S}_{ab}$, as well as the initial conditions. We will apply Proposition 4 to analyse the closed-loop system behaviour on S_{good} ; on the other hand, we will easily obtain bounds on the system behaviour on S_{bad} . Before doing so, we partition the time index $\{j \in \mathbf{Z} : j \geq t_0\}$ 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) $k_1 = t_0$;
- (ii) $[k_i, k_{i+1})$ either belongs to S_{good} or S_{bad} ; and
- (iii) if $k_{i+1} \neq \infty$ 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 analyse the behaviour during each interval.

Step 3 The closed-loop behaviour on S_{bad} .

Let $j \in [k_i, k_{i+1})$ be arbitrary. In this case, either $\phi(j) = 0$ or $\frac{[\bar{w}(j)]^2}{\|\phi(j)\|^2} \geq \xi$ holds. In either case, we have

$$\|\phi(j)\| \leq \frac{1}{\xi^{1/2}} |\bar{w}(j)|, \quad j \in [k_i, k_{i+1}). \tag{29}$$

From the crude model (16) and Proposition 2, we have

$$\begin{aligned} \|\phi(j+1)\| &\leq \frac{c_1}{\xi^{1/2}} |\bar{w}(j)| + c_1 |y^*(j+d+1)| + c_1 |w(j+1)| \\ &\leq \left(\frac{c_1}{\xi^{1/2}} + c_1 \right) |\tilde{w}(j)|, \quad j \in [k_i, k_{i+1}). \end{aligned}$$

If we combine this with (29), we end up with

$$\|\phi(j)\| \leq \begin{cases} \frac{1}{\xi^{1/2}} |\tilde{w}(j)| & \text{if } j = k_i \\ c_1 \left(1 + \frac{1}{\xi^{1/2}} \right) |\tilde{w}(j-1)| & \text{if } j = k_i + 1, \dots, k_{i+1}. \end{cases} \tag{30}$$

Step 4 The closed-loop behaviour on S_{good} .

Suppose that $[k_i, k_{i+1})$ lies in S_{good} ; this case is much more involved than in the proof of [20,25] since the bound on $\|\Delta(t)\|$ provided by Claim 2 extends both forward and backward in time, occasionally outside S_{good} . Furthermore, the difference equation for $\bar{\phi}$ given in (25) only holds for $t \geq t_0 + 2d - 2$, so if $k_i = t_0$ then we cannot use

⁸ If the disturbance is zero, then S_{good} may be the whole time-line $[t_0, \infty)$.

it until $k_i + 2d - 2$. For this reason, we will handle the first $2d$ and last $2d$ time units separately.

To this end, first suppose that $k_{i+1} \leq k_i + 4d$. Then by Claim 1, we see that there exists a constant γ_5 so that

$$\|\phi(k)\| \leq \gamma_5 \lambda^{k-k_i} \|\phi(k_i)\| + \sum_{j=k_i}^{k-1} \gamma_5 \lambda^{k-1-j} |\tilde{w}(j)|, \quad k_i \leq k \leq k_{i+1}. \quad (31)$$

Now suppose that $k_{i+1} > k_i + 4d$. Define $\bar{k}_i := k_i + 2d$ and $\underline{k}_{i+1} := k_{i+1} - 2d$. By the definition of S_{good} and the fact that $\rho_\delta(\cdot, \cdot) \in [0, 1]$, it follows that

$$\rho_\delta(\phi(j), e(j+d)) \frac{\bar{w}(j)^2}{\|\phi(j)\|^2} < \xi, \quad j \in [k_i, k_{i+1}),$$

so

$$\rho_\delta(\phi(j-d+1), e(j+1)) \frac{\bar{w}(j-d+1)^2}{\|\phi(j-d+1)\|^2} < \xi, \quad j \in [k_i+d-1, k_{i+1}+d-1). \quad (32)$$

By Proposition 1, we see that

$$\begin{aligned} V(\bar{k}) &\leq V(\underline{k}) + \sum_{j=\underline{k}}^{\bar{k}-1} \rho_\delta(\phi(j-d+1), e(j+1)) \\ &\quad \times \left[-\frac{1}{2} \frac{[e(j+1)]^2}{\|\phi(j-d+1)\|^2} + 2 \frac{[\bar{w}(j-d+1)]^2}{\|\phi(j-d+1)\|^2} \right], \quad k_i \leq \underline{k} < \bar{k} \leq k_{i+1}. \end{aligned}$$

Using the fact that the first term in the sum is $v(j)$, using (32) to provide a bound on the second term in the sum, and using the fact that

$$V(\underline{k}) = \|\hat{\theta}(\underline{k}) - \theta^*\| \leq 4\|\mathcal{S}\|^2,$$

it follows that

$$\sum_{j=\underline{k}}^{\bar{k}-1} \|v(j)\|^2 \leq 8\|\mathcal{S}\|^2 + 4\xi(\bar{k} - \underline{k}), \quad k_i + d - 1 \leq \underline{k} < \bar{k} \leq k_{i+1} + d - 2. \quad (33)$$

We would like to leverage this bound on $v(j)$ to obtain a bound on $\|\Delta(j)\|$. From Claim 2

$$\sum_{j=\underline{k}}^{\bar{k}-1} \|\Delta(j)\| \leq \gamma_4 \left[\sum_{j=\underline{k}-d+1}^{\bar{k}+d-1} \|v(j)\|^2 \right]^{1/2} (\bar{k} - \underline{k})^{1/2}, \quad \bar{k} > \underline{k} \geq t_0 + d - 1;$$

as long as we keep

$$\underline{k} - d + 1 \geq k_i + d - 1 \geq t_0 + d - 1 \text{ and } \bar{k} + d - 1 \leq k_{i+1} + d - 2$$

or equivalently

$$\underline{k} \geq k_i + 2d - 2 \geq t_0 + 2d - 2 \text{ and } \bar{k} \leq k_{i+1} - 1,$$

then we can use (33) to provide a bound on the RHS; this will definitely be the case if we restrict

$$\bar{k}_i \leq \underline{k} < \bar{k} \leq \underline{k}_{i+1},$$

resulting in

$$\sum_{j=\underline{k}}^{\bar{k}-1} \|\Delta(j)\| \leq \gamma_4[8\|\mathcal{S}\|^2 + 4\xi(\bar{k} - \underline{k} + 2d - 1)]^{1/2}(\bar{k} - \underline{k})^{1/2}, \bar{k}_i \leq \underline{k} < \bar{k} \leq \underline{k}_{i+1}.$$

If we restrict $\xi \leq 1$, then we obtain

$$\begin{aligned} \sum_{j=\underline{k}}^{\bar{k}-1} \|\Delta(j)\| &\leq \gamma_4[8\|\mathcal{S}\|^2 + 4\xi(2d - 1) + 4\xi(\bar{k} - \underline{k})]^{1/2}(\bar{k} - \underline{k})^{1/2} \\ &\leq \gamma_4[8\|\mathcal{S}\|^2 + 4(2d - 1)]^{1/2}(\bar{k} - \underline{k})^{1/2} \\ &\quad + 2\gamma_4\xi^{1/2}(\bar{k} - \underline{k}), \bar{k}_i \leq \underline{k} < \bar{k} \leq \underline{k}_{i+1}; \end{aligned}$$

if we define

$$\gamma_6 := \gamma_4[8\|\mathcal{S}\|^2 + 4(2d - 1)]^{1/2} + 2\gamma_4,$$

then for $\xi \leq 1$ we have

$$\sum_{j=\underline{k}}^{\bar{k}-1} \|\Delta(j)\| \leq \gamma_6(\bar{k} - \underline{k})^{1/2} + \gamma_6\xi^{1/2}(\bar{k} - \underline{k}), \bar{k}_i \leq \underline{k} < \bar{k} \leq \underline{k}_{i+1}.$$

Now we will apply Proposition 4: we set

$$\beta_0 = 0, \beta_1 = \gamma_6, \beta_2 = \gamma_6\xi^{1/2}, \mu = \lambda, \gamma = \gamma_2;$$

we need $\beta_2 < \frac{1}{\gamma}(\mu - \sigma)$, or equivalently

$$\gamma_6\xi^{1/2} < \frac{1}{\gamma_2}(\lambda - \sigma)$$

$$\Leftrightarrow \xi < \left(\frac{\lambda - \sigma}{\gamma_2 \gamma_6} \right)^2,$$

so we set

$$\xi := \min \left\{ 1, \frac{1}{2} \left(\frac{\lambda - \sigma}{\gamma_2 \gamma_6} \right)^2 \right\}.$$

So from Proposition 4, we see that there exists a constant γ_7 so that the state transition matrix $\Phi_{\bar{A}_g + \Delta}(t, \tau)$ satisfies

$$\|\Phi_{\bar{A}_g + \Delta}(t, \tau)\| \leq \gamma_7 \lambda^{t-\tau}, \quad \bar{k}_i \leq \tau \leq t \leq \underline{k}_{i+1}.$$

In order to solve (25), we need a bound on $\eta(t)$. But (33) implies that

$$|\nu(j)|^2 \leq 8\|\mathcal{S}\|^2 + 4\xi \leq 8\|\mathcal{S}\|^2 + 4, \quad \bar{k}_i \leq j \leq \underline{k}_{i+1},$$

so from (26) we see that

$$\|\eta(t)\| \leq \gamma_1 [1 + d(8\|\mathcal{S}\|^2 + 4)^{1/2}] |\tilde{w}(t)|, \quad \bar{k}_i \leq t \leq \underline{k}_{i+1}.$$

If we use this in the difference equation (25), we see that there exists a constant γ_8 so that

$$\|\bar{\phi}(k)\| \leq \gamma_8 \lambda^{k-\bar{k}_i} \|\bar{\phi}(\bar{k}_i)\| + \sum_{j=\bar{k}_i}^{k-1} \gamma_8 \lambda^{k-1-j} |\tilde{w}(j)|, \quad \bar{k}_i \leq k \leq \underline{k}_{i+1}. \tag{34}$$

However, we would like a bound on the whole interval $[k_i, k_{i+1})$, and we'd like it on ϕ rather than $\bar{\phi}$.

Claim 3 There exists a constant γ_9 so that

$$\|\phi(k)\| \leq \gamma_9 \lambda^{k-k_i} \|\phi(k_i)\| + \sum_{j=k_i}^{k-1} \gamma_9 \lambda^{k-1-j} |\tilde{w}(j)|, \quad k_i \leq k \leq k_{i+1}.$$

Proof of Claim 3 First of all, on the interval $[k_i, \bar{k}_i]$ we can apply Claim 1:

$$\|\phi(k)\| \leq \gamma_3 \lambda^{k-k_i} \|\phi(k_i)\| + \sum_{j=k_i}^{k-1} \gamma_3 \lambda^{k-1-j} |\tilde{w}(j)|, \quad k_i \leq k \leq \bar{k}_i.$$

This provides a bound of the desired form on the first sub-interval of $[k_i, k_{i+1})$.

Next, from (34) we have

$$\|\phi(k)\| \leq \gamma_8 \lambda^{k-\bar{k}_i} \|\bar{\phi}(\bar{k}_i)\| + \sum_{j=\bar{k}_i}^{k-1} \gamma_8 \lambda^{k-1-j} |\tilde{w}(j)|, \quad \bar{k}_i \leq k \leq \underline{k}_{i+1}. \quad (35)$$

Using the definition of $\bar{\phi}$ and Claim 1 (with $\underline{t} = k_i$), we have

$$\begin{aligned} \|\bar{\phi}(\bar{k}_i)\| &\leq \|\phi(\bar{k}_i)\| + \|\phi(\bar{k}_i - 1)\| + \dots + \|\phi(\bar{k}_i - d + 1)\| \\ &\leq \gamma_3 (\lambda^{\bar{k}_i - k_i} + \lambda^{\bar{k}_i - k_i + 1} + \dots + \lambda^{\bar{k}_i - k_i + d - 1}) \|\phi(k_i)\| \\ &\quad + \gamma_3 \left[\sum_{j=k_i}^{\bar{k}_i - 1} \lambda^{\bar{k}_i - 1 - j} |\tilde{w}(j)| + \sum_{j=k_i}^{\bar{k}_i - 2} \lambda^{\bar{k}_i - 2 - j} |\tilde{w}(j)| + \dots + \sum_{j=k_i}^{\bar{k}_i - d} \lambda^{\bar{k}_i - d - j} |\tilde{w}(j)| \right] \\ &\leq \underbrace{\frac{\gamma_3 d}{\lambda^{d-1}}}_{=: \bar{\gamma}_3} \lambda^{\bar{k}_i - k_i} \|\phi(k_i)\| + \gamma_3 \left[\sum_{j=k_i}^{\bar{k}_i - 1} \lambda^{\bar{k}_i - 1 - j} |\tilde{w}(j)| \right. \\ &\quad \left. + \frac{1}{\lambda} \sum_{j=k_i}^{\bar{k}_i - 2} \lambda^{\bar{k}_i - 1 - j} |\tilde{w}(j)| + \dots + \frac{1}{\lambda^{d-1}} \sum_{j=k_i}^{\bar{k}_i - d} \lambda^{\bar{k}_i - 1 - j} |\tilde{w}(j)| \right] \\ &\leq \bar{\gamma}_3 \lambda^{\bar{k}_i - k_i} \|\phi(k_i)\| + \bar{\gamma}_3 \sum_{j=k_i}^{\bar{k}_i - 1} \lambda^{\bar{k}_i - 1 - j} |\tilde{w}(j)|. \end{aligned}$$

If we now substitute this into (35), then we have

$$\begin{aligned} \|\phi(k)\| &\leq \gamma_8 \bar{\gamma}_3 \lambda^{k-\bar{k}_i} \|\phi(k_i)\| + \gamma_8 \bar{\gamma}_3 \lambda^{k-\bar{k}_i} \sum_{j=k_i}^{\bar{k}_i - 1} \lambda^{\bar{k}_i - 1 - j} |\tilde{w}(j)| \\ &\quad + \gamma_8 \sum_{j=k_i}^{k-1} \lambda^{k-1-j} |\tilde{w}(j)|, \quad \bar{k}_i \leq k \leq \underline{k}_{i+1}; \end{aligned}$$

if we define $\bar{\gamma}_8 := \gamma_8 + \gamma_8 \bar{\gamma}_3$, then it follows that

$$\|\phi(k)\| \leq \bar{\gamma}_8 \lambda^{k-k_i} \|\phi(k_i)\| + \bar{\gamma}_8 \sum_{j=k_i}^{k-1} \lambda^{k-1-j} |\tilde{w}(j)|, \quad \bar{k}_i \leq k \leq \underline{k}_{i+1}. \quad (36)$$

This provides a bound of the desired form on the second sub-interval of $[k_i, k_{i+1}]$.

Last of all, we would like to obtain a bound of the desired form on $[\underline{k}_{i+1}, k_{i+1}]$. By Claim 1 we have

$$\|\phi(k)\| \leq \gamma_3 \lambda^{k-k_{i+1}} \|\phi(k_{i+1})\| + \sum_{j=k_{i+1}}^{k-1} \gamma_3 \lambda^{k-1-j} |\tilde{w}(j)|, \quad k_{i+1} \leq k \leq k_{i+1}.$$

Using (36) to obtain a bound on $\|\phi(k_{i+1})\|$, we obtain

$$\begin{aligned} \|\phi(k)\| &\leq \gamma_3 \lambda^{k-k_{i+1}} \left[\bar{\gamma}_8 \lambda^{k_{i+1}-k_i} \|\phi(k_i)\| + \bar{\gamma}_8 \sum_{j=k_i}^{k_{i+1}-1} \lambda^{k_{i+1}-1-j} |\tilde{w}(j)| \right] \\ &\quad + \sum_{j=k_{i+1}}^{k-1} \gamma_3 \lambda^{k-1-j} |\tilde{w}(j)| \\ &\leq \gamma_3 \bar{\gamma}_8 \lambda^{k-k_i} \|\phi(k_i)\| + (\gamma_3 \bar{\gamma}_8 + \gamma_3) \sum_{j=k_i}^{k-1} \lambda^{k-1-j} |\tilde{w}(j)|, \quad k_{i+1} \leq k \leq k_{i+1}. \end{aligned}$$

So if we set

$$\gamma_9 := \max\{\gamma_3, \bar{\gamma}_8, \gamma_3 \bar{\gamma}_8 + \gamma_3\},$$

then the result holds. □

Step 5 Analysing the whole time-line.

At this point, we glue together the bounds obtained on S_{bad} and S_{good} together to obtain a bound which holds on all of $[t_0, \infty)$.

Claim 4 There exists a constant γ_{10} so that the following bound holds:

$$\|\phi(k)\| \leq \gamma_{10} \lambda^{k-t_0} \|\phi(t_0)\| + \sum_{j=t_0}^k \gamma_{10} \lambda^{k-j} |\tilde{w}(j)|, \quad k \geq t_0. \tag{37}$$

Proof of Claim 4 If $[k_1, k_2) = [t_0, k_2) \subset S_{\text{good}}$, then (37) holds for $k \in [t_0, k_2]$ by Claim 3 as long as

$$\gamma_{10} \geq \max\left\{\gamma_9, \frac{\gamma_9}{\lambda}\right\}. \tag{38}$$

If $[k_1, k_2) = [t_0, k_2) \subset S_{\text{bad}}$, then from (30) we see that the bound holds as long as

$$\gamma_{10} \geq \max\left\{\frac{1}{\xi^{1/2}}, \frac{c_1}{\lambda} \left(1 + \frac{1}{\xi^{1/2}}\right)\right\}. \tag{39}$$

We now use induction - suppose that (37) holds for $k \in [k_1, k_i]$; we need to prove that it holds for $k \in (k_i, k_{i+1}]$ as well. If $[k_i, k_{i+1}) \subset S_{\text{bad}}$, then from (30) we see that the bound holds on $(k_i, k_{i+1}]$ as long (39) holds. On the other hand, if $[k_i, k_{i+1}) \subset S_{\text{good}}$, then $k_i - 1 \in S_{\text{bad}}$; from (30) we have that

$$\|\phi(k_i)\| \leq c_1 \left(1 + \frac{\gamma_3}{\xi^{1/2}}\right) |\tilde{w}(k_i - 1)|;$$

combining this with Claim 3, which provides a bound on $\|\phi(k)\|$ for $k \in [k_i, k_{i+1})$, we have

$$\begin{aligned} \|\phi(k)\| &\leq \gamma_9 \lambda^{k-k_i} \|\phi(k_i)\| + \sum_{j=k_i}^{k-1} \gamma_9 \lambda^{k-1-j} |\tilde{w}(j)| \\ &\leq \gamma_9 \lambda^{k-k_i} c_1 \left(1 + \frac{\gamma_3}{\xi^{1/2}}\right) |\tilde{w}(k_i - 1)| + \sum_{j=k_i}^{k-1} \gamma_9 \lambda^{k-1-j} |\tilde{w}(j)| \\ &\leq \left[\gamma_9 c_1 \left(1 + \frac{\gamma_3}{\xi^{1/2}}\right) + \gamma_9 \right] \sum_{j=k_i-1}^{k-1} \gamma_9 \lambda^{k-1-j} |\tilde{w}(j)|, \quad k \in [k_i, k_{i+1}). \end{aligned}$$

So the bound holds in this case as long as

$$\gamma_{10} \geq \frac{1}{\lambda} \left[\gamma_9 c_1 \left(1 + \frac{\gamma_3}{\xi^{1/2}}\right) + \gamma_9 \right].$$

□

Step 6 Obtaining a bound of the desired form

The last step is to convert the bound proven in Claim 4 to one of the desired form, i.e. we need to replace \tilde{w} with w and y^* . First of all, using the definition of \tilde{w} given in (26), we see that

$$\tilde{w}(t) = \sum_{j=1}^{d+1} |y^*(t + j)| + \sum_{j=1-d}^{d+1} (|w(t + j)| + |\bar{w}(t + j)|).$$

From the definition of $\bar{w}(t)$, we have that

$$|\bar{w}(t)| \leq |f_0| |w(t + d)| + \dots + |f_{d-1}| |w(t + 1)|;$$

the f_i 's arise from long division and are continuous functions of $\theta_{ab}^* \in \mathcal{S}_{ab}$, so it follows that exists a constant γ_{11} so that for all $\theta_{ab}^* \in \mathcal{S}_{ab}$, we have

$$|\tilde{w}(t)| \leq \sum_{j=1}^{d+1} |y^*(t + j)| + \gamma_{11} \sum_{j=1-d}^{2d+1} |w(t + j)|.$$

It turns out that some of the terms in the sum are superfluous:

- (i) We see from the control law (11) that $y^*(t)$ affects $\phi(t)$ only via $u(t)$. Indeed, $u(t)$ explicitly depends on $y^*(t + d)$ which means, by causality, that $y^*(t + d + 1)$ has no effect on $\phi(t)$.
- (ii) We see from the original plant equation (1) that $w(t)$ affects $\phi(t)$ via $y(t)$. Indeed, $y(t)$ explicitly depends on $w(t)$, which means, by causality, that $\sum_{j=1}^{2d+1} |w(t + j)|$ can have no effect on $\phi(t)$.

If we combine these observations with Claim 4, we see that the bound becomes

$$\begin{aligned} \|\phi(k)\| \leq & \gamma_{10}\lambda^{k-t_0}\|\phi(t_0)\| + \sum_{j=t_0}^k \gamma_{10}\lambda^{k-j} [|y^*(j+1)| + \dots + |y^*(j+d)| \\ & + \gamma_{11}|w(j-d+1)| + \dots + \gamma_{11}|w(j)|]; \end{aligned}$$

if we examine each term in the summation, it is straight-forward to verify that

$$\begin{aligned} \|\phi(k)\| \leq & \gamma_{10}\lambda^{k-t_0}\|\phi(t_0)\| + \frac{\gamma_{10}d}{\lambda^{d-1}} \sum_{j=t_0-d+1}^k \lambda^{k-j}|y^*(j+d)| \\ & + \frac{\gamma_{10}\gamma_{11}d}{\lambda^{d-1}} \sum_{j=t_0-d+1}^k \lambda^{k-j}|w(j)| \\ \leq & \gamma_{10}\lambda^{k-t_0}\|\phi(t_0)\| + \frac{\gamma_{10}d(1+\gamma_{11})}{\lambda^{d-1}} \sum_{j=t_0}^k \lambda^{k-j} (|y^*(j+d)| + |w(j)|) \\ & + \underbrace{\frac{\gamma_{10}d}{\lambda^{d-1}} \sum_{j=t_0-d+1}^{t_0-1} \lambda^{k-j}|y^*(j+d)|}_{\psi_1(k)} + \underbrace{\frac{\gamma_{10}\gamma_{11}d}{\lambda^{d-1}} \sum_{j=t_0-d+1}^{t_0-1} \lambda^{k-j}|w(j)|}_{=: \psi_2(k)}. \end{aligned}$$

Now y^* affects ϕ via the control signal, and it is clear that $y^*(k)$ for $k < t_0 + d$ has no effect on $\phi(k)$, $k \geq t_0$, so the $\psi_1(k)$ term can be ignored.⁹ On the other hand, the $\psi_2(k)$ term does have an effect: first observe that ψ_2 can be rewritten as

$$\psi_2(k) = \frac{\gamma_{10}\gamma_{11}d}{\lambda^{d-1}} \left[\sum_{j=t_0-d+1}^{t_0-1} \lambda^{t_0-j}|w(j)| \right] \lambda^{k-t_0};$$

second of all, it follows from the plant equation (1) that each of $\{w(t_0 - d + 1), \dots, w(t_0 - 1)\}$ can be rewritten as a linear function of x_0 , so it follows that there exists a constant γ_{12} so that

$$\|\psi_2(k)\| \leq \gamma_{12}\lambda^{k-t_0}\|x_0\|, \quad k \geq t_0.$$

Since $\|\phi(t_0)\| \leq \|x_0\|$, we conclude that

$$\begin{aligned} \|\phi(k)\| \leq & (\gamma_{10} + \gamma_{12})\lambda^{k-t_0}\|x_0\| \\ & + \frac{\gamma_{10}d(1+\gamma_{11})}{\lambda^{d-1}} \sum_{j=t_0}^k \lambda^{k-j} (|y^*(j+d)| + |w(j)|), \quad k \geq t_0, \end{aligned}$$

as desired. □

⁹ Alternatively, we can argue that the closed-loop system response for $k \geq t_0$ is identical if we were to replace y^* by another signal which is the same for $t \geq t_0 + d$ and zero elsewhere, in which case the corresponding $\psi_1(k) = 0$ for $k \geq t_0$.

5 Robustness to time-variations and unmodelled dynamics

It turns out that the exponential stability property and the convolution bounds proven in Theorem 1 will guarantee robustness to a degree of time-variations and unmodelled dynamics; in this way, the approach has a lot in common with an LTI closed-loop system, which also enjoys this feature. Indeed, as we have recently proven in [36], this is true even in a more general situation than that considered here. At this point, we will state the key result and then simply refer to that paper for the proof.

To proceed, consider a time-varying version of the reparameterized plant (2), in which θ^* is now time-varying and there are some unmodelled dynamics which enter the system via $\bar{w}_\Delta(t)$:

$$y(t + d) = \theta^*(t)^T \phi(t) + \bar{w}(t) + \bar{w}_\Delta(t). \tag{40}$$

- (i) *Time variation model* Since the mapping from \mathcal{S}_{ab} to $\mathcal{S}_{\alpha\beta}$ is analytic and both sets are compact, we may as well consider the time-variations in the latter setting. We adopt a common model of acceptable time-variations used in adaptive control, e.g. see [10]: with $c_0 \geq 0$ and $\epsilon > 0$, we let $s(\mathcal{S}_{\alpha\beta}, c_0, \epsilon)$ denote the subset of $\ell_\infty(\mathbf{R}^{n+m+d})$ whose elements θ^* satisfy $\theta^*(t) \in \mathcal{S}_{\alpha\beta}$ for every $t \in \mathbf{Z}$ as well as

$$\sum_{t=t_1}^{t_2-1} \|\theta^*(t + 1) - \theta^*(t)\| \leq c_0 + \epsilon(t_2 - t_1), \quad t_2 > t_1 \tag{41}$$

for every $t_1 \in \mathbf{Z}$.

- (ii) *Unmodelled dynamics* We also adopt a common model of unmodelled dynamics:

$$m(t + 1) = \beta m(t) + \beta \|\phi(t)\|, \quad m(t_0) = m_0 \tag{42}$$

$$\bar{w}_\Delta(t) \leq \mu m(t) + \mu \|\phi(t)\|. \tag{43}$$

As argued in [25], this encapsulates a large class of additive, multiplicative and uncertainty in a coprime factorization, which is common in the robust control literature, e.g. see [46], and is commonly used in the adaptive control literature, e.g. see [11].

We will now show that if the time-variations are slow enough and the size of the unmodelled dynamics is small enough, then the closed-loop system retains exponential stability as well as the convolution bounds.

Theorem 2 *For every $\delta \in (0, \infty]$, $c_0 \geq 0$ and $\beta \in (0, 1)$, there exist an $\epsilon > 0$, $\mu > 0$, $\lambda \in (\max\{\beta, \lambda\}, 1)$ and $c > 0$ so that for every $t_0 \in \mathbf{Z}$, plant parameter $\theta^* \in s(\mathcal{S}_{\alpha\beta}, c_0, \epsilon)$, exogenous signals y^* , $w \in \ell_\infty$, estimator initial condition $\theta_0 \in \mathcal{S}$, and plant initial condition*

$$x_0 = [y(t_0 - 1) \cdots y(t_0 - n - d + 1) \ u(t_0 - 1) \cdots u(t_0 - m - 2d + 1)]^T,$$

when the adaptive controller (8), (9) and (11) is applied to the plant (40) with \bar{w}_Δ satisfying (42), (43), the following bound holds:

$$\|\phi(k)\| \leq c\lambda^{k-t_0} \left\| \begin{bmatrix} x_0 \\ m_0 \end{bmatrix} \right\| + \sum_{j=t_0}^k c\lambda^{k-j} (|y^*(j+d)| + |w(j)|), \quad k \geq t_0. \quad (44)$$

Proof This was proven in [25] for the case of pole placement and then extended to a more general setting in Theorems 1 and 3 of [36]; the setup here is a special case of the latter, so the result follows immediately from that. \square

6 Tracking

We now move from the stability problem to the much harder tracking problem. We first derive a very useful bound on the tracking error in terms of the prediction error. Following that, we analyse the case when the disturbance is absent: we start with the original LTI case, and then we move to the situation in which the parameters are slowly time-varying. Last of all, we consider the original LTI case with a disturbance.

6.1 A useful bound

The d -step-ahead control law (11) can be rewritten as

$$y^*(t) = \hat{\theta}(t-d)^\top \phi(t-d), \quad t \geq t_0 + d.$$

Since the tracking error is $\varepsilon(t) = y(t) - y^*(t)$, we have

$$\varepsilon(t) = y(t) - \hat{\theta}(t-d)^\top \phi(t-d), \quad t \geq t_0 + d.$$

Combining the above with the prediction error definition in (5), we easily obtain

$$\varepsilon(t) = e(t) + \phi(t-d)^\top \left[\hat{\theta}(t-1) - \hat{\theta}(t-d) \right], \quad t \geq t_0 + d. \quad (45)$$

Observe that the relationship in (45) is true irrespective of the plant model, i.e. it need not satisfy (1) or (2).

Now we turn to the prediction and tracking errors scaled by the state ϕ . For $t \geq t_0 + d$, if $\|\phi(t-d)\| \neq 0$ then it follows from (45) that

$$\frac{|\varepsilon(t)|}{\|\phi(t-d)\|} \leq \frac{|e(t)|}{\|\phi(t-d)\|} + \|\hat{\theta}(t-1) - \hat{\theta}(t-d)\|;$$

multiplying both sides by $\rho_\delta(\phi(t - d), e(t))$ yields

$$\begin{aligned} \rho_\delta(\phi(t - d), e(t)) \times \frac{|\varepsilon(t)|}{\|\phi(t - d)\|} &\leq \rho_\delta(\phi(t - d), e(t)) \times \frac{|e(t)|}{\|\phi(t - d)\|} \\ &\quad + \rho_\delta(\phi(t - d), e(t)) \times \|\hat{\theta}(t - 1) - \hat{\theta}(t - d)\| \\ &\leq \rho_\delta(\phi(t - d), e(t)) \times \frac{|e(t)|}{\|\phi(t - d)\|} \\ &\quad + \|\hat{\theta}(t - 1) - \hat{\theta}(t - d)\|, \\ &\quad t \geq t_0 + d. \end{aligned}$$

It turns out that the estimator property in (10) holds irrespective of the plant model; if we use this to yield a bound on the second term on the right-hand side above, then we obtain

$$\begin{aligned} &\rho_\delta(\phi(t - d), e(t)) \frac{|\varepsilon(t)|}{\|\phi(t - d)\|} \\ &\leq \sum_{j=0}^{d-1} \rho_\delta(\phi(t - d - j), e(t - j)) \frac{|e(t - j)|}{\|\phi(t - d - j)\|}, \quad t \geq t_0 + d. \end{aligned}$$

By Cauchy–Schwartz

$$\begin{aligned} &\rho_\delta(\phi(t - d), e(t)) \frac{|\varepsilon(t)|^2}{\|\phi(t - d)\|^2} \\ &\leq \left(\sum_{j=0}^{d-1} \rho_\delta(\phi(t - d - j), e(t - j)) \frac{|e(t - j)|}{\|\phi(t - d - j)\|} \right)^2 \\ &\leq \left(\sum_{j=0}^{d-1} \rho_\delta(\phi(t - d - j), e(t - j)) \frac{|e(t - j)|^2}{\|\phi(t - d - j)\|^2} \right) \left(\sum_{j=0}^{d-1} 1 \right) \\ &= d \sum_{j=0}^{d-1} \rho_\delta(\phi(t - d - j), e(t - j)) \frac{|e(t - j)|^2}{\|\phi(t - d - j)\|^2}, \quad t \geq t_0 + d. \end{aligned}$$

Hence, it follows that for $\bar{t} \geq t_0 + 2d - 1$ we obtain

$$\begin{aligned} &\sum_{j=\bar{t}}^t \rho_\delta(\phi(j - d), e(j)) \frac{|\varepsilon(j)|^2}{\|\phi(j - d)\|^2} \\ &\leq \sum_{j=\bar{t}}^t \left(d \sum_{q=0}^{d-1} \rho_\delta(\phi(j - d - q), e(j - q)) \frac{|e(j - q)|^2}{\|\phi(j - d - q)\|^2} \right) \end{aligned}$$

$$\leq d^2 \sum_{j=\bar{t}-d+1}^t \rho_\delta(\phi(j-d), e(j)) \frac{|e(j)|^2}{\|\phi(j-d)\|^2},$$

$$t \geq \bar{t} \geq t_0 + 2d - 1. \tag{46}$$

The inequality in (46) will prove to be useful in the forth-coming analysis.

In the next two sub-sections, we analyse the case when the disturbance is absent; we start with the case with the original LTI case, and then we move to the situation in which the parameters are slowly time-varying. After that, we analyse the general case.

6.2 The LTI case: no disturbance

In the literature, it is typically proven that the tracking error is square summable, e.g. see [5]. Here we can prove an explicit bound on the 2-norm of the error signal in terms of the plant initial condition and the size of the reference signal, which is a significant improvement.

Theorem 3 *For every $\delta \in (0, \infty]$ and $\lambda \in (\underline{\lambda}, 1)$ there exists a constant $c > 0$ so that for every $t_0 \in \mathbf{Z}$, plant parameter $\theta_{ab}^* \in \mathcal{S}_{ab}$, exogenous signal $y^* \in \ell_\infty$, estimator initial condition $\theta_0 \in \mathcal{S}$, and plant initial condition*

$$x_0 = [y(t_0 - 1) \cdots y(t_0 - n - d + 1) u(t_0 - 1) \cdots u(t_0 - m - 2d + 1)]^T,$$

when the adaptive controller (8), (9) and (11) is applied to the plant (1) in the presence of a zero disturbance w , the following bound holds:

$$\sum_{k=t_0+2d-1}^\infty \varepsilon(k)^2 \leq c(\|x_0\|^2 + \sup_{j \geq t_0+d} |y^*(j)|^2).$$

Proof Fix $\delta \in (0, \infty]$ and $\lambda \in (\underline{\lambda}, 1)$. Let $t_0 \in \mathbf{Z}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, $y^* \in \ell_\infty$, $\theta_0 \in \mathcal{S}$, and x_0 be arbitrary. Now suppose that $w = 0$; for this case, by the definition of the function ρ_δ , for $t \geq t_0 + d$ we have

$$\rho_\delta(\phi(t-d), e(t)) = 0 \iff \|\phi(t-d)\| = 0.$$

If we incorporate this into the relation (46) with $\bar{t} = t_0 + 2d - 1$, then for $T > t_0 + 2d - 1$ we obtain

$$\begin{aligned} \sum_{t=t_0+2d-1, \phi(t-d) \neq 0}^T \frac{|\varepsilon(t)|^2}{\|\phi(t-d)\|^2} &\leq d^2 \sum_{t=t_0+d, \phi(t-d) \neq 0}^T \frac{|e(t)|^2}{\|\phi(t-d)\|^2} \\ &\leq d^2 \sum_{t=t_0+d, \phi(t-d) \neq 0}^\infty \frac{|e(t)|^2}{\|\phi(t-d)\|^2} \\ &\leq 8d^2 \|\mathcal{S}\|^2 \text{ (by Proposition 1)}. \end{aligned}$$

Since $\varepsilon(t) = 0$ if $\phi(t - d) = 0$, if we now apply the bound on $\phi(t)$ proven in Theorem 1, we conclude that there exists constants $c_1 > 0$ and $\lambda \in (0, 1)$ so that

$$\begin{aligned} \sum_{t=t_0+2d-1}^{\infty} \varepsilon(t)^2 &\leq 8d^2 \|\mathcal{S}\|^2 \times \sup_{j \geq t_0} \|\phi(j)\|^2 \\ &\leq 16d^2 \|\mathcal{S}\|^2 c_1^2 \left[\|x_0\|^2 + \left(\frac{1}{1-\lambda}\right)^2 \sup_{j \geq t_0+d} |y^*(j)|^2 \right], \end{aligned}$$

which yields the desired result. □

6.3 The slowly time-varying case: no disturbance

Now we turn to the case in which the plant parameter is slowly time-varying (with no jumps in the parameters) and the disturbance $w(t)$ is zero. We should not expect to get exact tracking; we will be able to prove, roughly speaking, that the average tracking error is small on average if the time-variation is small. To proceed, we consider the time-varying plant of (40) without the unmodelled dynamics and with zero noise:

$$y(t + d) = \theta^*(t)^T \phi(t). \tag{47}$$

Theorem 4 *For every $\delta \in (0, \infty]$, there exist constants $\bar{\epsilon} > 0$ and $\gamma > 0$ so that for every $t_0 \in \mathbf{Z}$, $\epsilon \in (0, \bar{\epsilon})$, $\theta^* \in s(\mathcal{S}_{\alpha\beta}, 0, \epsilon)$, $y^* \in \ell_\infty$, $\theta_0 \in \mathcal{S}$, and plant initial condition*

$$x_0 = [y(t_0 - 1) \ \cdots \ y(t_0 - n - d + 1) \ u(t_0 - 1) \ \cdots \ u(t_0 - m - 2d + 1)]^T,$$

when the adaptive controller (8), (9) and (11) is applied to the time-varying plant (47), the following holds:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |\varepsilon(j)|^2 \leq \gamma \epsilon^{2/3} \|y^*\|_\infty^2.$$

Proof Fix $\delta \in (0, \infty]$ and $\lambda_1 \in (\lambda, 1)$, and set $w = 0$. By Theorem 2, there exist constants $\gamma_1 > 0$ and

$$\bar{\epsilon} \in (0, \max\{2\|\mathcal{S}\|, 2^{3/2}d^{1/2}\|\mathcal{S}\|\})$$

so that for every $t_0 \in \mathbf{Z}$, $y^* \in \ell_\infty$, $\theta_0 \in \mathcal{S}$, initial condition x_0 , and $\theta^* \in s(\mathcal{S}_{\alpha\beta}, 0, \bar{\epsilon})$, when the adaptive controller (8), (9) and (11) is applied to the plant (47), the following holds:

$$\|\phi(t)\| \leq \gamma_1 \lambda_1^{t-t_0} \|x_0\| + \frac{\gamma_1}{1-\lambda_1} \sup_{j \in [t_0, t]} |y^*(j+d)|, \quad t \geq t_0. \tag{48}$$

Now, let $t_0 \in \mathbf{Z}$, $y^* \in \ell_\infty$, $\theta_0 \in \mathcal{S}_{ab}$, $x_0, \epsilon \in (0, \bar{\epsilon})$ and $\theta^* \in s(\mathcal{S}_{\alpha\beta}, 0, \epsilon)$ be arbitrary. We first want to find a bound on the tracking error in terms of the prediction

error. Note that the relationship in (45) and (46) holds irrespective of the plant model. Since $w = 0$, it follows that for $t \geq t_0 + d$:

$$\rho_\delta(\phi(t - d), e(t)) = 0 \iff \|\phi(t - d)\| = 0.$$

Now apply (46); by changing the index to facilitate analysis, we obtain

$$\sum_{j=\bar{t}+d-1, \phi(j) \neq 0}^{t-1} \frac{|\varepsilon(j + d)|^2}{\|\phi(j)\|^2} \leq d^2 \sum_{j=\bar{t}, \phi(j) \neq 0}^{t-1} \frac{|e(j + d)|^2}{\|\phi(j)\|^2}, \quad t \geq \bar{t} + d, \quad \bar{t} \geq t_0. \quad (49)$$

Now we incorporate time-variation. Let $t_i \geq t_0$ be arbitrary; then from plant equation (47) we have

$$y(t + d) = \phi(t)^\top \theta^*(t_i) + \underbrace{\phi(t)^\top [\theta^*(t) - \theta^*(t_i)]}_{=:\Delta_i(t)}. \quad (50)$$

Note from (41) that $\|\Delta_i(t)\| \leq \epsilon \|\phi(t)\| (t - t_i)$. Define $\tilde{\theta}_i(t) := \hat{\theta}(t) - \theta^*(t_i)$. Since $w = 0$, by applying Proposition 1 to the plant (50), we obtain¹⁰

$$\begin{aligned} \sum_{j=t_i, \|\phi(j)\| \neq 0}^{t-1} \frac{|e(j + d)|^2}{\|\phi(j)\|^2} &\leq 2\|\tilde{\theta}_i(t_i)\|^2 + \sum_{j=t_i, \|\phi(j)\| \neq 0}^{t-1} \frac{4\|\Delta_i(j)\|^2}{\|\phi(j)\|^2} \\ &\leq 2\|\tilde{\theta}_i(t_i)\|^2 + 4\epsilon^2 \sum_{j=t_i}^{t-1} (j - t_i)^2 \\ &= 2\|\tilde{\theta}_i(t_i)\|^2 + 4\epsilon^2 \sum_{k=0}^{t-t_i-1} k^2 \\ &\leq 8\|S\|^2 + 4\epsilon^2(t - t_i - 1)^3, \quad t \geq t_i + 1, \quad t_i \geq t_0. \end{aligned}$$

Using the above bound in (49) with $\bar{t} = t_i$, we obtain

$$\begin{aligned} \sum_{j=t_i+d-1, \phi(j) \neq 0}^{t-1} \frac{|\varepsilon(j + d)|^2}{\|\phi(j)\|^2} &\leq d^2 \sum_{j=t_i, \phi(j) \neq 0}^{t-1} \frac{|e(j + d)|^2}{\|\phi(j)\|^2} \\ &\leq d^2 \left[8\|S\|^2 + 4\epsilon^2(t - t_i - 1)^3 \right], \quad t \geq t_i + d, \quad t_i \geq t_0. \end{aligned} \quad (51)$$

¹⁰ Observe that $\sum_{k=1}^m k^2 = \frac{m(m+1)(2m+1)}{6} = \frac{m^3}{3} + \frac{m^2}{2} + \frac{m}{6} \leq m^3$.

Since the disturbance is zero here, it follows that $\|\phi(j)\| = 0$ implies that $\varepsilon(j + d) = 0$.¹¹ So from (51), we conclude that

$$\sum_{j=t_i+d-1}^{t-1} |\varepsilon(j + d)|^2 \leq 4d^2 \left[2\|\mathcal{S}\|^2 + \epsilon^2(t - t_i - 1)^3 \right] \times \sup_{j \in [t_i+d-1, t)} \|\phi(j)\|^2, \quad t \geq t_i + d, \quad t_i \geq t_0. \tag{52}$$

As $t_i \geq t_0$ is arbitrary, we can relabel the indexes and obtain

$$\sum_{j=t_i}^{t-1} |\varepsilon(j + d)|^2 \leq 4d^2 \left[2\|\mathcal{S}\|^2 + \epsilon^2(t - t_i - 1 + d - 1)^3 \right] \times \sup_{j \in [t_i, t)} \|\phi(j)\|^2, \tag{53}$$

$$t \geq t_i + 1, \quad t_i \geq t_0 + d - 1.$$

We now analyse the average tracking error. From (53), we obtain

$$\frac{1}{t - t_i} \sum_{j=t_i}^{t-1} |\varepsilon(j + d)|^2 \leq 4d^2 \left[\frac{2\|\mathcal{S}\|^2}{t - t_i} + \epsilon^2 \frac{(t - t_i + d - 2)^3}{t - t_i} \right] \times \sup_{j \in [t_i, t)} \|\phi(j)\|^2, \tag{54}$$

$$t \geq t_i + 1, \quad t_i \geq t_0 + d - 1.$$

Now define

$$\beta_\epsilon := \left(\epsilon \times d\|\mathcal{S}\|^2 \right)^{\frac{2}{3}} \tag{55}$$

and $T_\beta \in \mathbb{N}$ by

$$T_\beta := \left\lceil \frac{2d\|\mathcal{S}\|^2}{\beta_\epsilon} \right\rceil;$$

this means that

$$\frac{2d\|\mathcal{S}\|^2}{T_\beta} \leq \beta_\epsilon.$$

Since $\epsilon \leq 2\|\mathcal{S}\|$ by design, we can easily check that $T_\beta \geq d$, which means that

$$\frac{(T_\beta + d - 2)^3}{T_\beta} \leq 8T_\beta^2.$$

¹¹ Observe that this is true even when $\theta^*(t)$ is time-varying.

Incorporating this and the definition of T_β into (54), by choosing $t = t_i + T_\beta$ we have

$$\begin{aligned} \frac{1}{T_\beta} \sum_{j=t_i}^{t_i+T_\beta-1} |\varepsilon(j+d)|^2 &\leq 4d \left[\frac{2d\|\mathcal{S}\|^2}{T_\beta} + 8\epsilon^2 d T_\beta^2 \right] \times \sup_{j \in [t_i, t_i+T_\beta)} \|\phi(j)\|^2 \\ &\leq 4d \left[\beta_\epsilon + 8\epsilon^2 d T_\beta^2 \right] \times \sup_{j \in [t_i, t_i+T_\beta)} \|\phi(j)\|^2, \quad t_i \geq t_0 + d - 1. \end{aligned} \tag{56}$$

We would like to obtain a bound on $\epsilon^2 T_\beta^2$. But it follows from (55) that

$$\beta_\epsilon^3 = \epsilon^2 d^2 \|\mathcal{S}\|^4,$$

so

$$\epsilon^2 = \beta_\epsilon \left(\frac{\beta_\epsilon}{d\|\mathcal{S}\|^2} \right)^2 = 4\beta_\epsilon \left(\frac{\beta_\epsilon}{2d\|\mathcal{S}\|^2} \right)^2;$$

if we define $x = \frac{\beta_\epsilon}{2d\|\mathcal{S}\|^2}$, then we see that

$$\begin{aligned} \epsilon^2 T_\beta^2 &= 4\beta_\epsilon x^2 \left(\lceil \frac{1}{x} \rceil \right)^2 \\ &\leq 4\beta_\epsilon x^2 \left(\frac{1}{x} + 1 \right)^2 \\ &= 4\beta_\epsilon (x + 1)^2. \end{aligned}$$

But $\epsilon \in (0, 2^{3/2} d^{1/2} \|\mathcal{S}\|)$ by hypothesis, so

$$x = \frac{\beta_\epsilon}{2d\|\mathcal{S}\|^2} = \frac{(\epsilon d \|\mathcal{S}\|^2)^{2/3}}{2d\|\mathcal{S}\|^2} < \frac{(2^{3/2} d^{1/2} \|\mathcal{S}\| \times d \|\mathcal{S}\|^2)^{2/3}}{2d\|\mathcal{S}\|^2} = 1,$$

so

$$\epsilon^2 T_\beta^2 \leq 16\beta_\epsilon.$$

Substituting this into (56) and simplifying yields

$$\frac{1}{T_\beta} \sum_{j=t_i}^{t_i+T_\beta-1} |\varepsilon(j+d)|^2 \leq 516d^2 \beta_\epsilon \times \sup_{j \in [t_i, t_i+T_\beta)} \|\phi(j)\|^2, \quad t_i \geq t_0 + d - 1. \tag{57}$$

We now analyse the average tracking error over the whole time horizon; we do so by considering time intervals of length T_β . From (57) we easily obtain

$$\frac{1}{iT_\beta} \sum_{j=\bar{i}}^{\bar{i}+iT_\beta-1} |\varepsilon(j+d)|^2 \leq 516d^2\beta_\epsilon \times \sup_{j \in [\bar{i}, \bar{i}+iT_\beta)} \|\phi(j)\|^2, \quad i \in \mathbf{N}, \bar{i} \geq t_0 + d - 1. \tag{58}$$

The bound in (58) provides a bound on the average tracking error over time intervals of lengths that are multiples of T_β . To extend this to intervals of arbitrary length, first observe that (58) can be rewritten as

$$\sum_{j=\bar{i}}^{\bar{i}+iT_\beta-1} |\varepsilon(j+d)|^2 \leq iT_\beta \left(\underbrace{516d^2}_{=: \gamma_2} \beta_\epsilon \right) \sup_{j \in [\bar{i}, \bar{i}+iT_\beta)} \|\phi(j)\|^2, \quad i \in \mathbf{N}, \bar{i} \geq t_0 + d - 1.$$

For $k \in \{0, 1, \dots, T_\beta - 1\}$, this inequality implies that

$$\sum_{j=\bar{i}+k}^{\bar{i}+iT_\beta-1+k} |\varepsilon(j+d)|^2 \leq i\gamma_2 T_\beta \beta_\epsilon \sup_{j \in [\bar{i}+k, \bar{i}+iT_\beta+k)} \|\phi(j)\|^2, \quad i \in \mathbf{N}, \bar{i} \geq t_0 + d - 1;$$

adding these two inequalities and simplifying yields

$$\begin{aligned} \sum_{j=\bar{i}}^{\bar{i}+iT_\beta-1+k} |\varepsilon(j+d)|^2 &\leq 2i\gamma_2 T_\beta \beta_\epsilon \sup_{j \in [\bar{i}, \bar{i}+iT_\beta+k)} \|\phi(j)\|^2, \quad i \in \mathbf{N}, \\ k &\in \{0, 1, \dots, T_\beta - 1\}, \bar{i} \geq t_0 + d - 1. \end{aligned}$$

Changing variables to enhance clarity, we see that this implies that

$$\sum_{j=\bar{i}}^{\bar{i}+T-1} |\varepsilon(j+d)|^2 \leq 2\gamma_2 T \beta_\epsilon \sup_{j \in [\bar{i}, \bar{i}+T-1)} \|\phi(j)\|^2, \quad T \geq T_\beta, \bar{i} \geq t_0 + d - 1,$$

which means that

$$\frac{1}{T} \sum_{j=\bar{i}}^{\bar{i}+T-1} |\varepsilon(j+d)|^2 \leq 2\gamma_2 \beta_\epsilon \sup_{j \in [\bar{i}, \bar{i}+T-1)} \|\phi(j)\|^2, \quad T \geq T_\beta, \bar{i} \geq t_0 + d - 1. \tag{59}$$

This means that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=\bar{i}}^{\bar{i}+T-1} |\varepsilon(j+d)|^2 \leq 2\gamma_2 \beta_\epsilon \times \limsup_{j \rightarrow \infty} \|\phi(j)\|^2, \quad \bar{i} \geq t_0 + d - 1.$$

From (48), see that

$$\limsup_{j \rightarrow \infty} \|\phi(j)\| \leq \frac{\gamma_1}{1 - \lambda_1} \|y^*\|_\infty;$$

the boundedness of the tracking error ensures that for every $\bar{i} \geq t_0$ we have

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=\bar{i}}^{\bar{i}+T-1} |\varepsilon(j+d)|^2 = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |\varepsilon(j+d)|^2, \quad \bar{i} \geq t_0 + d - 1,$$

so we conclude that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |\varepsilon(j+d)|^2 \leq \left(\frac{\gamma_1}{1 - \lambda_1}\right)^2 2\gamma_2\beta_\epsilon \times \|y^*\|_\infty^2.$$

Since $\beta_\epsilon = (d\|S\|^2)^{2/3}\epsilon^{2/3}$, the result follows. □

6.4 Tracking in the presence of a disturbance

Now we turn to the much harder problem of tracking in the presence of a disturbance; throughout this sub-section, we assume that the plant is LTI. Our goal is to show that if the noise is small, then the tracking error is small; this is a stringent requirement, since in adaptive control it is usually only proven that if the noise is bounded, then the error is bounded. We can, of course, measure signal sizes in a variety of ways, with the 2-norm and the ∞ -norm the most common; given that a large disturbance can lead the estimator astray and cause “temporary instability”, the 2-norm seems to be the most appropriate here.

If the closed-loop system were LTI, then by Parseval’s Theorem we could conclude that the average power of the tracking error is bounded by the average power of the disturbance, i.e. there exists a constant c so that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [\varepsilon(t)]^2 \leq c \times \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [w(t)]^2; \tag{60}$$

unfortunately, while the closed-loop system has some desirable linear-like closed-loop properties, the controller is nonlinear so the closed-loop system is clearly not LTI. However, we can prove this bound in two extreme cases:

- if $y^* = 0$, then $y = \varepsilon$, so with c and λ as given in Theorem 1, it is easy to see that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [\varepsilon(t)]^2 \leq \frac{c^2}{(1 - \lambda)^2} \times \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [w(t)]^2;$$

- on the other hand, if $y^* \neq 0$ and $w = 0$, then from Theorem 3 we see that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} [\varepsilon(j)]^2 = 0.$$

In the general case, we will prove something weaker than (60), but with much the same flavour; it is, however, much stronger than the standard result in the literature.

Theorem 5 For every $\delta \in (0, \infty]$, there exists a $\gamma > 0$ so that for every $t_0 \in \mathbf{Z}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, y^* , $w \in \ell_\infty$, $\theta_0 \in \mathcal{S}$, and plant initial condition

$$x_0 = [y(t_0 - 1) \ \cdots \ y(t_0 - n - d + 1) \ u(t_0 - 1) \ \cdots \ u(t_0 - m - 2d + 1)]^T,$$

when the adaptive controller (8), (9) and (11) is applied to the plant (1) with $\liminf_{t \rightarrow \infty} |y^*(t)| > 0$ then the following holds:

$$\begin{aligned} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |\varepsilon(j)|^2 &\leq \gamma \times \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=t_0}^{t_0+T-1} |w(j)|^2 \\ &\times \frac{\limsup_{t \rightarrow \infty} |y^*(t)|^2 + \limsup_{t \rightarrow \infty} |w(t)|^2}{\liminf_{t \rightarrow \infty} |y^*(t)|^2}. \end{aligned} \tag{61}$$

Remark 9 So we see that the bound proven here is similar to that of (60) which holds in the LTI case, although we have an extra term multiplied on the RHS:

$$\frac{\limsup_{t \rightarrow \infty} |y^*(t)|^2 + \limsup_{t \rightarrow \infty} |w(t)|^2}{\liminf_{t \rightarrow \infty} |y^*(t)|^2}.$$

If the reference signal is larger than the noise, which is what one would normally expect, then this would be bounded by

$$2 \frac{\limsup_{t \rightarrow \infty} |y^*(t)|^2}{\liminf_{t \rightarrow \infty} |y^*(t)|^2};$$

if $|y^*(t)| \in \{-1, 1\}$ then this is exactly two. It is curious that the quantity gets large if $y^*(t)$ gets close to zero; we suspect that this is an artifact of the proof, since all simulations indicate that the LTI-like bound (60) holds.

Proof Fix $\delta \in (0, \infty]$ and $\lambda \in (\underline{\lambda}, 1)$. Let $t_0 \in \mathbf{Z}$, $\theta_{ab}^* \in \mathcal{S}_{ab}$, y^* , $w \in \ell_\infty$, $\theta_0 \in \mathcal{S}$, and x_0 be arbitrary, but so that $\liminf_{t \rightarrow \infty} |y^*(t)| > 0$. Before proceeding, choose $t \geq t_0 + 2d - 1$ so that

$$\inf\{|y^*(t)| : t \geq t\} > 0.$$

Now, by using (46) and applying Proposition 1, for $\bar{t} \geq t_0 + 2d - 1$ we obtain

$$\begin{aligned}
 & \sum_{j=\bar{t}}^t \rho_\delta(\phi(j-d), e(j)) \frac{|\varepsilon(j)|^2}{\|\phi(j-d)\|^2} \\
 & \leq d^2 \sum_{j=\bar{t}-d+1}^t \rho_\delta(\phi(j-d), e(j)) \frac{|e(j)|^2}{\|\phi(j-d)\|^2} \\
 & \leq 8d^2 \|\mathcal{S}\|^2 + 4d^2 \sum_{j=\bar{t}-d+1}^t \rho_\delta(\cdot, \cdot) \frac{|\bar{w}(j-d)|^2}{\|\phi(j-d)\|^2} \\
 & = 4d^2 \|\mathcal{S}\|^2 \left(2 + \frac{1}{\|\mathcal{S}\|^2} \sum_{j=\bar{t}-d+1}^t \rho_\delta(\cdot, \cdot) \frac{|\bar{w}(j-d)|^2}{\|\phi(j-d)\|^2} \right) \\
 & = 4d^2 \|\mathcal{S}\|^2 \left(2 + \sum_{j=\bar{t}-d+1}^t \rho_\delta(\cdot, \cdot) \frac{|\bar{w}(j-d)|^2}{(\|\mathcal{S}\| \|\phi(j-d)\|)^2} \right), \\
 & \quad t \geq \bar{t} \geq t_0 + 2d - 1.
 \end{aligned} \tag{62}$$

From the controller equation (11), we have

$$y^*(t) = \hat{\theta}(t-d)^\top \phi(t-d), \quad t \geq t_0 + d,$$

which means that

$$|y^*(t)| \leq \|\hat{\theta}(t-d)\| \|\phi(t-d)\| \leq \|\mathcal{S}\| \times \|\phi(t-d)\|, \quad t \geq t_0 + d;$$

if we substitute this into (62), then we obtain

$$\begin{aligned}
 & \sum_{j=\bar{t}}^t \rho_\delta(\phi(j-d), e(j)) \frac{|\varepsilon(j)|^2}{\|\phi(j-d)\|^2} \\
 & \leq 4d^2 \|\mathcal{S}\|^2 \left(2 + \sum_{j=\bar{t}-d+1}^t \rho_\delta(\phi(j-d), e(j)) \frac{|\bar{w}(j-d)|^2}{|y^*(j)|^2} \right), \\
 & \quad t \geq \bar{t} \geq \underline{t}.
 \end{aligned} \tag{63}$$

Now we analyse the above bound for two cases: when the estimator is turned on, i.e. when $\rho_\delta(\cdot, \cdot) = 1$ and when the estimator is turned off, i.e. when $\rho_\delta(\cdot, \cdot) = 0$. Before proceeding, we define some notation: for $t_2 \geq t_1 \geq t_0$, we define

$$\underline{\mathbf{y}}^*_{[t_1, t_2]} := \inf_{j \in [t_1, t_2], \rho_\delta(\phi(j-d), e(j))=1} |y^*(j)|^2.$$

Case 1 The estimator is turned on: $\rho_\delta(\phi(j-d), e(j)) = 1$.

From (63), we have

$$\sum_{j=\bar{t}, \rho_\delta(\phi(j-d), e(j))=1}^t \frac{|\varepsilon(j)|^2}{\|\phi(j-d)\|^2} \leq 4d^2 \|\mathcal{S}\|^2 \left(2 + \frac{1}{\underline{y}_{[\bar{t}-d+1, t]}^*} \sum_{j=\bar{t}-d+1}^t |\bar{w}(j-d)|^2 \right), \quad t \geq \bar{t} \geq \underline{t}, \tag{64}$$

which means that

$$\sum_{j=\bar{t}, \rho_\delta(\phi(j-d), e(j))=1}^t |\varepsilon(j)|^2 \leq 4d^2 \|\mathcal{S}\|^2 \left(\sup_{j \in [\bar{t}-d, t-d]} \|\phi(j)\|^2 \right) \times \left(2 + \frac{1}{\underline{y}_{[\bar{t}-d+1, t]}^*} \sum_{j=\bar{t}-d+1}^t |w(j-d)|^2 \right), \quad t \geq \bar{t} \geq \underline{t}. \tag{65}$$

Case 2 The estimator is turned off : $\rho_\delta(\phi(j-d), e(j)) = 0$.

In this case, we know from the definition of ρ_δ that when $\rho_\delta(\phi(t-d), e(t)) = 0$:

- if $\delta = \infty$ then $\phi(t-d) = 0$, so

$$\|\phi(t-d)\| \leq \frac{1}{\delta} |\bar{w}(t-d)|;$$

- if $\delta \in (0, \infty)$, then we have that

$$|e(t)| \geq (2\|\mathcal{S}\| + \delta) \|\phi(t-d)\|;$$

using the formula for the prediction error given in (13) we see that

$$\begin{aligned} |e(t)| &\leq \|\phi(t-d)\| \times \|\tilde{\theta}(t-1)\| + |\bar{w}(t-d)| \\ &\leq 2\|\mathcal{S}\| \times \|\phi(t-d)\| + |\bar{w}(t-d)|; \end{aligned}$$

combining these two equations yields

$$\|\phi(t-d)\| \leq \frac{1}{\delta} |\bar{w}(t-d)|.$$

Using the formula for the tracking error given in (12), we have

$$\begin{aligned} |\varepsilon(t)| &\leq \|\phi(t-d)\| \|\tilde{\theta}(t-d)\| + |\bar{w}(t-d)| \\ &\leq \frac{2\|\mathcal{S}\|}{\delta} |\bar{w}(t-d)| + |\bar{w}(t-d)|, \quad t \geq t_0 + 2d - 1. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{j=\bar{t}, \rho_\delta(\phi(j-d), e(j))=0}^t |\varepsilon(j)|^2 &\leq \left(1 + \frac{2\|\mathcal{S}\|}{\delta}\right)^2 \sum_{j=\bar{t}, \rho_\delta(\phi(j-d), e(j))=0}^t |\bar{w}(j-d)|^2 \\ &\leq \left(1 + \frac{2\|\mathcal{S}\|}{\delta}\right)^2 \sum_{j=\bar{t}}^t |\bar{w}(j-d)|^2, \quad t \geq \bar{t} \geq t_0 + 2d - 1. \end{aligned} \tag{66}$$

We can now combine (65) and (66) of Case 1 and Case 2, respectively, to yield

$$\begin{aligned} \sum_{j=\bar{t}}^t |\varepsilon(j)|^2 &\leq 8d^2\|\mathcal{S}\|^2 \left(\sup_{j \in [\bar{t}, t]} \|\phi(j+d)\|^2\right) \\ &\quad + \max \left\{ \left(1 + \frac{2\|\mathcal{S}\|}{\delta}\right)^2, 4d^2\|\mathcal{S}\|^2 \left(\frac{\sup_{j \in [\bar{t}, t]} \|\phi(j+d)\|^2}{\underline{y}_{[\bar{t}-d+1, t]}^*}\right) \right\} \\ &\quad \times \left(\sum_{j=\bar{t}-d+1}^t |\bar{w}(j-d)|^2\right), \quad t \geq \bar{t} \geq \underline{t}. \end{aligned} \tag{67}$$

By Theorem 1, there exists constants $c > 0$ and $\lambda \in (0, 1)$ so that

$$\|\phi(t+d)\| \leq c\lambda^{t+d-t_0}\|x_0\| + \sum_{j=t_0}^{t+d} c\lambda^{t+d-j}(|y^*(j)| + |w(j)|), \quad t \geq \underline{t},$$

so we can choose $\bar{t} \geq \underline{t}$ (which depends implicitly on $x_0, y^*, \theta_0,$ and θ^*) such that

$$\|\phi(t+d)\| \leq \frac{2c}{1-\lambda} \limsup_{k \rightarrow \infty} (|y^*(\bar{t}+k)| + |w(\bar{t}+k)|), \quad t \geq \bar{t},$$

as well as

$$|y^*(t)|^2 \geq \underbrace{\frac{1}{2} \liminf_{k \rightarrow \infty} |y^*(k)|^2}_{=: \underline{y}^*}, \quad t \geq \bar{t} - d + 1.$$

If we incorporate this into (67), then we obtain

$$\begin{aligned} \sum_{j=\bar{t}}^t |\varepsilon(j)|^2 &\leq 8d^2\|\mathcal{S}\|^2 \left(\sup_{j \in [\bar{t}, t]} \|\phi(j+d)\|^2\right) \\ &\quad + \max \left\{ \left(\frac{2\|\mathcal{S}\|}{\delta} + 1\right)^2, \left(\frac{4cd\|\mathcal{S}\|}{1-\lambda}\right)^2 \times \frac{\limsup_{k \rightarrow \infty} (|y^*(\bar{t}+k)|^2 + |\bar{w}(\bar{t}+k)|^2)}{0.5\underline{y}^*} \right\} \end{aligned}$$

$$\times \left(\sum_{j=\bar{t}-d+1}^{\bar{t}} |\bar{w}(j-d)|^2 \right), \quad t \geq \bar{t} - d + 1,$$

which means that

$$\begin{aligned} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=\bar{t}}^{\bar{t}+T-1} |\varepsilon(j)|^2 &\leq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{j=\bar{t}-d+1}^{\bar{t}+T-1} |\bar{w}(j-d)|^2 \\ &\times \max \left\{ \left(\frac{2\|S\|}{\delta} + 1 \right)^2, \left(\frac{4cd\|S\|}{1-\lambda} \right)^2 \times \frac{\limsup_{k \rightarrow \infty} (|y^*(\bar{t}+k)|^2 + |w(\bar{t}+k)|^2)}{0.5\underline{y}^*} \right\}. \end{aligned}$$

But $\bar{w}(t)$ is a weighted sum of $\{w(t+1), \dots, w(t+d)\}$, and the boundedness of all variables makes the starting point of the average sums irrelevant, so the desired bound (61) follows. □

7 A simulation example

Here we provide several simulation examples to illustrate the results of this paper. Consider the second-order time-varying plant with relative degree one:

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

with $a_1(t) \in [-3, 3]$, $a_2(t) \in [-4, 4]$, $b_0(t) \in [1.5, 5]$ and $b_1(t) \in [-1, 1]$. When the parameters are fixed, we see that the corresponding system has a zero in $[-\frac{2}{3}, \frac{2}{3}]$ and poles which may be stable or unstable.

7.1 Simulation 1: the benefit of vigilant estimation

In the first simulation, we illustrate that the vigilant estimation algorithm provides better performance than the classical algorithm does. More specifically, we compare the adaptive controller using the vigilant estimator (8), (9) (with $\delta = \infty$) with the adaptive controller using the classical estimation algorithm (7) suitably modified to incorporate projection onto \mathcal{S} (and with $\bar{\alpha} = \bar{\beta} = 1$). In each case, we apply the adaptive controller to this plant with the parameters chosen as

$$\begin{aligned} a_1(t) &= 3 \cos(0.01t), \\ a_2(t) &= 4 \sin(0.007t), \\ b_0(t) &= 3.25 - 1.75 \cos(0.0045t), \\ b_1(t) &= -\cos(0.002t). \end{aligned}$$

We set $y^* = 0$ (so $\varepsilon(t) = y(t)$) and the noise to

$$w(t) = \begin{cases} 0 & 0 \leq t \leq 100 \\ 0.05 \cos(10t) & \text{otherwise;} \end{cases}$$

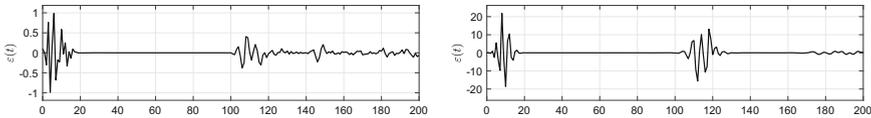


Fig. 1 The left plot shows the tracking error for Simulation 1 for $t \leq 200$ using vigilant estimation; the right plot shows the tracking error for Simulation 1 for $t \leq 200$ using classical estimation

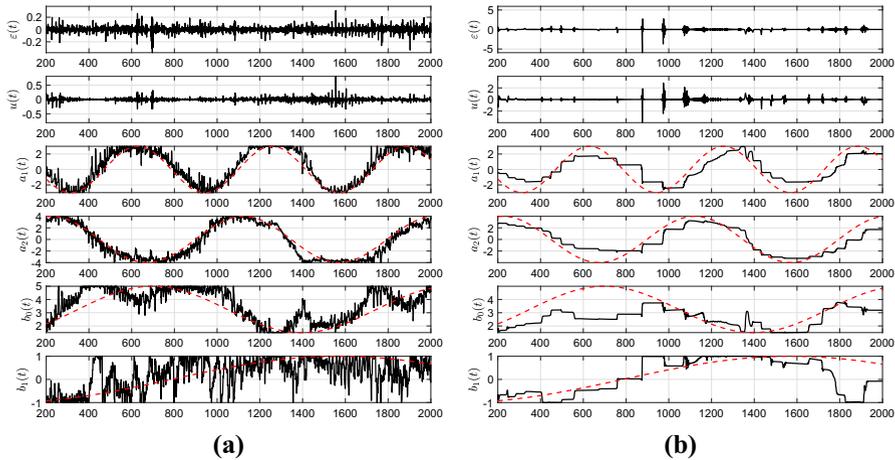


Fig. 2 Column **a** shows results using vigilant estimation, and column **b** shows results using classical estimation. In each column, the upper plot shows the tracking error for $t \geq 200$; the next plot shows the control signal for $t \geq 200$; the last four plots show the parameter estimates (solid) and actual parameters (dashed) for $t \geq 200$

we set $y(-1) = y(0) = -0.1, u(-1) = u(-2) = 0$, and the initial parameter estimates to the midpoint of the respective intervals. Fig. 1 shows tracking errors during the transient phase of 200 steps, and Fig. 2 shows the tracking error and parameter estimation for the rest of the simulation time of 2000 steps. We can clearly see that the proposed controller provides better transient performance as well as better disturbance rejection than the classical algorithm. Furthermore, you can see that the proposed controller does a much better job of tracking the parameter variations.

7.2 Simulation 2: illustrating robustness

In the second simulation, we illustrate the tolerance to time-variation and unmodelled dynamics. We apply the proposed adaptive controller (with $\delta = \infty$) to the plant when the parameters are time-varying:

$$\begin{aligned}
 a_1(t) &= -3 \cos(0.002t), \\
 a_2(t) &= -4 \sin(0.0015t), \\
 b_0(t) &= 3.25 - 1.75 \cos(0.0035t),
 \end{aligned}$$

$$b_1(t) = \begin{cases} 0.5 & 0 \leq t \leq 1500 \\ -0.5 & 1500 < t \leq 3000 \\ 1 & 3000 < t, \end{cases}$$

and when the unmodelled dynamics (in the associated plant model (40)) are described by

$$\begin{aligned} m(t + 1) &= 0.5m(t) + 0.5\|\phi(t)\|, & m(0) &= 0 \\ \bar{w}_\Delta(t) &= \begin{cases} 0.1m(t) + 0.1\|\phi(t)\| & t \geq 2500 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We set the reference signal to

$$y^*(t) = \cos(0.015t)$$

and the noise signal to

$$w(t) = 0.01 \cos(10t);$$

we choose initial conditions of $y(-1) = y(0) = -1$ and $u(-1) = u(-2) = 0$, and the initial parameter estimates to $\theta_0 = [-2 \ -2 \ 2 \ 1]^\top$; Fig. 3 shows the results. The adaptive controller clearly shows robust performance to both unmodelled dynamics and time-variations, including parameter jumps: the tracking is quite good, and the parameter estimator (approximately) tracks the time-varying parameters.

7.3 Simulation 3: tracking with no noise

Theorem 3 says that there exists a constant c so that

$$\sum_{j=t_0+2d-1}^{\infty} \varepsilon(j)^2 \leq c(\|y^*\|_\infty^2 + \|x_0\|^2),$$

i.e. the 2-norm of the tracking error is not only finite, but is bounded by a constant time the sizes of the reference signal and plant initial condition. In this simulation, we apply the proposed adaptive controller (with $\delta = \infty$) to the plant when

$$a_1(t) = -2, \quad a_2(t) = 3, \quad b_0(t) = 3.25, \quad b_1(t) = -1.$$

We set the disturbance signal to zero and the reference signal to

$$y^*(t) = A_0 \cos(0.025t)$$

with amplitudes of

$$A_0 \in \{10^{-3}, 10^{-2}, 0.1, 1, 10, 10^2, 10^3, 10^4\};$$

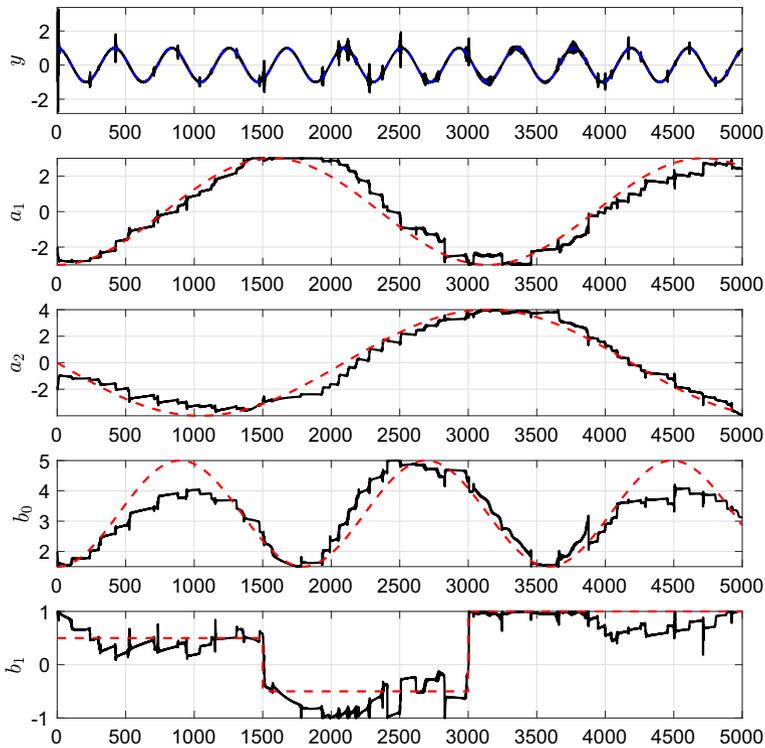


Fig. 3 The upper plot shows both the output signal (solid) and the reference signal (dashed) for Simulation 2; the next four plots show the parameter estimates (solid) and actual parameters (dashed) for Simulation 2

we set the plant initial condition to $y(-1) = y(0) = -1$ and $u(-1) = u(-2) = 0$, and we set the initial parameter estimate to the midpoint of the respective intervals. We simulate for 100,000 steps. Because there is no noise, we get asymptotic tracking; we compute the square sum of the tracking error and compare it to the square of the reference amplitude and the square of the norm of the plant initial condition; the results are in Table 1. We see that the result is consistent with Theorem 3: the gain is bounded above by approximately 1000.

7.4 Simulation 4: tracking with slowly varying parameters

In our fourth simulation, we illustrate the result in Theorem 4; we show that the average tracking error is proportional to the speed of the parameter variation. We apply the proposed controller (with $\delta = \infty$) with plant parameters of

$$a_2(t) = 1, \quad b_0(t) = 3.25, \quad b_1(t) = 1, \\ a_1(t) = 2 \sin(\omega_0 t),$$

Table 1 Simulation 3 results

A_0	$\left(\sum_{j=1}^{100000} \varepsilon(j)^2 \right) / (A_0^2 + \ x_0\ ^2)$
10^{-3}	21.1
10^{-2}	29.4
10^{-1}	31.2
1	186.8
10	968.9
10^2	783.5
10^3	788.5
10^4	789.2

with

$$\omega_0 \in \{0.001, 0.002, 0.005, 0.01\}.$$

We choose plant initial conditions of $y(-1) = y(0) = -1$ and $u(-1) = u(-2) = 0$, and the initial parameter estimate equal to the midpoint of the respective intervals. With a zero disturbance and

$$y^*(t) = \cos(0.015t),$$

we simulate the closed-loop system for $T = 5000$ steps; we plot the tracking error for the last 3000 steps in Fig. 4, i.e. after the transient effect has been eliminated. We see that, consistent with Theorem 4, average tracking error increases with the speed of the plant parameter variation.

7.5 Simulation 5: tracking with noise

In the final simulation, we illustrate the tracking result in the presence of noise, namely Theorem 5. We show that, on average, the tracking error is proportional to the size of the exogenous noise. We apply the proposed adaptive controller (with $\delta = \infty$) to the plant when

$$a_1 = -2, \quad a_2 = 3, \quad b_0 = 3.25, \quad b_1 = -1.$$

We choose an initial condition of $y(-1) = y(0) = -1$ and $u(-1) = u(-2) = 0$ and set the initial parameter estimates to the midpoint of the respective intervals. We set the noise to

$$w(t) = W_0 \cos(10t)$$

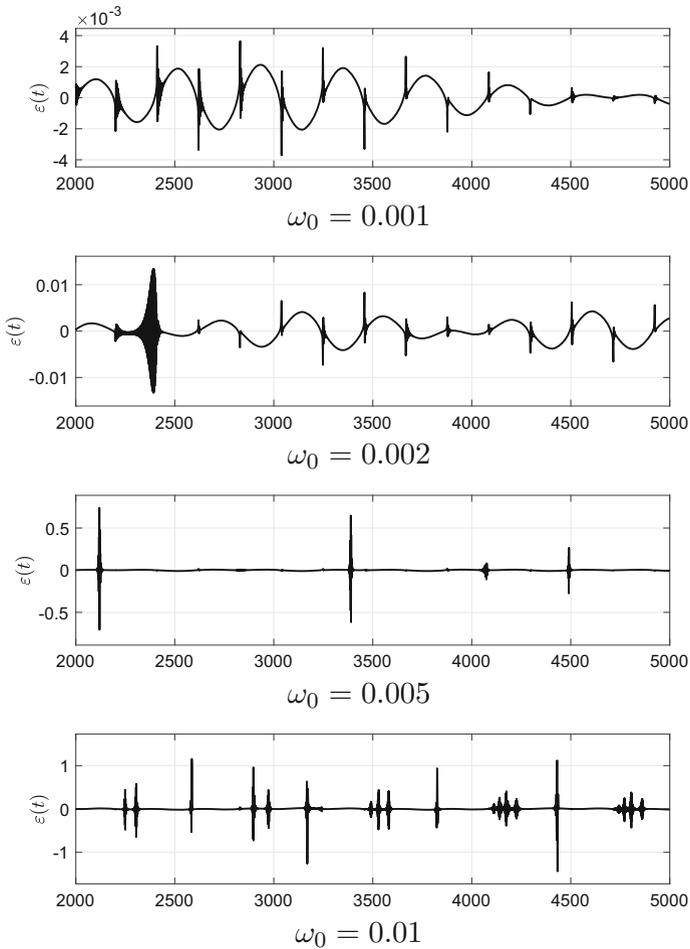


Fig. 4 The plots show the tracking error for Simulation 4

with different amplitudes:

$$W_0 \in \{0.001, 0.01, 0.1, 1\};$$

we choose the reference signal y^* to be a square wave of amplitude one and period 300; observe that $\liminf_{t \rightarrow \infty} |y^*(t)| = 1$, $\|y^*\|_\infty = 1$ and $\|w\|_\infty = |W_0|$. We simulate for $T = 5000$ steps; in Fig. 5, we plot the tracking error for the last 3000 steps, so that we focus on the “steady-state behaviour”. We clearly see that the average tracking error magnitude is roughly proportional to the average noise signal magnitude.

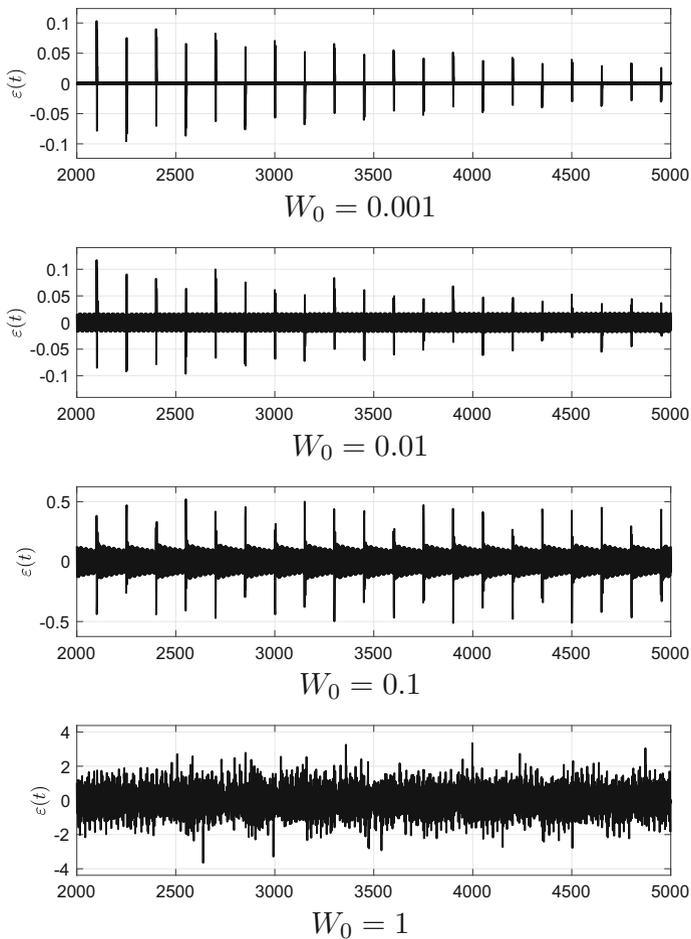


Fig. 5 The “steady-state” tracking error for Simulation 5 for different noise magnitudes

8 Summary and conclusions

Under common assumptions on the plant model, in this paper we use a modified version of the original, ideal, projection algorithm (termed a **vigilant estimator**) to carry out parameter estimation; the corresponding d -step-ahead adaptive controller **guarantees linear-like convolution bounds on the closed-loop behaviour**, which confers exponential stability and a bounded noise gain, unlike almost all other parameter adaptive controllers. This is then leveraged in a modular way to prove tolerance to unmodelled dynamics and plant parameter variation. We examine the tracking ability of the approach and are able to prove properties which most adaptive controllers do not enjoy:

- (i) in the absence of a disturbance, we obtain an explicit 2-norm on the size of the tracking error in terms of the size of the initial condition and the reference signal;

- (ii) if there is no noise but there are slow time-variations, then we prove that we can bound the size of the average tracking error by the size of the time-variation;
- (iii) if there is noise, then under some technical conditions we bound the size of the average tracking error in terms of the size of the average disturbance times another complicated quantity.

We are working on several extensions of the approach. First of all, we would like to use a multi-estimator approach to reduce the amount of structural information on the plant; we have proven this already in the case of first-order systems [35], but extending this to the general case has proven to be challenging. Second of all, we would like to obtain a crisper bound on the average tracking error than that discussed in (iii) above, i.e. prove that the size of the average tracking error is bounded above by a constant time the size of the average disturbance; while all of our simulations confirm this, as of yet we have not been able to prove it.

9 Appendix

Proof of Proposition 3 To proceed, we analyse the good closed-loop model of Sect. 3.1. From (15), it is clear that we need to obtain a bound on the terms $B_1\varepsilon(t + 1)$ and $B_2\varepsilon(t + j)$ for $j = 1, \dots, d + 1$. It will be convenient to define an intermediate quantity¹²:

$$\bar{v}(t - 1) := \rho_\delta(\phi(t - d), e(t)) \times \frac{\phi(t - d)}{\|\phi(t - d)\|^2} \varepsilon(t), \quad t \geq t_0.$$

Step 1 Obtain a desirable bound on $B_i\varepsilon(t)$ in terms of $\bar{v}(t - 1)$, $\phi(t - d)$ and $\bar{w}(t - d)$.

First of all, for $i = 1, 2$ define

$$\bar{\Delta}_i(t) := \rho_\delta(\phi(t - d), e(t)) \frac{\varepsilon(t)}{\|\phi(t - d)\|^2} B_i \phi(t - d)^T, \quad t \geq t_0.$$

It is easy to see that

$$\bar{\Delta}_i(t) \phi(t - d) = \rho_\delta(\phi(t - d), e(t)) B_i \varepsilon(t).$$

So

$$\begin{aligned} B_i \varepsilon(t) &= \rho_\delta(\phi(t - d), e(t)) B_i \varepsilon(t) + \underbrace{[1 - \rho_\delta(\phi(t - d), e(t))] \varepsilon(t)}_{=: \eta_0(t)} B_i \\ &= \bar{\Delta}_i(t) \phi(t - d) + B_i \eta_0(t), \quad t \geq t_0. \end{aligned} \tag{68}$$

Using (12) and (13) and the definition of $\rho_\delta(\phi(t - d), e(t))$, it is easy to show that

$$|\eta_0(t)| \leq \left(1 + \frac{4\|\mathcal{S}\|}{\delta}\right) |\bar{w}(t - d)|, \quad t \geq t_0 + d; \tag{69}$$

¹² It is similar to $v(t - 1)$ except for the $\varepsilon(t)$ rather than $e(t)$ at the end.

furthermore, it is clear that

$$\|\bar{\Delta}_i(t)\| = \rho_\delta(\phi(t-d), e(t)) \frac{|e(t)|}{\|\phi(t-d)\|} = \|\bar{v}(t-1)\|, \quad t \geq t_0. \quad (70)$$

Step 2 Bound $\bar{v}(t-1)$ in terms of $v(t), \dots, v(t-d)$.

It follows from the formulas for $\varepsilon(t)$ and $e(t)$ given in (12) and (13) that

$$\varepsilon(t) = e(t) + \phi(t-d)^T [\hat{\theta}(t-1) - \hat{\theta}(t-d)], \quad t \geq t_0 + d.$$

Using the definition of $\bar{v}(t-1)$, we have

$$\begin{aligned} \|\bar{v}(t-1)\| &\leq \rho_\delta(\phi(t-d), e(t)) \frac{|e(t)|}{\|\phi(t-d)\|} + \|\hat{\theta}(t-1) - \hat{\theta}(t-d)\| \\ &= \|v(t-1)\| + \|\hat{\theta}(t-1) - \hat{\theta}(t-d)\|, \quad t \geq t_0 + d. \end{aligned}$$

Now it follows from the estimator update law that

$$\begin{aligned} \|\hat{\theta}(t-1) - \hat{\theta}(t-d)\| &\leq \|\hat{\theta}(t-1) - \hat{\theta}(t-2)\| + \dots + \|\hat{\theta}(t-d+1) - \hat{\theta}(t-d)\| \\ &\leq \rho_\delta(\phi(t-d-1), e(t-1)) \frac{|e(t-1)|}{\|\phi(t-d-1)\|} + \dots \\ &\quad + \rho_\delta(\phi(t-2d+1), e(t-d+1)) \frac{|e(t-d+1)|}{\|\phi(t-2d+1)\|} \\ &\leq \sum_{j=2}^d \|v(t-j)\|. \end{aligned}$$

We conclude that

$$\|\bar{v}(t-1)\| \leq \sum_{j=1}^d \|v(t-j)\|, \quad t \geq t_0 + d. \quad (71)$$

Step 3 Obtain a bound on $B_i \varepsilon(t)$ in terms of $\{v(t), \dots, v(t-d)\}, \phi(t-d)$ and $\bar{w}(t-d)$.

If we combine (68), (69), (70) and (71), we see that

$$B_i \varepsilon(t) = \bar{\Delta}_i(t) \phi(t-d) + B_i \eta_0(t)$$

with

$$\|\bar{\Delta}_i(t)\| \leq \sum_{j=1}^d \|v(t-j)\|$$

and

$$|\eta_0(t)| \leq \left(1 + \frac{4\|S\|}{\delta}\right) |\bar{w}(t - d)|, \quad t \geq t_0 + d.$$

Step 4 Apply the result of Step 3 to (15).

We can apply the above result to each of the terms on the RHS of (15) containing $\varepsilon(\cdot)$. So from Step 3, we have

$$B_1\varepsilon(t + 1) = \bar{\Delta}_1(t + 1)\phi(t + 1 - d) + B_1\eta_0(t + 1) \tag{72}$$

and

$$\begin{aligned} \frac{a_{d-j}}{b_0} B_2\varepsilon(t + 1 + j) &= \frac{a_{d-j}}{b_0} \bar{\Delta}_2(t + 1 + j)\phi(t + 1 - d + j) \\ &\quad + \frac{a_{d-j}}{b_0} B_2\eta_0(t + 1 + j), \quad j = 0, 1, \dots, d. \end{aligned} \tag{73}$$

Each term except one is of the desired form: the case of $j = d$ is problematic since it contains a $\phi(t + 1)$ term. However, we can use the crude model given in (16) to see that

$$\begin{aligned} &\frac{a_0}{b_0} \bar{\Delta}_2(t + d + 1)\phi(t + 1) \\ &= \frac{a_0}{b_0} \bar{\Delta}_2(t + d + 1) \times [A_b(t)\phi(t) + B_3(t)y^*(t + d + 1) + B_4(t)w(t + 1)]. \end{aligned} \tag{74}$$

If we now combine (72), (73) and (74), we see that we should define

$$\begin{aligned} \Delta_0(t) &= \frac{a_0}{b_0} \bar{\Delta}_2(t + d + 1)A_b(t) + \frac{a_1}{b_0} \bar{\Delta}_2(t + d), \\ \Delta_j(t) &= \frac{a_{j+1}}{b_0} \bar{\Delta}_2(t + d - j), \quad j = 1, \dots, d - 2, \end{aligned}$$

and

$$\Delta_{d-1}(t) = \bar{\Delta}_1(t + 1) + \frac{a_d}{b_0} \bar{\Delta}_2(t + 1).$$

It is clear from Step 3 that this choice of Δ_i has the desired property for $t \geq t_0 + d - 1$. Last of all, we group all of the remaining terms into $\eta(t)$:

$$\begin{aligned} \eta(t) &:= B_1y^*(t + 1) + B_2\frac{1}{b_0} \sum_{j=0}^d a_{d-j}y^*(t + 1 + j) - \frac{1}{b_0} B_2w(t + d + 1) \\ &\quad + B_1\eta_0(t + 1) + \frac{a_0}{b_0} \bar{\Delta}_2(t + d + 1)[B_3(t)y^*(t + d + 1) \end{aligned}$$

$$+ B_4(t)w(t+1)] + B_2 \sum_{j=0}^{d-1} \frac{a_{d-j}}{b_0} \eta_0(t+1+j).$$

If we apply Proposition 2 and use the bound on $\eta_0(t)$ given in Step 3, then we see that $\eta(t)$ has the desired property for $t \geq t_0 + d - 1$ as well. \square

References

1. Akhtar S, Bernstein DS (2005) Lyapunov-stable discrete-time model reference adaptive control. In: 2005 American control conference, Portland, OR, USA, pp 3174–3179
2. Anderson BDO, Johnson CR (1982) Exponential convergence of adaptive identification and control algorithms. *Automatica* 18(1):1–13
3. Feuer A, Morse AS (1978) Adaptive control of single-input, single-output linear systems. *IEEE Trans Autom Control* 23(4):557–569
4. Fu M, Barmish BR (1986) Adaptive stabilization of linear systems via switching control. *IEEE Trans Autom Control* AC–31:1097–1103
5. Goodwin GC, Ramadge PJ, Caines PE (1980) Discrete time multivariable control. *IEEE Trans Autom Control* AC–25:449–456
6. Goodwin GC, Sin KS (1984) Adaptive filtering prediction and control. Prentice Hall, Englewood Cliffs, NJ
7. Hespanha JP, Liberzon D, Morse AS (2003) Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica* 39:263–272
8. Hespanha JP, Liberzon D, Morse AS (2003) Overcoming the limitations of adaptive control by means of logic-based switching. *Syst Control Lett* 49(1):49–65
9. Ioannou PA, Tsakalis KS (1986) A robust direct adaptive controller. *IEEE Trans Autom Control* AC–31(11):1033–1043
10. Kreisselmeier G (1986) Adaptive control of a class of slowly time-varying plants. *Syst Control Lett* 8:97–103
11. Kreisselmeier G, Anderson BDO (1986) Robust model reference adaptive control. *IEEE Trans Autom Control* AC–31:127–133
12. Kristic M, Kokotovic PV, Kanellakopoulos I (1993) Transient-performance improvement with a new class of adaptive controllers. *Syst Control Lett* 21(6):451–461
13. Li Y, Chen H-F (1996) Robust adaptive pole placement for linear time-varying systems. *IEEE Trans Autom Control* AC–41:714–719
14. Middleton RH et al (1988) Design issues in adaptive control. *IEEE Trans Autom Control* 33(1):50–58
15. Middleton RH, Goodwin GC (1988) Adaptive control of time-varying linear systems. *IEEE Trans Autom Control* 33(2):150–155
16. Miller DE (2003) A new approach to model reference adaptive control. *IEEE Trans Autom Control* AC–5:743–757
17. Miller DE (2006) Near optimal LQR performance for a compact set of plants. *IEEE Trans Autom Control* 9:1423–1439
18. Miller DE (2016) An adaptive controller to provide near optimal LQR performance. In: 2016 IEEE 55th conference on decision and control, Las Vegas, Nevada, pp 3813–3818
19. Miller DE (2017) A parameter adaptive controller which provides exponential stability: the first order case. *Syst Control Lett* 103:23–31
20. Miller DE (2017) Classical discrete-time adaptive control revisited: exponential stabilization. In: 1st IEEE conference on control technology and applications, Mauna Lani, HI, pp 1975–1980
21. Miller DE (1996) On necessary assumptions in discrete-time model reference adaptive control. *Int J Adapt Control Signal Process* 10(6):589–602
22. Miller DE, Davison EJ (1991) An adaptive controller which provides an arbitrarily good transient and steady-state response. *IEEE Trans Autom Control* AC–36:68–81
23. Miller DE, Davison EJ (1989) An adaptive controller which provides Lyapunov stability. *IEEE Trans Autom Control* AC–34:599–609

24. Miller DE, Mansouri N (2010) Model reference adaptive control using simultaneous probing, estimation, and control. *IEEE Trans Autom Control* 55(9):2014–2029
25. Miller DE, Shahab MT (2018) Classical pole placement adaptive control revisited: linear-like convolution bounds and exponential stability. *Math Control Signals Syst* 30(4):19
26. Miller DE, Shahab MT (2019) Classical d-step-ahead adaptive control revisited: linear-like convolution bounds and exponential stability. In: 2019 American control conference, Philadelphia, pp 417–411
27. Morse AS (1980) Global stability of parameter-adaptive control systems. *IEEE Trans Autom Control* AC–25:433–439
28. Morse AS (1996) Supervisory control of families of linear set-point controllers—part 1: exact matching. *IEEE Trans Autom Control* AC–41:1413–1431
29. Morse AS (1997) Supervisory control of families of linear set-point controllers—part 2: robustness. *IEEE Trans Autom Control* AC–42:1500–1515
30. Morse AS (1998) A bound for the disturbance to tracking error gain of a supervised set-point control system. In: Normand-Cyrot D (ed) *Perspectives in control*. Springer, London
31. Naik SM, Kumar PR, Ydstie BE (1992) Robust continuous-time adaptive control by parameter projection. *IEEE Trans Autom Control* AC–37:182–297
32. Narendra KS, Lin YH (1980) Stable discrete adaptive control. *IEEE Trans Autom Control* AC–25(3):456–461
33. Narendra KS, Lin YH, Valavani LS (1980) Stable adaptive controller design, part II: proof of stability. *IEEE Trans Autom Control* AC–25:440–448
34. Rohrs CE et al (1985) Robustness of continuous-time adaptive control algorithms in the presence of unmodelled dynamics. *IEEE Trans Autom Control* AC–30:881–889
35. Shahab MT, Miller DE (2018) Multi-estimator based adaptive control which provides exponential stability: the first-order case. In: 2018 IEEE conference on decision and control, Miami Beach, USA, pp 2223–2228
36. Shahab MT, Miller DE (2019) A convolution bound implies tolerance to time-variations and unmodelled dynamics. [arXiv:1910.02112](https://arxiv.org/abs/1910.02112)
37. Tsakalis KS, Ioannou PA (1989) Adaptive control of linear time-varying plants: a new model reference controller structure. *IEEE Trans Autom Control* 34(10):1038–1046
38. Vale JR, Miller DE (2011) Step tracking in the presence of persistent plant changes. *IEEE Trans Autom Control* 1:43–58
39. Vu L, Chatterjee D, Liberzon D (2007) Input-to-state stability of switched systems and switching adaptive control. *Automatica* 43:639–646
40. Vu L, Liberzon D (2011) Supervisory control of uncertain linear time-varying systems. *IEEE Trans Autom Control* 1:27–42
41. Wen C (1994) A robust adaptive controller with minimal modifications for discrete time-varying systems. *IEEE Trans Autom Control* AC–39(5):987–991
42. Wen C, Hill DJ (1992) Global Boundedness of discrete-time adaptive control using parameter projection. *Automatica* 28(2):1143–1158
43. Ydstie BE (1989) Stability of discrete-time MRAC revisited. *Syst Control Lett* 13:429–439
44. Ydstie BE (1992) Transient performance and robustness of direct adaptive control. *IEEE Trans Autom Control* 37(8):1091–1105
45. Zhivoglyadov PV, Middleton RH, Fu M (2001) Further results on localization-based switching adaptive control. *Automatica* 37:257–263
46. Zhou K, Doyle JC, Glover K (1995) *Robust and optimal control*. Prentice Hall, Englewood Cliffs, NJ