Weighted l_2-norms for Analysis of an Adaptive Control Loop based on a Non-linear Model

Tor A. Johansen

Department of Engineering Cybernetics,
Norwegian Institute of Technology, 7034 Trondheim, Norway.
Tor.Arne.Johansen@regtek.sintef.no

Abstract

Here it is shown that the use of weighted l_2-norms is a useful tool for the analysis of adaptive control loops based on non-linear models. We apply a simplified feedback linearizing controller based on a linearly parameterized non-linear discrete-time input/output model with slowly time-varying parameters. The main assumption that is required in order to apply weighted l_2-norms for analysis is that the non-linearities must satisfy some global growth conditions.

1 Introduction

The application of input/output stability theory [1], and in particular weighted L_2-norms in the continuous-time case [2, 3], and weighted l_2-norms in the discrete time case [4], has turned out to be attractive for the analysis of stability and robustness of adaptive control loops based on linear models.

The purpose of the present paper is to illustrate the application of weighted l_2-norms for analysis of adaptive controllers based on non-linear models. The control structure we have chosen is a simple feedback linearizing controller [5, 6, 7] based on a non-linear discrete-time input/output model with slowly time-varying parameters. The parameter estimation algorithm applies normalization [8] and a relative dead-zone [9] to ensure robustness.

The paper is organized as follows: First, we review the essentials of weighted l_2-norms. Then we present the model representation in section 3, and the adaptive control structure in section 4. The parameter estimation algorithm and the closed loop stability are analyzed in sections 5 and 6, respectively. The paper ends with some discussion of the assumptions made, and some concluding remarks.

2 Preliminaries

The Euclidean norm of a vector $x \in R^n$ is $||x|| = \sqrt{x^T x}$. The difference sequence Δs of a sequence $s = (s(t))$ is defined by $\Delta s(t) = s(t) - s(t-1)$ for all $t \geq 0$. The truncation of a sequence s at time t is denoted s_t. The exponentially weighted l_2-norm of a sequence is defined by

$$||s_t||_2^\delta \quad = \quad \sqrt{\sum_{\tau=0}^{t} \delta^{t-\tau} ||s(\tau)||^2}$$

for any $\delta \in [0,1]$ and $t \geq 0$. The norm of a discrete-time transfer function $H(q)$, where q is the one-step prediction operator and $H(q)$ has all poles strictly inside the unit disc, is defined by

$$||H(q)||_\infty \quad = \quad \sup_{\omega \in [0,2\pi]} \left| H\left(e^{j\omega}\right) \right|$$

If $H(q)$ is proper and has all poles on or inside the disc $|c| \leq \sqrt{\delta}$, then the exponentially weighted norm is defined by $||H(q)||_\infty^\delta = ||H(\sqrt{\delta} q)||_\infty$. The mixed notation $s_2(t) = H(q^{-1}) s_1(t)$, where $H(q^{-1})$ is an exponentially stable transfer function, means the convolution $s_2(t) = h(t) \star s_1(t)$, where $h(t)$ is the inverse z-transform of the transfer function $H(q^{-1})$. The exponentially decaying term, which may appear as a consequence of non-zero initial conditions, is consistently neglected in this work, as it does not affect the stability analysis. The following lemmas will be used frequently:

Lemma 1 *Let $s_2(t) = H(q^{-1}) s_1(t)$, where $H(q^{-1})$ is a proper causal transfer function that has all poles in the disc $|c| \leq \sqrt{\delta}$, where $\delta \in (0,1]$, and suppose $\sum_{\tau=0}^{t} s_1^2(\tau) < \infty$. Then*

$$||(s_2)_t||_2^\delta \quad \leq \quad ||H(q^{-1})||_\infty^\delta ||(s_1)_t||_2^\delta$$

for all $t \geq 0$

Proof: Given in [4].
□

Lemma 2 *Let s_1 be a positive sequence, i.e. $s_1(t) \geq 0$ for all t. Let the transfer function $H(q^{-1})$ be defined by $H(q^{-1}) = K(1 - \sigma q^{-1})^{-1}$ for some $K \geq 0$ and $\sigma \in [0,1)$. Suppose the sequence s_2 satisfy*

$$s_2(t) \ \leq \ H(q^{-1}) \left(\gamma_0 s_1(t) + \gamma_1 s_1(t-1) + \ldots + \gamma_n s_1(t-n) \right)$$

for some non-negative integer n and constants $\gamma_0, \ldots ,\gamma_n \geq 0$. Then,

$$s_2(t) \leq \quad H(q^{-1}) \left(\gamma_0 + \gamma_1 \sigma^{-1} + \ldots + \gamma_n \sigma^{-n} \right) s_1(t)$$

Proof: Simple.
□

3 Model Representation

We consider linearly parameterized non-linear discrete-time input/output models, written in the predictor form

$$\begin{aligned} y(t+d) \ = \ & \varphi^T(y(t), \ldots, y(t-m), u(t), \ldots, u(t-r)) \theta^\star(t) \\ & + \nu(t+d) \end{aligned} \tag{1}$$

with scalar input and output sequences u and y, and m and r are non-negative integers. The integer $d \geq 1$ is the system's time-delay. Alternatively, it can be interpreted as the system's relative degree [5]. The sequence ν is unstructured uncertainty, which contains unmodeled dynamics (modeling error and reduced order model), disturbances, noise, effects due to sampling of continuous-time signals etc. The function φ is possibly non-linear, and the parameter sequence $\theta^\star$ may be slowly time-varying.

AP. There exists a constant $\underline{\theta} \geq 0$, such that $||\Delta\theta^\star(t)|| \leq \underline{\theta}$ for all $t \geq 0$. In addition, we know a convex and compact set Θ such that $\theta^\star(t) \in \Theta$ for all $t \geq 0$. $\square$

This means that the parametric variation from time $t-1$ to t is bounded in norm by the small constant $\underline{\theta}$. Contrary to the time-invariant case, we need to assume that Θ is compact, in order to show boundedness of the parameter estimate. As a consequence, there exists a $\overline{\theta} \geq 0$ such that $\theta_1, \theta_2 \in \Theta$ implies $||\theta_1 - \theta_2|| \leq \overline{\theta}$.

AB. The sign of the high-frequency gain is uniformly constant and known. Without loss of generality, we assume here that it is positive:

$$\frac{\partial\varphi^T(y(t),...,y(t-m),u(t),...,u(t-r))\theta}{\partial u(t)} \geq \beta_0 > 0 \tag{2}$$

holds uniformly for all $\theta \in \Theta$, input sequences u, and output sequences y. It is assumed that φ is a continuously differentiable function of $u(t)$. Moreover, φ is bounded by an affine function, i.e. there exists a constant $C \geq 0$ such that

$$\begin{aligned} ||\varphi(\cdot)|| \quad &\leq \quad C(1+|y(t)|+...+|y(t-m)| \\ &\qquad +|u(t)|+...+|u(t-r)|) \end{aligned}$$

$\square$

AB contains a strong global controllability assumption. In the usual global controllability definition, it is required that any state can be reached from any initial state within an unspecified, but finite time [7]. Our definition requires this time to be d, which is the shortest possible. On the other hand, we are only concerned about the output, not the full state.

AR. Let $y^\star$ be a bounded reference sequence that is known d steps ahead in time. The bound is denoted $K^\star = \sup_{t\geq 0}|y^\star(t)|$. $\square$

The problem we address here is the one of asymptotically tracking this reference. In other words, if the tracking error is defined as $\tilde{y} = y^\star - y$, then the control objective is to make $|\tilde{y}(t)|$ as small as possible as $t \to \infty$.

4 Adaptive Control Structure

With the above model representation, the adaptive certainty equivalence feedback linearizing controller is defined by the implicit equation

$$v(t) = \varphi^T(y(t),...,y(t-m),u(t),...,u(t-r))\hat{\theta}(t) \tag{3}$$

where $\hat{\theta}(t)$ is an estimate of $\theta^\star(t)$. From (2) it follows that there always exists a control input $u(t)$ that satisfies (3), and it is unique. The closed loop behavior is now described by

$$y(t+d) \quad = \quad v(t) + \varphi^T(t)(\theta^\star(t) - \hat{\theta}(t)) + \nu(t+d)$$

The following linear control algorithm is applied to control this nominally linear system

$$v(t) \quad = \quad y^\star(t+d) + G(q^{-1})(y^\star(t) - y(t)) \tag{4}$$

where $G(q^{-1}) = P(q^{-1})/Q(q^{-1})$. Now the closed loop satisfy

$$\tilde{y}(t) = M(q^{-1})\left(\varphi^T(t-d)(\hat{\theta}(t-d) - \theta^\star(t-d)) - \nu(t)\right) \tag{5}$$

where

$$M(q^{-1}) = \frac{1}{1+G(q^{-1})q^{-d}} = \frac{Q(q^{-1})}{Q(q^{-1}) + q^{-d}P(q^{-1})}$$

The effect of the parametric and unstructured uncertainty in (5) will be analyzed in the remaining of this paper.

5 Parameter Estimation

The estimate $\hat{\theta}(t)$ of $\theta^\star(t)$ is based on the predictor $\hat{y}(t|t-1) = \varphi^T(t-d)\hat{\theta}(t-1)$ which gives the prediction error

$$\begin{aligned} e(t) &= \hat{y}(t|t-1) - y(t) \\ &= \varphi^T(t-d)\hat{\theta}(t-1) - \varphi^T(t-d)\theta^\star(t-d) - \nu(t) \end{aligned} \tag{6}$$

and normalized prediction error $\varepsilon(t) = e(t)/n(t)$, where the scalar normalizing sequence n is defined by

$$n^2(t) \quad = \quad \delta n^2(t-1) + 1 + ||\varphi(t-d)||^2$$

with $n(0) > 0$ and $\delta \in (0,1)$ given. The normalization is introduced to ensure that the normalized unstructured uncertainty is bounded [8]:

AU. The unstructured uncertainty satisfies $|\nu(t)| \leq \mathcal{V}n(t)$ for some constant $\mathcal{V} \geq 0$. $\square$

We apply a parameter estimation algorithm with normalization and relative dead-zone [9]:

$$\begin{aligned} \hat{\theta}'(t) \quad &= \quad \hat{\theta}(t-1) \\ &\qquad -\lambda\frac{\varphi(t-d)}{1+\varphi^T(t-d)\varphi(t-d)}n(t)\mathcal{D}\left(\varepsilon(t)\right) \end{aligned} \tag{7}$$

$$\hat{\theta}(t) \quad = \quad \mathcal{P}_\Theta\left(\hat{\theta}'(t)\right) \tag{8}$$

$$\mathcal{D}(\varepsilon(t)) \quad = \quad \begin{cases} \varepsilon(t) + d_0 & \text{if } \varepsilon(t) < -d_0 \\ 0 & \text{if } |\varepsilon(t)| \leq d_0 \\ \varepsilon(t) - d_0 & \text{if } \varepsilon(t) > d_0 \end{cases} \tag{9}$$

where $\mathcal{P}_\Theta$ is a continuous parameter projection that projects its argument to the closest point (using Euclidean norm) in Θ. The continuous function $\mathcal{D}$ is referred to as a dead-zone function, the constant $d_0 \geq 0$ is the dead-zone, and $\lambda > 0$ is the estimator gain. For convenience, the parameter error sequence is defined by $\tilde{\theta}(t) = \hat{\theta}(t) - \theta^\star(t)$. Next, we examine the properties of this algorithm:

Lemma 3 *Suppose the outputs are generated by*

$$y(t+d) \quad = \quad \varphi^T(t)\theta^\star(t) + \nu(t+d)$$

and consider the parameter estimation algorithm (7)-(9) with initial estimate $\hat{\theta}(0) \in \Theta$. If AP and AU hold, the gain satisfies $\lambda \in (0,2)$, and the dead-zone is chosen such that $d_0 = \mathcal{V}$, then the algorithm has the properties

$$\begin{aligned}\sum_{\tau=0}^{t} \mathcal{D}^2(\varepsilon(\tau)) \quad &\leq \quad \Bigg(\frac{2}{\lambda(2-\lambda)}\left((d\underline{\theta})^2\right.\\ &\left.+2(1+\lambda)d\underline{\theta}\overline{\theta}\right) + 2(d-1)^2\underline{\theta}^2\Bigg)t + \frac{2}{\lambda(2-\lambda)}\overline{\theta}^2 \quad (10)\\ |\varepsilon(t)| \quad &\leq \quad \mathcal{V} + |\mathcal{D}(\varepsilon(t))| \quad (11)\\ \sum_{\tau=0}^{t} ||\Delta\hat{\theta}(\tau)||^2 \quad &\leq \quad \Bigg(\frac{\lambda}{2-\lambda}\left((d\underline{\theta})^2 + 2(1+\lambda)d\underline{\theta}\overline{\theta}\right)\\ &+(d-1)^2\underline{\theta}^2 + 2\lambda(d-1)\underline{\theta}\overline{\theta}\Bigg)t + \frac{\lambda}{2-\lambda}\overline{\theta}^2 \quad (12)\end{aligned}$$

for all $t \geq 0$.

Proof: Rewriting the equation for the prediction error (6) gives

$$\begin{aligned}\varepsilon(t) \quad &= \quad \frac{\varphi^T(t-d)\tilde{\theta}(t-1)}{n(t)}\\ &-\frac{\varphi^T(t-d)}{n(t)}\left(\theta^\star(t-d) - \theta^\star(t-1)\right) - \frac{\nu(t)}{n(t)}\end{aligned}$$

Since $d_0 = \mathcal{V}$ there exists a sequence β with the properties $0 \leq \beta(t) \leq 1$ and

$$\begin{aligned}\mathcal{D}(\varepsilon(t)) \quad &= \quad \beta(t)\frac{\varphi^T(t-d)\tilde{\theta}(t-1)}{n(t)}\\ &-\beta(t)\frac{\varphi^T(t-d)}{n(t)}\left(\theta^\star(t-d) - \theta^\star(t-1)\right)\end{aligned}$$

From (7) we find

$$\tilde{\theta}'(t) \; = \; \left(I - \lambda\beta(t)\frac{\varphi(t-d)\varphi^T(t-d)}{1+\varphi^T(t-d)\varphi(t-d)}\right)\tilde{\theta}(t-1) + \gamma(t) \quad (13)$$

where

$$\begin{aligned}\gamma(t) \quad &= \quad -\Delta\theta^\star(t) + \lambda\beta(t)\frac{\varphi(t-d)\varphi^T(t-d)}{1+\varphi^T(t-d)\varphi(t-d)}\\ &\cdot\left(\theta^\star(t-d) - \theta^\star(t-1)\right)\end{aligned}$$

Defining the function $V(t) = ||\tilde{\theta}(t)||^2$, we find $V(t) \leq ||\tilde{\theta}'(t)||^2$ from the convexity of Θ. Hence,

$$\begin{aligned}&V(t) - V(t-1) \quad \leq\\ &\left(\gamma^T(t) + \tilde{\theta}^T(t-1)\left(I - \lambda\beta(t)\frac{\varphi(t-d)\varphi^T(t-d)}{1+\varphi^T(t-d)\varphi(t-d)}\right)\right)\\ &\cdot\left(\left(I - \lambda\beta(t)\frac{\varphi(t-d)\varphi^T(t-d)}{1+\varphi^T(t-d)\varphi(t-d)}\right)\tilde{\theta}(t-1) + \gamma(t)\right)\\ &-\tilde{\theta}^T(t-1)\tilde{\theta}(t-1)\end{aligned}$$

Assume $\beta(t) \neq 0$, then

$$\begin{aligned}V(t) - V(t-1) \quad &\leq \quad \gamma^T(t)\gamma(t)\\ &+2\gamma^T(t)\left(I - \lambda\beta(t)\frac{\varphi(t-d)\varphi^T(t-d)}{1+\varphi^T(t-d)\varphi(t-d)}\right)\tilde{\theta}(t-1)\\ &-\left(\frac{2\lambda}{\beta(t)} - \lambda^2\right)\left(\frac{n^2(t)}{1+\varphi^T(t-d)\varphi(t-d)}\right)\\ &\cdot\left(\mathcal{D}(\varepsilon(t)) + \beta(t)\frac{\varphi^T(t-d)}{n(t)}\left(\theta^\star(t-d) - \theta^\star(t-1)\right)\right)^2\end{aligned}$$

or

$$\begin{aligned}&\left(\mathcal{D}(\varepsilon(t)) + \beta(t)\frac{\varphi^T(t-d)}{n(t)}\left(\theta^\star(t-d) - \theta^\star(t-1)\right)\right)^2\\ &\leq \frac{\beta(t)}{\lambda(2-\beta(t)\lambda)}\\ &\cdot\left(||\gamma(t)||^2 + 2(1+\lambda)||\gamma(t)|| \cdot ||\tilde{\theta}(t-1)|| - \Delta V(t)\right) \quad (14)\end{aligned}$$

If $\beta(t) = 0$, then $V(t) = V(t-1)$ and $\mathcal{D}(\varepsilon(t)) = 0$. Hence, from the properties of β it follows that

$$\begin{aligned}\mathcal{D}^2(\varepsilon(t)) \quad &\leq \quad \frac{2}{\lambda(2-\lambda)}\Big(||\gamma(t)||^2\\ &+2(1+\lambda)||\gamma(t)|| \cdot ||\tilde{\theta}(t-1)|| - \Delta V(t)\Big)\\ &+2||\theta^\star(t-d) - \theta^\star(t-1)||^2\end{aligned}$$

Summing, and using the fact that $||\tilde{\theta}(t)|| \leq \overline{\theta}$, which implies $V(t) \leq \overline{\theta}^2$ for all $t \geq 0$, together with the fact that AP implies $||\gamma(t)|| \leq d\underline{\theta}$, we get (10). By the definition of the dead-zone function, (11) follows. Finally, we get from (13) that

$$\begin{aligned}\sum_{\tau=0}^{t} ||\Delta\hat{\theta}(\tau)||^2 \quad &\leq \quad \sum_{\tau=0}^{t} ||\hat{\theta}'(\tau) - \hat{\theta}(\tau-1)||^2\\ \leq \quad &\sum_{\tau=0}^{t}\Big((\gamma(\tau) + \Delta\theta^\star(\tau))^T(\gamma(\tau) + \Delta\theta^\star(\tau))\\ &-2(\gamma(\tau) + \Delta\theta^\star(\tau))^T\lambda\beta(\tau)R(\tau)\tilde{\theta}(\tau-1)\\ &\lambda^2\beta^2(\tau)\tilde{\theta}^T(\tau-1)R(\tau)R(\tau)\tilde{\theta}(\tau-1)\Big)\\ \leq \quad &\sum_{\tau=0}^{t}\Big(||\gamma(\tau) + \Delta\theta^\star(\tau)||^2\\ &+2\lambda||\gamma(\tau) + \Delta\theta^\star(\tau)|| \cdot ||\tilde{\theta}(\tau-1)||\\ &+\lambda^2\frac{n^2(\tau)}{1+\varphi^T(\tau-d)\varphi(\tau-d)}\\ &\cdot\left(\mathcal{D}(\varepsilon(\tau)) + \beta(\tau)\frac{\varphi^T(\tau-d)}{n(\tau)}\left(\theta^\star(\tau-d) - \theta^\star(\tau-1)\right)\right)^2\Big)\end{aligned}$$

where

$$R(\tau) \quad = \quad \frac{\varphi(\tau-d)\varphi^T(\tau-d)}{1+\varphi^T(\tau-d)\varphi(\tau-d)}$$

Now, (12) follows from (14).

□

The relative dead-zone is a modification that turns the parameter estimator off when the normalized prediction error becomes small, and avoids therefore drift phenomena that otherwise might be excitated by the unstructured uncertainty. There are, however, several other modifications available, that leads to parameter estimation algorithms with similar properties and a robust adaptive control system. These include different variations of σ-modification, ϵ-modification, and parameter projection [10, 2, 4, 11].

6 Closed Loop Stability

With the feedback linearizing control structure, it is well known that certain states may be unobservable [12, 5]. In order to ensure boundedness of these states, we must study the behavior of the inverse system. By the Implicit Function Theorem, e.g. [7], and assumption AB, it is evident that the inverse system is globally defined by the function g:

$$\begin{aligned} u(t) \quad = \quad & g(u(t-1), ..., u(t-r), y(t+d), y(t), \\ & ..., y(t-m), \nu(t+d), \theta^\star(t)) \end{aligned} \tag{15}$$

where y and ν are viewed as inputs, and u as the output.

AI. The inverse system (15) is globally uniformly exponentially stable, in the sense that its impulse response coefficients are bounded by an exponentially decaying sequence. □

Under this assumption, we can prove the following theorem, which essentially shows that the input is bounded by the output and the unstructured uncertainty:

Lemma 4 *Suppose AB and AI hold, then there exist constants $\sigma \in [0,1)$ and $K_\sigma \geq 0$ such that for any $\kappa \in (\sigma, 1)$*

$$||(q^{-d}u)_t||_2^\kappa \quad \leq \quad \frac{K_\sigma}{1-\sqrt{\sigma/\kappa}} \left(||y_t||_2^\kappa + ||\nu_t||_2^\kappa\right)$$

Proof. It follows directly from AI that for some $K_\sigma \geq 0$

$$|u(t)| \quad \leq \quad K_\sigma \sum_{\tau=0}^{t} \sqrt{\sigma}^{t-\tau} \left(|y(\tau+d)| + |\nu(\tau+d)|\right)$$

the result follows from Lemma 1 together with $||(1-\sqrt{\sigma}q^{-1})^{-1}||_\infty^\kappa = (1-\sqrt{\sigma/\kappa})^{-1}$.

□

We are now in position to prove the main result:

Theorem 1 *Suppose the system (1) is controlled by (3) and (4), and the algorithm (7)-(9) is applied to estimate the parameter sequence $\theta^\star$. Assume AB, AR, AP, AU and AI hold, and let $\rho \in [0,1)$ be a constant such that the impulse response coefficients of $M(q^{-1})$ are bounded by $\sqrt{\rho}^t$. Suppose the dead-zone is chosen as $d_0 = \mathcal{V}$. Let κ be arbitrary in the open interval $\max(\delta, \sigma, \rho) < \kappa < 1$. If $\underline{\theta}$ and $\mathcal{V}$ are both sufficiently small, then for arbitrary initial conditions and $\hat{\theta}(0) \in \Theta$, all variables in the closed loop are bounded, and the tracking error satisfies*

$$\frac{1}{t}\sum_{\tau=0}^{t} \tilde{y}^2(\tau) \quad \leq \quad c_1 + \frac{1}{t}c_2 \tag{16}$$

for all $t > 0$, where c_2 is proportional to $\overline{\theta}^2$, and $c_1 \to 0$ when $\underline{\theta} \to 0$ and $\mathcal{V} \to 0$.

Proof: Let us define

$$\tilde{Y}(t) \quad = \quad \varphi^T(t-d)\tilde{\theta}(t-d) - \nu(t) \tag{17}$$

Combining (6) and (17), we get

$$\frac{\tilde{Y}(t)}{n(t)} = \varepsilon(t) + \frac{\varphi^T(t-d)}{n(t)}\left(\hat{\theta}(t-d) - \hat{\theta}(t-1)\right) \tag{18}$$

From the definition of the normalizing signal, assumption AB and Lemma 2, it is evident that there exists a constant C such that

$$n^2(t) \leq \frac{1}{1-\delta} + C^2 \sum_{\tau=0}^{t} \kappa^{t-\tau}\left(1 + |y(\tau-d)| + |u(\tau-d)|\right)^2$$

Using Lemma 4 and the triangle inequality, we find

$$\begin{aligned} n^2(t) \quad \leq \quad & \frac{1+3C^2}{1-\kappa} + 3C^2\left(\left|\left|\left(q^{-d}y\right)_t\right|\right|_2^\kappa\right)^2 \\ & + \frac{3C^2K_\sigma^2}{(1-\sqrt{\sigma/\kappa})^2}\left(\left|\left|(|y|+|\nu|)_t\right|\right|_2^\kappa\right)^2 \end{aligned}$$

because $\kappa > \sigma$. Using Lemma 2, we get

$$n^2(t) \quad \leq \quad \frac{1+3C^2}{1-\kappa} + \left(K_y\,||y_t||_2^\kappa\right)^2 + \left(K_\nu\,||\nu_t||_2^\kappa\right)^2$$

where $K_y = \sqrt{3}C(\kappa^{-d/2} + \sqrt{2}K_\sigma(1-\sqrt{\sigma/\kappa})^{-1})$ and $K_\nu = \sqrt{6}CK_\sigma(1-\sqrt{\sigma/\kappa})^{-1}$. By the triangle inequality, the bound on $y^\star$, and (5), we get

$$n^2(t) \leq K_0 + \left(K_y(1-\sqrt{\rho/\kappa})^{-1}||\tilde{Y}_t||_2^\kappa\right)^2 + \left(K_\nu||\nu_t||_2^\kappa\right)^2$$

where $K_0 = (1 + 3C^2 + 3C^2K_y^2(K^\star)^2)/(1-\kappa)$. This can be written

$$\begin{aligned} n^2(t) \quad \leq \quad & K_0 + K_y^2\frac{\kappa}{(\sqrt{\kappa}-\sqrt{\rho})^2}\sum_{\tau=0}^{t}\kappa^{t-\tau}\tilde{Y}^2(\tau) \\ & + K_\nu^2\sum_{\tau=0}^{t}\kappa^{t-\tau}\nu^2(\tau) \end{aligned}$$

After some straightforward manipulations and using (18)

$$\begin{aligned} n^2(t) \quad \leq \quad & K_0 + K_y^2\frac{\kappa}{(\sqrt{\kappa}-\sqrt{\rho})^2}\sum_{\tau=0}^{t}\kappa^{t-\tau}s^2(\tau)n^2(\tau) \\ & + K_\nu^2\sum_{\tau=0}^{t}\kappa^{t-\tau}\nu^2(\tau) \end{aligned}$$

where we have defined

$$s^2(t) = \left(\varepsilon(t) + \frac{\varphi^T(t-d)}{n(t)}\left(\hat{\theta}(t-d) - \hat{\theta}(t-1)\right)\right)^2 \tag{19}$$

It is straightforward to show that there exists a finite constant $B \geq 1$ such that

$$\frac{n^2(t+1)}{n^2(t)} \quad \leq \quad B$$

for all $t \geq 0$. Hence

$$n^2(t) \quad \leq \quad K_0 + \sum_{\tau=0}^{t-1} \kappa^{t-\tau} w^2(\tau) n^2(\tau) \qquad (20)$$

where $w^2(t) = B\kappa^{-1}K_\nu^2\mathcal{V}^2 + BK_y^2(\sqrt{\kappa}-\sqrt{\rho})^{-2}s^2(t)$. By the Gronwall lemma, e.g. [1, 4], and the relation between arithmetic and geometric means, e.g. [13], we get from (20)

$$\begin{aligned} n^2(t) \quad &\leq \quad K_0 + K_0 \sum_{\tau=0}^{t-1} \kappa^{t-\tau} w^2(\tau) \\ &\cdot \left(1 + \frac{\kappa^{-1}}{t-\tau-1} \sum_{i=\tau+1}^{t-1} w^2(i)\right)^{t-\tau-1} \end{aligned}$$

From Lemma 3, it follows after some manipulation that

$$\sum_{\tau=0}^{t} w^2(\tau) \quad \leq \quad \mu_0 t + \mu_1 \qquad (21)$$

where

$$\begin{aligned} \mu_0 \quad &= \quad \left(B\kappa^{-1}K_\nu^2 + \frac{2BK_y^2}{(\sqrt{\kappa}-\sqrt{\rho})^2}\right)\mathcal{V}^2 \\ &+ \frac{4Bd^2K_y^2}{(\sqrt{\kappa}-\sqrt{\rho})^2}\left(\frac{3+\lambda}{\lambda(2-\lambda)}\right)\underline{\theta}^2 \\ &+ \frac{2BdK_y^2}{(\sqrt{\kappa}-\sqrt{\rho})^2}\left(\frac{\lambda^3+4\lambda^2+4\lambda+4}{\lambda(2-\lambda)}\right)\overline{\theta}\underline{\theta} \qquad (22) \\ \mu_1 \quad &= \quad \frac{2BK_y^2}{(\sqrt{\kappa}-\sqrt{\rho})^2}\left(\frac{\lambda^2+2}{\lambda(2-\lambda)}\right)\overline{\theta}^2 \end{aligned}$$

Now,

$$\begin{aligned} n^2(t) \quad &\leq \quad K_0 + K_0 \sum_{\tau=0}^{t-1} \kappa^{t-\tau} w^2(\tau)(1+\kappa^{-1}\mu_0)^{t-\tau-1} \\ &\cdot \left(1 + \frac{\mu_1}{(\kappa+\mu_0)(t-\tau-1)}\right)^{t-\tau-1} \end{aligned}$$

Since $(1+x^{-1})^x \leq e$ for all $x > 0$, it follows that

$$\begin{aligned} n^2(t) \quad &\leq \quad K_0 + K_0(1+\kappa^{-1}\mu_0)^{-1} \\ &\cdot \exp\left(\frac{\mu_1}{\kappa+\mu_0}\right) \sum_{\tau=0}^{t-1} (\kappa+\mu_0)^{t-\tau} w^2(\tau) \end{aligned}$$

Taking $\underline{\theta}$ and $\mathcal{V}$ sufficiently small, it follows from (22) that we can make $\mu_0 + \kappa < 1$. Hence, (cf. Lemma 2.4 in [4])

$$n^2(t) \quad \leq \quad K_0 + \frac{K_0(\mu_0+\mu_1)\exp\left(\frac{\mu_1}{\kappa+\mu_0}\right)}{(1-\kappa-\mu_0)(1+\kappa^{-1}\mu_0)} \;=\; \overline{n}^2$$

and it is proved that n is a bounded sequence. From (5) and Lemma 1 it follows that

$$\begin{aligned} \sum_{\tau=0}^{t} \tilde{y}^2(\tau) \quad &\leq \quad \frac{1}{(1-\sqrt{\rho})^2} \sum_{\tau=0}^{t} \tilde{Y}^2(\tau) \\ &\leq \quad \frac{1}{(1-\sqrt{\rho})^2} \sum_{\tau=0}^{t} s^2(\tau) n^2(\tau) \\ &\leq \quad \frac{2\overline{n}^2}{(1-\sqrt{\rho})^2} \sum_{\tau=0}^{t} \left(\varepsilon^2(\tau) + (d-1)^2 \left|\left|\Delta\hat{\theta}(\tau)\right|\right|^2\right) \end{aligned}$$

From (19) and Lemma 3, it is evident that (16) holds, and c_1 and c_2 have the stated properties. It is easy to see that $\tilde{y}$ must be a bounded sequence. Hence, y is bounded since $y^\star$ is bounded. Since ν is bounded by n, it follows from Lemma 4 that u is bounded as well.

□

7 Discussion and Concluding Remarks

The contribution of this paper is a demonstration of the use of weighted l_2-norms [2, 4] for stability and robustness analysis of adaptive controllers based on non-linear models. Sufficient conditions for robust stability of the adaptive control loop are provided by Theorem 1. Stability requires the unstructured uncertainty $\mathcal{V}$ and parametric variations $\underline{\theta}$ both to be sufficiently small. The performance result (16) is conservative and should be interpreted in a qualitative way. The average squared tracking error will be bounded by a sum of two terms. The first term is the asymptotic performance bound that scales with the uncertainty and parameter variations. The second term is a bound on the transient performance that scales with the bound $\overline{\theta}^2$ on the parameter error, and vanishes asymptotically at the rate $1/t$.

It is clear that the model representation (1) contains two major limitations

1. Assumption AB requires $y(t+d)$ to depend strictly monotoneously on $u(t)$. Moreover, the model cannot contain actuator saturation. Typically, if this is violated, a solution to (3) will only exist for $v(t)$ (or $y^\star(t)$) restricted to some subset of R. Then it may be necessary to modify the controller for $v(t)$ to ensure that $v(t)$ remains within its prescribed subset, and analyze the effects of these modifications on the closed loop. Since these problems are of a more general nature and not related to the use of weighted l_2-norms, we have chosen to apply the restrictive assumption AB in order to simplify the presentation.

2. The predictor is linearly parameterized. This assumption can be relaxed to include non-linearly parameterized models

$$\begin{aligned} y(t+d) \;=\; & f(y(t),...,y(t-m),u(t),..., \\ & u(t-r),\theta^\star(t)) + \nu(t+d) \end{aligned} \qquad (23)$$

where f is a non-linear function that is twice continuously differentiable with respect to θ. With the certainty equivalence controller

$$v(t) = f(y(t), ..., y(t-m), u(t), ..., u(t-r), \hat{\theta}(t))$$

we get using Taylor's theorem

$$\tilde{y}(t) = M(q^{-1})\left(\frac{\partial f}{\partial \theta}(\cdot, \hat{\theta}(t-d))\tilde{\theta}(t-d) + \frac{1}{2}\tilde{\theta}(t-d)\frac{\partial^2 f}{\partial \theta^2}(\cdot, \overline{\theta}(t-d))\tilde{\theta}(t-d) + \nu(t)\right)$$

where $\overline{\theta}(t) = \alpha(t)\theta^\star(t) + (1-\alpha(t))\hat{\theta}(t)$ for some $\alpha(t) \in [0,1]$. Defining

$$\varphi(t) = \frac{\partial f}{\partial \theta}(\cdot, \hat{\theta}(t))$$
$$\eta(t) = \nu(t) + \frac{1}{2}\tilde{\theta}(t-d)\frac{\partial^2 f}{\partial \theta^2}(\cdot, \overline{\theta}(t-d))\tilde{\theta}(t-d)$$

we get

$$\tilde{y}(t) = M(q^{-1})(\varphi^T(t-d)\tilde{\theta}(t-d) + \eta(t))$$

Under similar conditions as before, it is clear that there exists a constant $K \geq 0$ such that $\eta(t) \leq \mathcal{V}'n(t)$ where

$$\mathcal{V}' = \mathcal{V} + K\overline{\theta}^2$$

Hence, if $\underline{\theta}, \overline{\theta}$ and $\mathcal{V}$ are all sufficiently small, then Theorem 1 still holds. The price to pay for the relaxed assumption is that the parametric uncertainty must be sufficiently small.

Assumptions AP, AU, and AR are standard in the certainty equivalence adaptive control literature. The invertability condition AI is imposed by the particular control algorithm, while the bound on φ, cf. AB, are directly related to the use of input/output global stability theory arguments, see also [15] for a related local Lyapunov approach. While the affine boundedness appears to strongly restrict the class of mathematical systems [6, 16], it is clear from a practical point of view that systems where the states do not diverge at a rate faster than exponential, will satisfy this assumption. We cannot think of any man-made physical system that does not satisfy this assumptions.

We have seen that the application of weighted l_2-norms gives a quite simple proofs under rather weak and transparent assumptions. However, while the results are qualitatively appealing and transparent, their quantitative significance is doubtful. This is related to the general conservativeness of global input/output stability theory applied to non-linear systems. More details and some extensions of this work can be found in [14].

References

[1] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properies*, Academic Press, New York, 1975.

[2] P. Ioannou and A. Datta, "Robust adaptive control: A unified approach", *Proceedings of the IEEE*, vol. 79, pp. 1736–1768, 1991, Also in *Foundations of Adaptive Control*, P. V. Kokotović, Ed., Springer-Verlag, Berlin, 1991.

[3] K. S. Tsakalis, "Robustness of model reference adaptive controllers: Input-Output properties", *IEEE Trans. Automatic Control*, vol. 37, pp. 556–565, 1992.

[4] A. Datta, "Robustness of discrete-time adaptive controllers: An input-output approach", *IEEE Trans. Automatic Control*, vol. 38, pp. 1852–1857, 1993.

[5] S. Monaco and D. Normand-Cyrot, "Minimum-phase nonlinear discrete-time systems and feedback stabilization", in *Proc. IEEE Conf. Decision and Control, Los Angeles, Ca.*, 1987, pp. 979–986.

[6] S. S. Sastry and A. Isidori, "Adaptive control of linearizable systems", *IEEE Trans. Automatic Control*, vol. 34, pp. 1123–1131, 1989.

[7] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.

[8] L. Praly, "Robust model reference adaptive controllers, Part I: Stability analysis", in *Proc. IEEE Conf. Decision and Control, Las Vegas, Nevada*, 1984, pp. 1009–1014.

[9] G. Kreisselmeier and B. D. O. Anderson, "Robust model reference adaptive control", *IEEE Trans. Automatic Control*, vol. 31, pp. 127–133, 1986.

[10] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[11] B. E. Ydstie, "Transient performance and robustness of direct adaptive control", *IEEE Trans. Automatic Control*, vol. 37, pp. 1091–1105, 1992.

[12] C. I. Byrnes and A. Isidori, "A frequency domain philosophy for nonlinear systems", in *Proceedings of the 1984 IEEE Conf. Decision and Control, Las Vegas, Nevada*, 1984, pp. 1569–1573.

[13] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1986.

[14] T. A. Johansen and M. M. Polycarpou, "Stable adaptive control of a general class of non-linear systems", in *Neural Adaptive Control Technology*, R. W. Zbikowski and K. Kunt, Eds., vol. 1. World Scientific, 1996.

[15] F.-C. Chen and H. K. Khalil, "Adaptive control of nonlinear systems using neural networks - a deadzone approach", in *Proc. American Control Conference, Boston*, 1991, pp. 667–672.

[16] C. A. Schwartz and I. M. Y. Mareels, "Comments on "Adaptive control of linearizable systems"", *IEEE Trans. Automatic Control*, vol. 37, pp. 698–701, 1992.