Available online at www.sciencedirect.com
SCIENCE DIRECT®

Systems & Control Letters 48 (2003) 15–25

SYSTEMS & CONTROL LETTERS
www.elsevier.com/locate/sysconle

Overparameterised adaptive controllers can reduce non-singular costs

F. Beleznay[1], M. French*

Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK

Received 25 March 2002; received in revised form 23 July 2002; accepted 13 August 2002

Abstract

By means of two examples we show that non-singular costs can been reduced for adaptive controllers by overparameterising the estimators. The examples are for scalar and second order systems respectively. In the second example the tuning function design and the overparameterised adaptive backstepping design are compared. In both cases a system is constructed for which the overparameterised design is superior w.r.t. a non-singular measure of transient performance.

Keywords: Overparameterisation; Back stepping; Performance

1. Introduction

Overparameterisation in adaptive controllers has often been considered to be a undesirable design feature. Generally overparameterisation leads to controllers of higher dynamic order with more complex dynamics and hence there is often a concern about a possible loss of robustness. There are a few results which partially support this intuitive claim, for example [7] presents concrete results showing parameter convergence for systems with fewer parameters, hence GAS of the closed loop and robustness to bounded disturbances. Often, informal reasonings for the need to reduce over-parameterisation are given as follows: as the dimensionality of the parameter space increases, parameter convergence is harder to achieve, which in turn can lead to robustness problems. Alternatively, it can be reasoned that because the parameter space is larger, the transient is larger. However, these arguments do not necessarily hold up to close scrutiny. For example, in the latter case, it is well known within the machine learning community that the size of parameters can be more critical for learning performance than the number of parameters, see e.g. [1]. Although the problem considered in [1] is not directly analogous, it exposes the weakness of the intuitive argument, by suggesting the 'size' of a learning problem may be dictated by measures of 'size' which are not as elementary as parameter counts.

On the other hand, overparameterisation can give the controller beneficial extra degrees of freedom in the design, see e.g. [5]. We exploit these extra degrees of freedom in the constructions which follow, which show that *the orthodoxy that one should not over-parameterise, is not, in general, valid.*[2] In

* Corresponding author. Tel.: +44-23-8059-2688; fax: +44-23-8059-4498.

E-mail addresses: beleznay@cs.elte.hu (F. Beleznay), mcf@ecs.soton.ac.uk (M. French).

[1] Present address: Department of analysis, Eötvös Lorand University, Budapest, Hungary.

[2] Related results in the context of function approximator based designs can be found in [4].

0167-6911/03/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0167-6911(02)00244-X

particular, we give two examples of systems and controllers whereby a sensible closed loop cost is lower for the overparameterised design. In this paper, we will consider the non-singular cost functionals

$$P = \|q^{\mathrm{T}}x\|_{L^\infty} + r\|u\|_{L^\infty},$$

$$P = \|q^{\mathrm{T}}x\|_{L^2}^2 + r\|u\|_{L^\infty}, \tag{1}$$

where x is the system state, u is the control and q, r are weighting parameters. We observe that it is critical to introduce a penalty on control effort in order to have a well-posed problem—the classes of systems considered in this paper are minimum phase and admit high gain solutions, which by trajectory initialisation [8] can yield arbitrarily small output responses at the price of high control effort. Furthermore, we need to be able to develop both upper and lower bounds for these costs. There are general results available for bounding the output transients of adaptive controllers, see e.g. [8], but in general the problem of bounding the control effort is much harder, however see e.g. [3,2].

The first example is a scalar system, and it is shown that the transient performance of a standard adaptive controller is improved when the estimator is over-parameterised. The essence of the example is that if an estimator has been driven to an overly high value (in this case by a large state initial condition), then it is desirable to 'forget' the 'large' value and restart adapting from a low value. Overparameterisation allows us to do this.

We then consider systems in strict feedback form, with adaptive backstepping controllers [6]. It is well known that such controllers are overparameterised (if n is the order of the system and p is the number of parameters then there are $n(n+1)p/2$ estimators). The tuning function design [7] eliminates the overparameterisation completely. The second contribution of this paper is to construct a situation in which the over-parameterisation of the adaptive backstepping design leads to a superior transient response to that of the tuning function design.

In particular, we construct a second-order system for which the following holds. Suppose we optimize the controllers for regulation to a point $y_{\mathrm{r}} = b$. Then the closed loop yield the same cost for either controller when applied with $y_{\mathrm{r}} = b$, whilst the adaptive backstepping controller has superior performance when $y_{\mathrm{r}} = a$. We interpret this result in a simple probabilistic setting. We are not in any way claiming that the adaptive backstepping design is superior to the tuning function design in general, as we fully expect that examples showing the contrary relationship between performance can also be established.

Ideally one would like to have results which fully characterise when to over-parameterise or not, or, more specifically, when to use the adaptive backstepping design and when to use the tuning function design. However, this remains an challenging open area of research. It is likely to be extremely difficult to achieve such characterisations, such an analysis would have to contend with the complexities of the backstepping transformations, the non-optimality of the controllers and the non-singularity of the cost, all of which lead to complex dynamics and to an extremely complex problem. An indication of the complexity of this problem may be obtained by considering the proof of Theorem 2 below. Whilst the system nonlinearities were chosen to simplify the problem as much as possible, the argument remains delicate.

Whilst the examples in this paper are extremely specialized, the paper makes a practical contribution: when designing adaptive controllers, do not dismiss over-parameterised designs; they may have superior performance!

2. A motivating scalar example

Consider the system

$$\Sigma(x_0, \theta)\colon\ \dot{x} = \theta f(x) + u, \quad x(0) = x_0, \tag{2}$$

where $\theta \in \mathbb{R}$ is an unknown parameter, x, u are real scalar valued signals, and $f : \mathbb{R} \to \mathbb{R}$ is a function which is of the form

$$f(x) = \phi_1(x) + \phi_2(x) = \phi_1(x) + x\psi(x), \tag{3}$$

where ϕ_1, ψ have the properties

1. ϕ_1 is continuous with support in $(-\infty, 1)$, and $\phi_1(0) \neq 0$.
2. ψ is continuous and non-negative with support in $(3, \infty)$, and $\psi(x) = 1\ \forall x \geqslant 4$.

We are interested in stabilising the system, i.e. achieving $x(t) \to 0$ as $t \to \infty$, whilst keeping all signals bounded. Consider the following two controllers:

$$\Xi^u(\alpha): u = -\hat{\theta} f(x) - x,$$
$$\dot{\hat{\theta}} = \alpha x f(x), \quad \hat{\theta}(0) = 0, \tag{4}$$

$$\Xi^o(\alpha_1, \alpha_2): u = -\hat{\theta}_1\phi_1(x) - \hat{\theta}_2\phi_2(x) - x,$$
$$\dot{\hat{\theta}}_1 = \alpha_1 x\phi_1(x), \quad \hat{\theta}_1(0) = 0,$$
$$\dot{\hat{\theta}}_2 = \alpha_2 x\phi_2(x), \quad \hat{\theta}_2(0) = 0. \tag{5}$$

Both controllers achieve stability $\forall x_0, \theta \in \mathbb{R}$. The proof is standard and can be achieved by considering the quadratic Lyapunov functions $V_u = \frac{1}{2}x^2 + (1/2\alpha)(\theta - \hat{\theta})^2$, $V_o = \frac{1}{2}x^2 + (1/2\alpha_1)(\theta - \hat{\theta}_1)^2 + (1/2\alpha_2)(\theta - \hat{\theta}_2)^2$, respectively.

We measure transient performance by the cost

$$P(\Sigma(\theta, x_0), \Xi(\alpha)) = \|x\|_{L^2}^2 + \|u\|_{L^\infty} \tag{6}$$

and it is convenient to define $P|_{[0,t]} = \|x\|_{L^2[0,t]}^2 + \|u\|_{L^\infty[0,t]}$, $P^o = P(\Sigma(\theta, x_0), \Xi^o(\alpha, \alpha))$, $P^u = P(\Sigma(\theta, x_0), \Xi^u(\alpha))$.

The result for this section is as follows:

Theorem 1. *For all* $\alpha > 0$, $\theta \in \mathbb{R}$,

$$\lim_{x_0 \to \infty} P^u - P^o = \infty. \tag{7}$$

The theorem therefore states that the difference between the basic design and its overparameterised variant becomes arbitrarily large as the size of the initial condition increases. In particular, the overparameterised design has superior performance.

Proof. Consider $\Sigma(\theta, x_0)$ with either Ξ^o or Ξ^u. Suppose $x_0 > 4$. Then by the fact that $x(t) \to 0$ as $t \to \infty$, and the definition of ϕ_1, ϕ_2 it follows that there is a unique time t^* at which $x(t^*) = 2$, since when $x(t) = 2$, then $\dot{x} = -x < 0$. Now observing that on $[0, t^*]$ Ξ^u, Ξ^o are identical, it follows that

$$P^u|_{[0,t^*]} = P^o|_{[0,t^*]}. \tag{8}$$

Now consider $(\Sigma(\theta, x_0), \Xi^o(\alpha, \alpha))$. Note that $\hat{\theta}_2$ is an increasing function, as its derivative is non-negative. We now claim that $\hat{\theta}_2(t^*) \to \infty$ as $x_0 \to \infty$. Suppose the contrary, i.e. that there exists $M > 0$, such that for some divergent sequence of points $\{x_{0i}\}_{i \geq 1}$, $\hat{\theta}_2(t^*) \leqslant M$. Let $t^{**} = \inf\{t \geqslant 0 : x(t) = 4\}$, and note that $0 \leqslant \hat{\theta}_2(t^{**}) \leqslant \hat{\theta}_2(t^*) \leqslant M$. Now

$$\begin{aligned} &\int_0^{t^{**}} x^2 \, \mathrm{d}t \\ &= \int_0^{t^{**}} -\dot{V} \, \mathrm{d}t = V(0) - V(t^{**}) \\ &\geqslant \frac{1}{2}x_{0_i}^2 - 8 + \frac{1}{2\alpha_2}(\theta^2 + (\theta + M)^2) \to \infty \\ &\quad \text{as } i \to \infty. \end{aligned} \tag{9}$$

Now, since $\dot{V} = -x^2$, we have

$$\begin{aligned} &\hat{\theta}_2(t^*) \geqslant \hat{\theta}_2(t^{**}) \\ &\quad = \int_0^{t^{**}} \dot{\hat{\theta}}_2 \, \mathrm{d}t = \int_0^{t^{**}} \alpha_2 x^2(t) \, \mathrm{d}t \to \infty \\ &\qquad \text{as } i \to \infty. \end{aligned} \tag{10}$$

This is a contradiction, hence $\hat{\theta}_2(t^*) \to \infty$ as $x_0 \to \infty$. The same argument for $(\Sigma(\theta, x_0), \Xi^u(\alpha))$ shows $\hat{\theta}(t^*) \to \infty$ as $x_0 \to \infty$.

It can easily be seen for the overparameterised controller Ξ^o, that $\|u\|_{L^\infty[t^*,\infty)}$ is independent of x_0 (this can be shown formally by considering $\Xi^r(0, \alpha)$ below), whilst

$$\begin{aligned} \|u\|_{L^\infty[0,t^*]} &\geqslant |-\hat{\theta}_2(t^{**})\phi_2(4) - 4| \\ &\geqslant 4\hat{\theta}_2(t^{**}) + 4 \to \infty \quad \text{as } x_0 \to \infty. \end{aligned} \tag{11}$$

It thus follows that for large x_0, the supremum for u is attained on $[0, t^*]$, and in particular it then follows for large x_0, that $\|u\|_{L^\infty}$ for Ξ^u is greater than or equal to that for Ξ^o. Since $\|x\|_{L^2}^2 = \|x\|_{L^2[0,t^*]}^2 + \|x\|_{L^2[t^*,\infty)}^2$, it then follows that for sufficiently large x_0,

$$\begin{aligned} P^u - P^o &\geqslant \int_{t^*}^{\infty} x_u^2 \, \mathrm{d}t - \int_{t^*}^{\infty} x_o^2 \, \mathrm{d}t \\ &\quad + \|u_u\|_{L^\infty(\mathbb{R}_+)} - \|u_o\|_{L^\infty(\mathbb{R}_+)} \\ &\geqslant \int_{t^*}^{\infty} x_u^2 \, \mathrm{d}t - \int_{t^*}^{\infty} x_o^2 \, \mathrm{d}t. \end{aligned} \tag{12}$$

We now claim that $\int_{t^*}^{\infty} x_u^2 \, \mathrm{d}t \to \infty$ as $x_0 \to \infty$, whilst $\int_{t^*}^{\infty} x_o^2 \, \mathrm{d}t$ is bounded independently of x_0. The result then follows. To show $\int_{t^*}^{\infty} x_o^2\,\mathrm{d}t$ is independent from x_0 it suffices to observe that the controller on $[t^*, \infty)$ is simply $\Xi^r(0, \alpha)$ for Ξ^o (and $\Xi^r(\hat{\theta}(t^*), \alpha)$ for Ξ^u), where $\Xi^r(\cdot, \cdot)$ is defined

$$\Xi^r(\mu_0, \alpha)\colon u = -\hat{\mu}\phi_1(x) - x, \qquad \dot{\hat{\mu}} = \alpha x \phi_1(x)\ \hat{\mu}(t^*) = \mu_0. \tag{13}$$

Clearly $\Xi^r(0, \alpha)$ is independent of x_0, hence so is $P^o|_{[t^*, \infty)}$. It remains to show that $\int_{t^*}^{\infty} x_u^2\,\mathrm{d}t \to \infty$ as $x_0 \to \infty$. By taking $\mu_0 = \hat{\theta}(t^*) \to \infty$ as $x_0 \to \infty$, it suffices to show that $\Xi^r(\mu_0, \alpha)$ causes $\int_{t^*}^{\infty} x^2\,\mathrm{d}t$ to diverge as $\mu_0 \to \infty$. Since $\phi_1(0) \neq 0$ and $\hat{\mu}, \theta$ is scalar, it follows that $\hat{\mu} \to \theta$ as $t \to \infty$ [7]. Now

$$\int_{t^*}^{\infty} x^2\,\mathrm{d}t = \int_{t^*}^{\infty} -\dot{V}\,\mathrm{d}t = V(t^*) - V(\infty) = 2 + \frac{1}{2\alpha}(\theta - \mu_0)^2 \to \infty \quad \text{as } \mu_0 \to \infty, \tag{14}$$

as required, thus completing the proof. □

3. Modification of the adaptive backstepping design

Consider a system in parametric strict feedback form, denoted by $\Sigma(\theta, \phi_1, \ldots, \phi_n, \mathbf{x}_0)$

$$\begin{aligned} \dot{x}_1 &= x_2 + \phi_1^{\mathrm{T}}(x_1)\theta, \\ &\ \vdots \qquad\qquad\qquad \mathbf{x}(0) = \mathbf{x}_0, \\ \dot{x}_{n-1} &= x_n + \phi_{n-1}^{\mathrm{T}}(x_1, \ldots, x_{n-1})\theta, \\ \dot{x}_n &= u + \phi_n^{\mathrm{T}}(x_1, \ldots, x_n)\theta, \\ y &= x_1, \end{aligned} \tag{15}$$

where $x_i\colon \mathbb{R}_+ \to \mathbb{R}$, $\theta \in \mathbb{R}^m$ for some m is an unknown parameter and $\phi_i\colon \mathbb{R}^i \to \mathbb{R}^m$. Our goal is to compare two designs that regulate the output signal y to some constant y_r (i.e. $\lim_{t\to\infty} y(t) = y_r$) using a controller Ξ with input $x_1, \ldots, x_n$ and output $u\colon \mathbb{R}_+ \to \mathbb{R}$. The first design we consider is a modification of the overparameterised adaptive backstepping controller introduced in [6], and e.g. in [8, Theorem 3.5]. It is straightforward to observe that there is no need to have one adaptation gain matrix Γ. We can introduce different matrices for the different estimates of θ to have more design freedom. This way we can get the following slightly modified adaptive backstepping controller for which the claims of [8, Theorem 3.5] are still true. We also add terms to be able to use the controller for tracking a reference signal $y_r(t)$. We denote this controller by $\Xi_{\mathrm{AB}}(\Gamma_1, \ldots, \Gamma_n, y_r)$

$$u = \alpha_n(x_1, \ldots, x_n, \hat{\theta}_1, \ldots, \hat{\theta}_n, y_{\mathrm{r}}^{(0)}, \ldots, y_{\mathrm{r}}^{(n-1)}) + y_{\mathrm{r}}^{(n)},$$

$$\dot{\hat{\theta}}_i = \Gamma_i \left(\phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j \right) z_i, \qquad \hat{\theta}_i(0) = 0,$$

$$z_i = x_i - y_{\mathrm{r}}^{(i-1)} - \alpha_{i-1}(x_1, \ldots, x_{i-1}, \hat{\theta}_1, \ldots, \hat{\theta}_{i-1}, y_{\mathrm{r}}^{(0)}, \ldots, y_{\mathrm{r}}^{(i-2)}),$$

$$\begin{aligned} \alpha_i = {} & -c_i z_i - z_{i-1} - \left(\phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j \right)^{\mathrm{T}} \hat{\theta}_i \\ & + \sum_{j=1}^{i-1} \left[\frac{\partial \alpha_{i-1}}{\partial x_j} x_{j+1} + \frac{\partial \alpha_{i-1}}{\partial y_{\mathrm{r}}^{(j-1)}} y_{\mathrm{r}}^{(j)} \right. \\ & \left. + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}_j} \Gamma_j \left(\phi_j - \sum_{k=1}^{j-1} \frac{\partial \alpha_{j-1}}{\partial x_k} \phi_k \right) z_j \right], \end{aligned} \tag{16}$$

where $c_i > 0, \hat{\theta}_i\colon \mathbb{R}_+ \to \mathbb{R}^m$ and $\Gamma_i = \Gamma_i^{\mathrm{T}} > 0$. The second controller we consider is the following version of the tuning function controller of [8], which we denote by $\Xi_{\mathrm{TF}}(\Gamma, y_r)$. The original design is summarized in Table 4.1 of [8]. To be able to compare the two designs we set the nonlinear damping term $\kappa = 0$:

$$u = \alpha_n(x_1, \ldots, x_n, \hat{\theta}, y_{\mathrm{r}}^{(0)}, \ldots, y_{\mathrm{r}}^{(n-1)}) + y_{\mathrm{r}}^{(n)},$$

$$\dot{\hat{\theta}} = \Gamma \sum_{i=1}^{n} \left(\phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j \right) z_i, \qquad \hat{\theta}(0) = 0,$$

$$z_i = x_i - y_{\mathrm{r}}^{(i-1)} - \alpha_{i-1}(x_1, \ldots, x_{i-1}, \hat{\theta}, y_{\mathrm{r}}^{(0)}, \ldots, y_{\mathrm{r}}^{(i-2)}),$$

$$\alpha_i = -c_i z_i - z_{i-1} - \left(\phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j\right)^{\mathrm{T}} \hat{\theta}$$
$$+ \sum_{j=1}^{i} \left[\frac{\partial \alpha_{i-1}}{\partial x_j} x_{j+1} + \frac{\partial \alpha_{i-1}}{\partial y_{\mathrm{r}}^{(j-1)}} y_{\mathrm{r}}^{(j)} \right.$$
$$\left. + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \Gamma \left(\phi_j - \sum_{k=1}^{j-1} \frac{\partial \alpha_{j-1}}{\partial x_k} \phi_k\right) z_j \right]$$
$$+ \sum_{k=2}^{i-1} \frac{\partial \alpha_{k-1}}{\partial \hat{\theta}} \Gamma \left(\phi_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \phi_j\right) z_k, \quad (17)$$

where $c_i > 0$, $\hat{\theta} : \mathbb{R}_+ \to \mathbb{R}^m$ and $\Gamma = \Gamma^{\mathrm{T}} > 0$. The cost function we use to measure the performance of the system, controller pair (Σ, Ξ) is

$$P(\Sigma, \Xi) = \|y\|_{L^\infty} + \lambda_1 \|x_2\|_{L^\infty} + \lambda_2 \|u\|_{L^\infty} \quad (18)$$

for some constants $\lambda_1, \lambda_2 > 0$. We fix the c_i's throughout the paper. Before stating the main result, we recall a definition. A controller is said to be ε-suboptimal if $P(\Sigma, \Xi(\mathbf{\Gamma}_\varepsilon)) \leqslant \inf_{\mathbf{\Gamma} > 0} P(\Sigma, \Xi(\mathbf{\Gamma})) + \varepsilon$, where $\mathbf{\Gamma} = \Gamma$ in the tuning function case, and $\mathbf{\Gamma} = (\Gamma_1, \ldots, \Gamma_n)$ in the adaptive backstepping case. We will prove the following theorem:

Theorem 2. *Suppose that the design objective is to regulate the output signal to a constant* $y_{\mathrm{r}} \in [a, b]$ *for some interval* $[a, b]$. *Then there exists a system* Σ *of degree two, an interval* $[a, b]$ *and constants* λ_1, λ_2, *such that the following hold. There is* $\varepsilon^* > 0$ *such that for all* $0 < \varepsilon < \varepsilon^*$, *if* $\Xi_{\mathrm{AB}}(\Gamma_1, \Gamma_2, b)$ *is an ε-suboptimal adaptive backstepping controller for* $y_{\mathrm{r}} = b$, *then with* $\Gamma = \Gamma_2$, $\Xi_{\mathrm{TF}}(\Gamma, b)$ *is an ε-suboptimal tuning function controller for* $y_{\mathrm{r}} = b$. *Moreover*

$$P(\Sigma, \Xi_{\mathrm{AB}}(\Gamma_1, \Gamma_2, b)) = P(\Sigma, \Xi_{\mathrm{TF}}(\Gamma, b)), \quad (19)$$

but

$$P(\Sigma, \Xi_{\mathrm{AB}}(\Gamma_1, \Gamma_2, a)) < P(\Sigma, \Xi_{\mathrm{TF}}(\Gamma, a)). \quad (20)$$

Remark. The theorem can be interpreted in the following manner. Suppose we do not know y_{r} a priori, but we expect $y_{\mathrm{r}} = b$ with high probability. We would therefore optimize for $y_{\mathrm{r}} = b$. If this expectation is incorrect and in fact $y_{\mathrm{r}} = a$ with high probability, then we would be in precisely the situation of the theorem.

By applying continuity arguments, we should therefore expect that we were in a situation where the controller will be asked to control to $y_{\mathrm{r}} = a$ with high probability, but with a small probability of $y_{\mathrm{r}} = b$, then the adaptive backstepping controller would give the lower expected cost. This type of scenario can be envisaged in robotics whereby a robot may be optimized for large movements, and occasionally asked to complete small movements.

4. Example

First we look at the design for $n = 2$ and $y_{\mathrm{r}}(t) = 0$. We will consider the following system $\Sigma(\phi_1, \phi_2, \theta, \mathbf{x}_0)$

$$\dot{x}_1 = x_2 + \theta \phi_1(x_1),$$
$$\dot{x}_2 = u + \theta \phi_2(x_1, x_2), \quad \mathbf{x}(0) = \mathbf{x}_0,$$
$$y = x_1, \quad (21)$$

where $x_1, x_2 : \mathbb{R}_+ \to \mathbb{R}$, $\theta \in \mathbb{R}_+$ and $\phi_i : \mathbb{R}^i \to \mathbb{R}$. The tuning function recursion using $c_1 = 1$ and $c_2 = 3$ results the following closed loop system $(\Sigma(\phi_1, \phi_2, \theta, \mathbf{x}_0), \Xi_{\mathrm{TF}}(\Gamma, 0))$:

$$\dot{x}_1 = x_2 + \theta \phi_1,$$
$$\dot{x}_2 = u + \theta \phi_2, \quad \mathbf{x}(0) = \mathbf{x}_0,$$
$$\dot{\hat{\theta}} = \Gamma(\phi_2 + \phi_1 + \phi_1' \hat{\theta} \phi_1)(x_2 + x_1 + \phi_1 \hat{\theta})$$
$$+ \Gamma \phi_1 x_1, \quad \hat{\theta}(0) = 0,$$
$$y = x_1,$$

$$u = -4x_2 - 4x_1 - 4\phi_1 \hat{\theta} - \phi_1' \hat{\theta} x_2$$
$$- \phi_1' \hat{\theta} \phi_1 \hat{\theta} - \phi_2 \hat{\theta} - \phi_1 \Gamma \phi_1 x_1$$
$$- \phi_1 \Gamma(\phi_2 + \phi_1 + \phi_1' \hat{\theta} \phi_1)(x_2 + x_1 + \phi_1 \hat{\theta}), \quad (22)$$

where $\Gamma > 0$ is the design constant. The modified adaptive backstepping recursion using $c_1 = 1$ and $c_2 = 3$ results the following closed loop system:

$(\Sigma(\phi_1, \phi_2, \theta, \mathbf{x}_0), \Xi_{\mathrm{AB}}(\Gamma_1, \Gamma_2, 0))$:

$$\dot{x}_1 = x_2 + \theta\phi_1,$$

$$\dot{x}_2 = u + \theta\phi_2, \quad \mathbf{x}(0) = \mathbf{x}_0,$$

$$\dot{\hat{\theta}}_1 = \Gamma_1 \phi_1 x_1, \quad \hat{\theta}_1(0) = 0,$$

$$\dot{\hat{\theta}}_2 = \Gamma_2(\phi_2 + \phi_1 + \phi_1' \hat{\theta}_1 \phi_1)(x_2 + x_1 + \phi_1 \hat{\theta}_1),$$

$$\hat{\theta}_2(0) = 0,$$

$$y = x_1,$$

$$u = -4x_2 - 4x_1 - 3\phi_1\hat{\theta}_1 - \phi_1\hat{\theta}_2 - \phi_1'\hat{\theta}_1 x_2 - \phi_1'\hat{\theta}_1\phi_1\hat{\theta}_2 - \phi_2\hat{\theta}_2 - \phi_1\Gamma_1\phi_1 x_1, \tag{23}$$

where $\Gamma_1, \Gamma_2 > 0$ are the design constants.

Let $b(x): \mathbb{R} \to \mathbb{R}$ be any twice differentiable "bump" function satisfying the following conditions: $b(x) = 0$ for $x \leqslant -\frac{1}{2}$ and $x \geqslant \frac{1}{2}$, $b(0) = 1$ and $0 \leqslant b(x) \leqslant 1$ for $-\frac{1}{2} \leqslant x \leqslant \frac{1}{2}$. Let $s(x): \mathbb{R} \to \mathbb{R}$ be any twice differentiable "step" function satisfying the following conditions: $s(x) = 0$ for $x \leqslant -\frac{1}{2}$, $s(x) = 1$ for $x \geqslant \frac{1}{2}$ and $0 \leqslant s(x) \leqslant 1$ for $-\frac{1}{2} \leqslant x \leqslant \frac{1}{2}$. We will consider systems $\Sigma(\phi_1, \phi_2, \theta, \mathbf{x}_0)$ defined by

$$\phi_1(x_1) = -Hb(x_1 - \tfrac{3}{2}) \tag{24}$$

for some $H > 0$,

$$\phi_2(x_1, x_2) = (-Ks(x_2 - x_1 + \tfrac{1}{2}) + Mx_1 b(x_2 - \tfrac{3}{2}))s(x_1 - 10) \tag{25}$$

for some $K, M > 0$, with initial conditions

$$\mathbf{x}_0 = (20, 100). \tag{26}$$

One can easily see that $\mathrm{supp}(\phi_1) = [1, 2]$, and that the support of ϕ_2 has two disjoint regions. We will refer to these regions, so let $\mathscr{R}_1 = [1, 2] \times \mathbb{R}$, and $\mathscr{R}_2$ and $\mathscr{R}_3$ be the two regions of the support of ϕ_2. $\mathscr{R}_2$ is $[9.5, \infty] \times [1, 2]$, where on the line $x_2 = \frac{3}{2}$ for $x_1 \geqslant 10.5$ $\phi_2(x_1, \frac{3}{2}) = Mx_1$. $\mathscr{R}_3$ is $\{(x_1, x_2)$: $x_2 \geqslant x_1 \geqslant 9.5\}$, where for $x_1 \geqslant 10.5$ and $x_2 - x_1 \geqslant 1$, $\phi_2(x_1, x_2) = -K$. Note that $\mathbf{x}_0 \in \mathscr{R}_3$, and that the supports of ϕ_1 and ϕ_2 are disjoint (i.e, that $\phi_1(x_1) \neq 0$ implies $\phi_2(x_1, x_2) = 0$, and $\phi_2(x_1, x_2) \neq 0$ implies $\phi_1(x_1) = 0$). The following lemma state the similarities of the $\mathbf{x}$-trajectories of the solutions. The state trajectories are illustrated in Figs. 1 and 2.

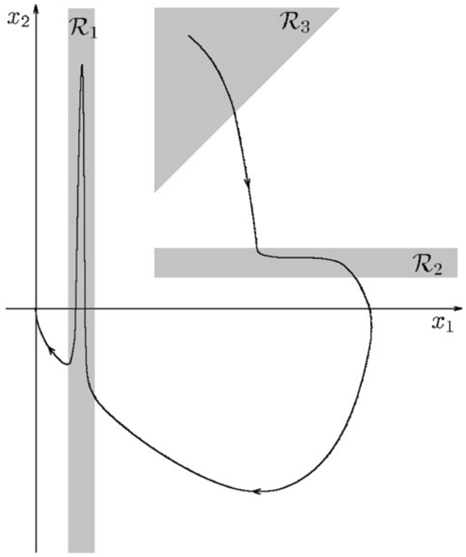

Fig. 1. $\mathbf{x}$-trajectory of the solution of the tuning function design.

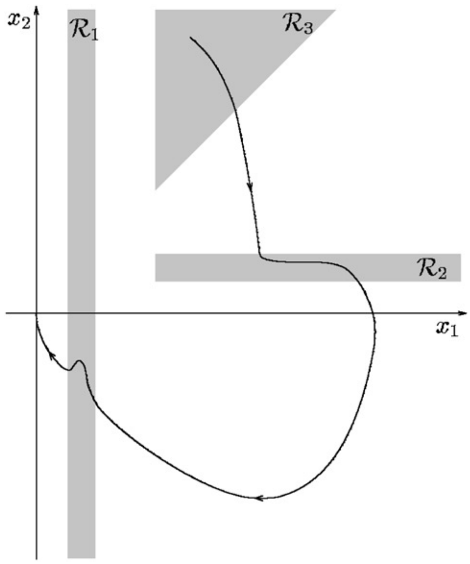

Fig. 2. $\mathbf{x}$-trajectory of the solution of the adaptive backstepping design.

Lemma 3. (i) *For both designs for any design constants the* $\mathbf{x}$*-trajectories start in* $\mathscr{R}_3$, *leave it through the* $x_1 = x_2$ *line, enter* $\mathscr{R}_2$ *through the* $x_2 = 2$ *line, leave it through the* $x_2 = 1$ *line, enter* $\mathscr{R}_1$ *with* $x_1 = 2, -4 \leqslant x_2 \leqslant -2$, *leave it with* $x_1 = 1$ *and then* *converge to the origin, not returning to any of these three regions.*

(ii) *If* $\Gamma = \Gamma_2$, *then* $\hat{\theta}_1 = 0$, $\hat{\theta} = \hat{\theta}_2$ *and the* **x**-*trajectories are the same for the two designs until they reach* $\mathscr{R}_1$.

(iii) *If* $\theta < 1$ *and* Γ *is small enough, then* $|\theta - \hat{\theta}| < 8/M$ *when the trajectory of the tuning function design enters* $\mathscr{R}_1$.

(iv) *If* $\theta < 1$ *and* Γ_2 *is small enough, then* $\hat{\theta}_1 = 0$ *and* $|\theta - \hat{\theta}_2| < 8/M$ *when the trajectory of the adaptive backstepping design enters* $\mathscr{R}_1$.

Proof. Look at the two systems outside $\mathscr{R}_1$. The tuning function design is

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = u + \theta\phi_2, \quad \mathbf{x}(0) = \mathbf{x}_0,$$

$$\dot{\hat{\theta}} = \Gamma(x_2 + x_1)\phi_2, \quad \hat{\theta}(0) = 0,$$

$$y = x_1,$$

$$u = -4x_2 - 4x_1 - \phi_2\hat{\theta}, \tag{27}$$

and the adaptive backstepping design is

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = u + \theta\phi_2, \quad \mathbf{x}(0) = \mathbf{x}_0,$$

$$\dot{\hat{\theta}}_1 = 0, \quad \hat{\theta}_1(0) = 0,$$

$$\dot{\hat{\theta}}_2 = \Gamma_2(x_2 + x_1)\phi_2, \quad \hat{\theta}_2(0) = 0,$$

$$y = x_1,$$

$$u = -4x_2 - 4x_1 - \phi_2\hat{\theta}_2, \tag{28}$$

(ii) is an immediate consequence of these forms.

(i) Consider (27). The **x**-trajectories start in $\mathscr{R}_3$ because $\mathbf{x}_0 \in \mathscr{R}_3$ and they leave $\mathscr{R}_3$, because $\lim_{t\to\infty} y(t) = 0$ along the solution (and since $y = x_1$ and $\mathscr{R}_3$ does not intersect the line $x_1 = 0$). In $\mathscr{R}_3, \dot{x}_1 = x_2 > 0$. Also, $\dot{\hat{\theta}} \leqslant 0$ and $\hat{\theta}(0) = 0$, so $\hat{\theta} \leqslant 0$. Since $\phi_2 \leqslant 0$ and $\theta \geqslant 0$, this implies that $u \leqslant 0$ and as a consequence $\dot{x}_2 \leqslant 0$ in $\mathscr{R}_3$. So the **x**-trajectory has decreasing x_2-coordinate and increasing x_1-coordinate in $\mathscr{R}_3$, so it leaves this region through the $x_1 = x_2$ line with $x_1 > x_1(0) = 20$. For $x_2 > 0$, $\dot{x}_1 > 0$, so until $x_2 > 0$ (i.e. until the **x**-trajectory reaches the $x_2 = 0$ line) the x_1-coordinate of the **x**-trajectory is increasing. Since along the solution $\lim_{t\to\infty} x_1(t) = 0$, this can only happen if the **x**-trajectory crosses $\mathscr{R}_2$ as stated in (i) (since the left border of $\mathscr{R}_2$ is at $x_1 = 9.5 < 20$), then reaches the $x_2 = 0$ axis with $x_1 > x_1(0) = 20$. Let T be the time, when $x_2(T) = 0$. Look now at the system

$$\dot{x}_1 = x_2, \quad x_1(T) = X > 20,$$

$$\dot{x}_2 = -4x_2 - 4x_1, \quad x_2(T) = 0, \tag{29}$$

which describes the behaviour of the **x**-trajectory of (27) for $x_2 \leqslant 0$, $x_1 \geqslant 2$ (i.e. after it has crossed the $x_2 = 0$ axis and before it reaches $\mathscr{R}_1$). The solution of this is $x_1(t + T) = Xe^{-2t}(2t + 1)$ and $x_2(t + T) = -4Xte^{-2t}$, from which we get that $\dot{x}_2(t + T) = 4Xe^{-2t}(2t - 1)$. This implies that the x_2-coordinate of the trajectory is decreasing for $t < \frac{1}{2}$ and increasing for $t > \frac{1}{2}$. At $t = \frac{1}{2}$, $x_1(T + \frac{1}{2}) = 2X/e > 2$, so the trajectory enters $\mathscr{R}_1$ at $t^* + T$ with $t^* > \frac{1}{2}$. Using this, and that $x_2(t^* + T) = -4t^* x_1(t^* + T)/(1 + 2t^*) = -4 + 4/(1 + 2t^*)$, we get that $-4 \leqslant x_2(t^* + T) \leqslant -2$, when the **x**-trajectory enters $\mathscr{R}_1$. By (ii), the behaviour of the **x**-trajectory of the adaptive backstepping design is the same until it reaches $\mathscr{R}_1$. Again, using that along the solution $\lim_{t\to\infty} x_1(t) = 0$, the **x**-trajectories leave $\mathscr{R}_1$ (through the $x_1 = 1$ line). From the form of the solution of the systems outside the supports of ϕ_1 and ϕ_2 it is easy to conclude that they do not return to these supports, and that they converge to the origin.

(iii) and (iv) are equivalent by (ii), so we prove (iii). ($\hat{\theta}_1 = 0$, because $\dot{\hat{\theta}}_1 = 0$ outside $\mathscr{R}_1$.) Since $\dot{\hat{\theta}} = 0$ between $\mathscr{R}_2$ and $\mathscr{R}_1$, we only have to show that $|\theta - \hat{\theta}| < 8/M$ when the **x**-trajectory leaves $\mathscr{R}_2$. Let T be the last time, when $x_2(T) = \frac{3}{2}$ (this T exists, since the **x**-trajectory enters $\mathscr{R}_2$ with $x_2 = 2$ and leaves it with $x_2 = 1$). Since this is the last time, $\dot{x}_2(T) \leqslant 0$. We already saw, that x_1 is increasing for $x_2 > 0$, so $x_1(T) > x_1(0) = 20$. Hence by the definition of ϕ_2, $\phi_2(x_1(T), x_2(T)) = Mx_1(T)$. Then

$$0 \geqslant \dot{x}_2(T) = -4x_1(T) - 4x_2(T) - (\hat{\theta}(T) - \theta)\phi_2$$

$$\geqslant -8x_1(T) - Mx_1(T)(\hat{\theta}(T) - \theta), \tag{30}$$

from which

$$\theta - \hat{\theta}(T) \leqslant 8/M \tag{31}$$

and since $\hat{\theta}$ is increasing in $\mathscr{R}_2$, the same inequality holds at the time, when the **x**-trajectory leaves $\mathscr{R}_2$. We also need to show, that $\hat{\theta}(T) - \theta \leqslant 8/M$ if Γ is small enough. Since $\dot{\hat{\theta}} = \Gamma(x_2 + x_1)\phi_2, \hat{\theta}$ increases in $\mathscr{R}_2$. If it does not reach θ, we are done by (31). After it reached $\theta, \dot{x}_2 \leqslant -4x_1 - 4x_2 < -4x_1$. But $|\dot{\hat{\theta}}| < 2\Gamma M x_1^2$ (since in $\mathscr{R}_2, 0 < x_2 < x_1$), so $|\dot{\hat{\theta}}/\dot{x}_2| < \Gamma M \|x_1\|_{L^\infty}/2$. We now give an upper bound on $\|x_1\|_{L^\infty}$. Look at the Lyapunov function

$$V = \tfrac{1}{2}x_1^2 + \tfrac{1}{2}(x_1 + x_2 + \phi_1\hat{\theta})^2 + \frac{1}{2\Gamma}(\theta - \hat{\theta})^2, \quad (32)$$

which has derivative $\dot{V} = -x_1^2 - 3(x_1 + x_2 + \phi_1\hat{\theta})^2 < 0$. This shows that $\|x_1\|_{L^\infty} < \sqrt{2V(0)} < v/\sqrt{\Gamma}$ for some constant v. (since we assumed that $\theta < 1$). Hence $|\dot{\hat{\theta}}/\dot{x}_2| < v\sqrt{\Gamma}M/2$. Let t_1 be the time when $\hat{\theta}(t_1) = \theta$, and let t_2 be the time when the **x**-trajectory leaves $\mathscr{R}_2$. Since $\dot{x}_2 < 0$ on $[t_1, t_2]$, we can conclude, that there is $t^* \in [t_1, t_2]$ such that

$$\left|\frac{\hat{\theta}(t_2) - \hat{\theta}(t_1)}{x_2(t_2) - x_2(t_1)}\right| = \left|\frac{\dot{\hat{\theta}}(t^*)}{\dot{x}_2(t^*)}\right| < \frac{v\sqrt{\Gamma}M}{2}, \quad (33)$$

which implies that $|\hat{\theta}(t_2) - \hat{\theta}(t_1)| < v\sqrt{\Gamma}M/2$ (since $|x_2(t_2) - x_2(t_1)| < 1$). If $\Gamma < 16^2/(v^2M^4)$, then $|\hat{\theta}(t_2) - \theta| = |\hat{\theta}(t_2) - \hat{\theta}(t_1)| < 8/M$ follows, which is what we wanted to prove. □

The following lemma states the difference between the **x**-trajectories of the solutions. Under certain circumstances and when θ and H (the height of ϕ_1), is varied, in $\mathscr{R}_1$ the adaptive backstepping trajectory remains uniformly bounded, on the other hand the tuning function design trajectory follows the curve $x_2 = -\theta\phi_1(x_1)$.

Lemma 4. (i) (*Tuning function design*) *There are* $E, \Gamma_0, M_0 > 0$, *such that if* $\Gamma < \Gamma_0$ *and* $M = M_0$, *then for all* $\theta, H > 0$ *and for all* t *such that* $1 \leqslant x_1(t) \leqslant 2, |(x_2 + \phi_1\theta)(t)| < E$.

(ii) (*Adaptive backstepping design*) *There are* $D, \Gamma_0, M_0 > 0$, *such that if* $\Gamma_1, \Gamma_2 < \Gamma_0$ *and* $M = M_0$, *then for all* $\theta, H > 0$ *and for all* t *such that* $1 \leqslant x_1(t) \leqslant 2, |x_2(t)| < D$.

Proof. (i) Let T be the first time, when the tuning function system reaches $\mathscr{R}_1$. By Lemma 3, $-4 \leqslant x_2(T) \leqslant -2$ and clearly $x_1(T) = 2$. Consider first the following modified system with initial condition as above:

$$\dot{x}_1 = x_2 + \phi_1\theta, \quad x_1(T) = 2,$$
$$\dot{x}_2 = -4x_2 - 4x_1 - 4\phi_1\theta - \phi_1'\theta(x_2 + \phi_1\theta),$$
$$-4 \leqslant x_2(T) = X \leqslant -2. \quad (34)$$

Equivalently, with $w_1 = x_1, w_2 = x_2 + \theta\phi_1$ we have

$$\dot{w}_1 = w_2, \quad w_1(T) = 2,$$
$$\dot{w}_2 = -4w_2 - 4w_1,$$
$$-4 \leqslant w_2(T) = W \leqslant -2. \quad (35)$$

The (w_1, w_2)-solution of this system depends continuously on W, and since the possible W values form a compact set, there are $E > 0$ and t_0 such that $w_1(T + t_0) < 1$ (i.e. the **w**-trajectory, which approaches the origin since

$$\begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}$$

is Hurwitz, leaves $\mathscr{R}_1$ for any possible W within t_0 time), and for all $T \leqslant t \leqslant t_0 + T$, $|x_2 + \theta\phi_1| = |w_2(t)| < E$. If Γ and $|\hat{\theta}(T) - \theta|$ are small enough, then the right hand side of the tuning function system and (34) are close to each other in an L^∞ sense, so applying Theorem 37 of Appendix C4 of [9] on $[T, T + t_0]$ (on the continuous dependence of the solution of a differential equation on the right-hand side of the equation) gives that the solutions of the two systems are arbitrarily close on $[T, T + t_0]$. By (iii) of Lemma 3, we know that with small enough Γ and large enough M, $|\hat{\theta}(T) - \theta|$ is arbitrarily small. Hence we can conclude that for small enough Γ and large enough M, $|x_2 + \theta\phi_1| \leqslant E$ in $\mathscr{R}_1$, which is what we wanted to prove.

(ii) Let now T be the time, when the adaptive backstepping system reaches $\mathscr{R}_1$. Using Theorem 37 of [9] again, we only have to show that for some $D > 0$ and $t_0 > 0$, $x_1(t_0) < 1$ and $|x_2(t)| < D$ for all $T \leqslant t \leqslant T + t_0$ are satisfied for the solution of the system

$$\dot{x}_1 = x_2 + \phi_1\theta, \quad x_1(T) = 2,$$
$$\dot{x}_2 = -4x_2 - 4x_1 - \phi_1\theta,$$
$$-4 \leqslant x_2(T) = X \leqslant -2. \quad (36)$$

Using the Lyapunov function $V = x_1^2/2 + (x_1 + x_2)^2/2$, which has derivative $\dot{V} = -x_1^2 - 3(x_1 + x_2)^2$ we can conclude that $x_1 \to 0$, hence the **x**-trajectory of this system indeed leaves $\mathscr{R}_1$ for any X. It also gives $t_0 > 0$, a bound on the time needed to cross $\mathscr{R}_1$. We show that $(x_2 + \phi_1\theta)(t) < 0$ for all $t \geqslant T$. Indeed, this is true for $t = T$. Suppose for a contradiction that t^* is the first time such that $(x_2 + \phi_1\theta)(t^*) = 0$. Then $\dot{x}_1(t^*) = 0$, hence the velocity vector of the **x**-trajectory of the solution at time t^* is vertical. Since the **x**-trajectory starts (at time T) below the graph of $x_2 = -\phi_1\theta$, when it reaches it the first time, the velocity vector cannot point down. Hence $\dot{x}_2(t^*) \geqslant 0$. On the other hand $\dot{x}_2(t^*) = -3x_2(t^*) - 4x_1(t^*) - (x_2 + \phi_1\theta)(t^*) < 0$. This contradiction show that $x_2 + \phi_1\theta < 0$. Now we compute $\mathrm{d}x_2/\mathrm{d}x_1$ as follows:

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = \frac{-4x_2 - 4x_1 - \phi_1\theta}{x_2 + \phi_1\theta} = -4 + \frac{3\phi_1\theta - 4x_1}{x_2 + \phi_1\theta} > -4 \tag{37}$$

and claim $x_2(t) \leqslant 2$ for all $t > T$ with $1 \leqslant x_1(t) \leqslant 2$. Indeed, if not, then for some $t^* > T$ with $1 \leqslant x_1(t^*) \leqslant 2$, $x_2(t^*) > 2$. Then for some $T < t^{**} < t^*$

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1}(t^{**}) = \frac{x_2(T) - x_2(t^*)}{x_1(T) - x_1(t^*)} < \frac{-2-2}{2 - x_1(t^*)} < -4,$$

which is a contradiction. On the other hand $x_2(t) > -4$, since the **x**-trajectory starts (at time T) above this line, hence if t^* is the first time when $x_2(t^*) = -4$, then $\dot{x}_2(t^*) \leqslant 0$. But $\dot{x}_2(t^*) = -4x_2(t^*) - 4x_1(t^*) - \phi_1\theta = 16 - 4x_1(t^*) - \phi_1\theta > 8$, which is a contradiction. Putting this together we get $-4 < x_2 < -2$ in $\mathscr{R}_2$, hence $D = 4$ completes the proof. □

Lemma 5. *For the Γ_0, M_0, E, D of Lemma 4, for any $\lambda_1, \lambda_2 > 0$, $G > D$ and for small enough ε there are K, θ, H such that $\theta H - E > G$ and if $\Xi_{\mathrm{TF}}(\Gamma, 10)$ is an ε-suboptimal tuning function controller designed to regulate the output to $y_\mathrm{r} = 10$, then $\Gamma < \Gamma_0$.*

Proof. We show first that since $\sup\{x_1 : (x_1, x_2) \in \mathscr{R}_1\} = 2 < 10$, if the controller is designed to regulate the output to $y_\mathrm{r} = 10$, then the **x**-trajectory of the solution does not enter $\mathscr{R}_1$. This follows easily from similar arguments we used to prove Lemma 3, noting that the only significant difference the y_r term makes is that after the **x**-trajectory enters the $x_2 \leqslant 0$, $x_1 \geqslant 0$ quaterplane at time T, it satisfies the differential equation system

$$\dot{x}_1 = x_2, \quad x_1(T) = X > 20,$$
$$\dot{x}_2 = -4x_2 - 4(x_1 - 10), \quad x_2(T) = 0, \tag{38}$$

which has solution $x_1(t+T) = X\mathrm{e}^{-2t}(2t+1) + 10 > 10$ and $x_2(t+T) = -4Xt\mathrm{e}^{-2t}$. We set $\phi_1 \equiv 0$. The consequence of the previous argument is that this change would not effect the solution, and hence the cost.

For a fixed constant γ and K let $\Gamma_K = 1/K^2, \theta_K = 1/K$, $M_K = M_0$, $H_K = \gamma K$, and use the notation $y^K = x_1^K$, $x_2^K, \hat{\theta}^K$ and u^K for the **x**-trajectory and control u of the closed loop system $(\Sigma, \Xi_{\mathrm{TF}}(\Gamma_K, 10))$. If γ is large enough, then $\theta_K H_K - E = \gamma - E > G > D$. First we show that there is a constant α, such that if K is large enough, then

$$P(\Sigma, \Xi_{\mathrm{TF}}(\Gamma_K, 10)) = \|y^K\|_{L^\infty} + \lambda_1\|x_2\|_{L^\infty} + \lambda_2\|u^K\|_{L^\infty} < \alpha. \tag{39}$$

The Lyapunov function

$$V = \tfrac{1}{2}(x_1 - 10)^2 + \tfrac{1}{2}(x_1 - 10 + x_2)^2 + \frac{1}{2\Gamma}(\theta - \hat{\theta})^2 \tag{40}$$

has derivative $\dot{V} = -(x_1 - 10)^2 - 3(x_1 - 10 + x_2)^2 < 0$ along the solution, hence for every $t \geqslant 0$, $|x_1^K(t) - 10| \leqslant \sqrt{2V(0)}$, $|(x_1(t) - 10 + x_2(t)| \leqslant \sqrt{2V(0)}$ and $(\theta - \hat{\theta}(t))^2 \leqslant 2\Gamma V(0) = 2V(0)/K^2$. Since $V(0)$ is constant, this means that there is a constant α_1, such that for every $t \geqslant 0$, $|x_1^K(t)| < \alpha_1$, $|\hat{\theta}^K(t)| < \alpha_1/K$. Since $x_2^K(t) \leqslant x_2^K(0)$, and for $x_2^K < 0$ we know the explicit form of the solution: $|x_2^K(t+T)| = 4Xt\mathrm{e}^{-2t} < 4\alpha_1 t\mathrm{e}^{-2t}$, there is a constant α_2, such that $|x_2^K(t)| < \alpha_2$. By the choice of ϕ_2 there is a constant α_3 such that $|\phi_2| \leqslant M_K|x_1^K| < \alpha_3 K$. Since $\phi_1 \equiv 0$, $u^K(t) = (-4x_2^K - 4(x_1^K - 10) - \phi_2\hat{\theta})(t)$, hence $|u^K|$ is indeed bounded by a constant, so there is a constant α satisfying (39).

To complete the proof, we now show that there is a K such that for any $\Gamma > \Gamma_0$, if u is the control of the system with $\theta_K = 1/K$, $M_K = M_0$ and $H_K = \gamma K$, then $\lambda_2\|u\|_{L^\infty(\mathscr{R}_3)} > \alpha$. This shows that for this setup the choice $\Gamma > \Gamma_0$ does not give ε-suboptimal tuning function controller for any small enough ε, which is what we wanted to show. Let t_K be the time when the $\mathbf{x}$-trajectory of the solution leaves the region $\mathscr{R}'_3 = \{(x_1,x_2)\colon x_2 - 1 \geqslant x_1 \geqslant 10.5\}$, where $\phi_2 = -K$. Suppose that there is a $t_0 > 0$ such that for all $K^* > 0$ there is $K > K^*$ such that $t_K > t_0$. In $\mathscr{R}'_3$, $x_1,x_2 > 10$ and $\phi_2 \equiv -K$, hence $\dot{\hat{\theta}} = \Gamma(x_2 + x_1)\phi_2 < -20\Gamma_0 K$. Therefore $\hat{\theta}(t_K) < -20t_0\Gamma_0 K$, hence $|u(t_K)| = 4x_2^K + 4(x_1^K - 10) + \phi_2\hat{\theta} > 20t_0\Gamma_0 K^2 > \alpha/\lambda_2$ if K is large enough. If on other hand $\lim_{K\to\infty} t_K = 0$, then $\lim_{K\to\infty}|x_1(t_K) - x_1(0)| = 0$, since $|\dot{x}_1| = |x_2| \leqslant x_2(0)$ on $\mathscr{R}'_3$. Since $x_1(t_K) = x_2(t_K) - 1$, this means, that $\lim_{K\to\infty}|x_2(t_K) - x_2(0)| = x_2(0) - x_1(0) - 1 > 0$, hence $\lim_{K\to\infty}(\sup_{t\in[0,t_K]}|\dot{x}_2|) = \infty$. But on $\mathscr{R}'_3$, $\dot{x}_2 = u + \theta_K\phi_2$, and $|\theta_K\phi_2| = 1$, this means that $\lim_{K\to\infty}(\sup|u|) = \infty$, which completes the proof. □

Proof of Theorem 2. Let $[a,b] = [0,10]$. The claim that if $\Xi_{AB}(\Gamma_1,\Gamma_2,10)$ is an ε-suboptimal adaptive backstepping controller, then for $\Gamma = \Gamma_2, \Xi_{TF}(\Gamma,10)$ is an ε-suboptimal tuning function controller (and that they have the same costs) follows from (ii) of Lemma 3, and the fact that in this case the $\mathbf{x}$-trajectories do not reach $\mathscr{R}_1$. According to Lemma 5, for any $\lambda_1,\lambda_2 > 0$ there is a system Σ, such that for small enough ε if $\Xi_{TF}(\Gamma,10)$ is an ε-suboptimal tuning function controller, then $\Gamma < \Gamma_0$. In this case we can apply Lemma 4 for the designs $(\Sigma,\Xi_{TF}(\Gamma,0))$ and $(\Sigma,\Xi_{AB}(\Gamma,\Gamma,0))$ to get a significant difference between $\|x_2^{TF}\|_{L^\infty}$ and $\|x_2^{AB}\|_{L^\infty}$. The solutions for the two designs agree until they reach $\mathscr{R}_1$, let T be the time, when they arrive there. Then

$$\|x_1^{AB}\|_{L^\infty[0,T]} + \lambda_1\|x_2^{AB}\|_{L^\infty[0,T]} + \lambda_2\|u^{AB}\|_{L^\infty[0,T]}$$
$$= \|x_1^{TF}\|_{L^\infty[0,T]} + \lambda_1\|x_2^{TF}\|_{L^\infty[0,T]} + \lambda_2\|u^{TF}\|_{L^\infty[0,T]}. \quad (41)$$

Moreover, $\|x_2^{TF}\|_{L^\infty(\mathscr{R}_1)} > \theta H - E > G$ is much bigger than these if γ of the construction of Lemma 5 is large enough, since $G > D$ can be chosen arbitrarily. This means, that for large enough γ it is enough to establish that the adaptive backstepping cost is less than the tuning function cost in $\mathscr{R}_1$.

$$\|x_1^{AB}\|_{L^\infty(\mathscr{R}_1)} = \|x_1^{TF}\|_{L^\infty(\mathscr{R}_1)} = 2, \quad (42)$$

$$\|x_2^{TF}\|_{L^\infty(\mathscr{R}_1)} > \theta H - E > G > D > \|x_2^{AB}\|_{L^\infty(\mathscr{R}_1)}. \quad (43)$$

We show now that there is a constant μ, such that

$$\|x_2^{TF}\|_{L^\infty(\mathscr{R}_1)} > \mu\|u^{AB}\|_{L^\infty(\mathscr{R}_1)}. \quad (44)$$

This is enough, since then by appropriately choosing λ_1 and λ_2 will result

$$\lambda_1\|x_2^{TF}\|_{L^\infty(\mathscr{R}_1)} > \lambda_1\|x_2^{AB}\|_{L^\infty(\mathscr{R}_1)} + \lambda_2\|u^{AB}\|_{L^\infty(\mathscr{R}_1)}, \quad (45)$$

which is enough to conclude that the tuning function cost is bigger. To prove (44) look at the control of the adaptive backstepping system in $\mathscr{R}_1$. According to Lemma 4, there is $\varepsilon^* > 0$ such that $\|\hat{\theta}_2 - \theta\| < \varepsilon^*$ and $\|\hat{\theta}_1\| < \varepsilon^*$ in $\mathscr{R}_1$. Then

$$\begin{aligned}|u^{AB}| &\leqslant 4|x_2^{AB}| + 4|x_1^{AB}| + 3|\phi_1\hat{\theta}_1| + |\phi_1\hat{\theta}_2| \\ &\quad + |\phi_1'\hat{\theta}_1 x_2^{AB}| + |\phi_1'\hat{\theta}_1\phi_1\hat{\theta}_2| + |\phi_1\Gamma_1\phi_1 x_1^{AB}| \\ &\leqslant 4D + 8 + 3H\varepsilon^* + H(\theta + \varepsilon^*) + |\phi_1'|D\varepsilon^* \\ &\quad + |\phi_1'|H(\theta + \varepsilon^*)\varepsilon^* + 2H^2\Gamma_1 \\ &\leqslant \mu' H\theta, \end{aligned} \quad (46)$$

for some $\mu' > 0$ if H and θ are large enough and Γ_1 is small enough. This completes the proof, since $\|x_2^{TF}\|_{L^\infty(\mathscr{R}_1)} > H\theta/2$. □

5. Summary

By means of two examples, we have shown that overparameterisation can be beneficial in adaptive control. This fully motivates a more general study into the whole question of when and when not to overparameterise, although as noted in the Introduction, this is likely to be a challenging task.

References

[1] P. Bartlett, The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network, IEEE Trans. Inform. Theory 44 (2) (1998) 525–536.

[2] M. French, An analytical comparison between the weighted LQ performance of a robust and an adaptive backstepping design, IEEE Trans. Automat. Control 47 (4) (2002) 670–675.
[3] M. French, Cs. Szepesvári, E. Rogers, Uncertainty, performance and model dependency in approximate adaptive nonlinear control, IEEE Trans. Automat. Control 45 (2) (2000) 353–358.
[4] M. French, Cs. Szepesvári, E. Rogers, An asymptotic scaling analysis of LQ performance for an approximate adaptive control design, Math. Control, Signals Systems 15 (2) (2002) 145–176.
[5] Z.-P. Jiang, I. Mareels, Robust nonlinear integral control, IEEE Trans. Automat. Control 15 (8) (2001) 1336–1342.
[6] I. Kanellakopoulos, P.V. Kokotović, A.S. Morse, Systematic design of adaptive controllers for feedback linearisable systems, IEEE Trans. Automat. Control 36 (11) (1991) 1241–1253.
[7] M. Krstić, I. Kanellakopoulos, P.V. Kokotović, Adaptive nonlinear control without overparameterization, Systems and Control Letters 19 (1992) 177–185.
[8] M. Krstić, I. Kanellakopoulos, P.V. Kokotović, Nonlinear and Adaptive Control Design, 1st Edition, Wiley, New York, 1995.
[9] E.D. Sontag, Mathematical Control Theory—Finite Dimensional Systems, Texts in Applied Mathematics, 1st Edition, Springer, New York, 1990.