PERGAMON

Automatica 38 (2002) 683–693

automatica

www.elsevier.com/locate/automatica

Brief Paper

LQ performance bounds for adaptive output feedback controllers for functionally uncertain nonlinear systems ☆

M. French[a,*], Cs. Szepesvári[b,1], E. Rogers[a]

[a] Department of Electronics and Computer Science, University of Southampton, Southampton S017 1BJ, UK
[b] Research Group on Artificial Intelligence, Hungarian Academy of Sciences–JATE, Szeged, Hungary

Received 12 August 1998; revised 3 April 2001; received in final form 2 August 2001

Abstract

We consider functionally uncertain systems which can be written in an output feedback form, where the nonlinearities are functions of the output only. The uncertainty is described by a weighted L^2 norm about a nominal system, and an approximate adaptive design is given which ensures output practical stability. The main result requires knowledge of the weighted L^2 uncertainty level. An upper bound on the LQ performance of the output transient and the control input is derived, where the cost penalises the output transient and the control effort on the time interval where the output lies outside the prescribed neighbourhood of zero to which we achieve convergence. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Adaptive control; Function approximation; Performance

1. Introduction

Adaptive output feedback designs for systems admitting a output feedback form and parametric uncertainty have been available from Marino and Tomei (1993a), see also e.g. Teel (1993) and Krstić and Kokotović (1996). The purpose of this paper is to generalise these adaptive designs to a case of non-parametric uncertainty. Importantly, we also bound an LQ-type cost functional which penalises both the output transient and the control effort. The approach taken is closely related to the neural network literature, where a neural network is used as an adaptive model to approximate a functional uncertainty, and the scheme is made robust to the 'disturbance' which arises from a residual approximation error, see e.g. Sanner and Slotine (1992). Essentially therefore, we have to give a robust adaptive output feedback design. Recently, a number of robust adaptive designs have been proposed for output feedback systems, see e.g. Ikhouane and Krstić (1998), Marino and Tomei (1997) and Jiang (1998) and the references therein. In contrast to these approaches in this paper we utilise a dead-zone modification to the nominal adaptive law; this is ideally suited to our problem since a uniform bound on the 'disturbance' (approximation error) terms can be obtained, and hence we can achieve stronger asymptotic behaviour in the presence of disturbances (i.e. practical asymptotic stabilisation, in contrast to simply uniform ultimate boundedness).[2] Thereafter, the trade-offs between various different robust modifications have been elucidated previously, e.g. Narendra and Annaswamy (1990).

In this paper, the only requirement on the adaptive model is that it is linearly parameterised, so we can apply the results in the paper to any of the rich variety of approximation schemas (polynomials, radial basis functions, splines, single-layer neural-nets, Fourier series,

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Frank L. Lewis under the direction of Editor Frank L. Lewis.

* Corresponding author.
E-mail addresses: mcf@ecs.soton.ac.uk, (M. French), szepes@mindmaker.hu (Cs. Szepesvári), etar@ecs.soton.ac.uk (E. Rogers).
[1] Present address: Mindmaker Ltd., Budapest, 1121 Konkoly Th. M. u. 29–33. Hungary.

[2] Although the dead-zone design proposed here does not have an ideal asymptotic behaviour when the disturbance is not present (contrast to Marino and Tomei (1997)), in the situation considered here the disturbance is generically present, and so this idealised property is not of interest.

PII: S0005-1098(01)00218-7

wavelets, etc). We will, however, take careful consideration of the restrictions that a canonical approximation theory places on the approximation properties of a model, for example, we can expect uniform approximation only over compacta with a finite dimensional model; global approximation requires infinite dimensional models. In particular, it will turn out that the (weighted) l^2 norm of the adaptive model's 'ideal' parameters is related to the transient behaviour of the output signal; additionally it will typically appear that as the model's domain or resolution is required to increase, so does this norm. The model's domain will be required to cover the output's range, and hence this coupling must be handled carefully. This is a major motivation for the introduction of the *functional* uncertainty models considered in this paper. In contrast to e.g. Jiang (1998), the uncertainty is described not by pointwise bounds, but by spatial L^2 bounds: this would appear to be the natural description of uncertainty when using approximate adaptive designs, as in the strict feedback and matched cases considered in French, Szepesvari, and Rogers (2000) and French and Rogers (1998).

Our results differ from other results using approximate models as we give completely constructive results where no parameters are left to be tuned as is typical in many neural network papers; it is necessary to give careful attention to the structure of the approximation errors, dimension of the approximating model and the transient behaviour of the system. The result differ from related work in approximate adaptive control where a robust term is added to the control law to control the system in the large (such as in Sanner & Slotine, 1992): in our results the system is controlled purely by the adaptive means, the only robust terms in the control law are small and are used solely to control small disturbances. It is, however, straightforward to introduce extra damping terms, as in Yao and Tomizuka (1997), and describe the uncertainty by mixed L^2/L^∞ uncertainty models as in French (1998), to achieve global results under similar assumptions.

The main contribution of this paper is to give a constructive bound on LQ costs for these adaptive designs. This result coupled with French (1998), French et al. (2000) and French and Rogers (1998) presents the first constructive bounds on a-priori determined performance costs in the adaptive control literature. It contrasts to the inverse optimal designs of Li and Krstić (1997), where optimal controllers are derived w.r.t. to a (meaningful) but not a-priori determined cost functional. The results of Li and Krstić (1997) are thus hard to interpret from a performance perspective, in particular it is hard to compare different adaptive models using these methods; however, it should of course be realised that the motivation for inverse optimal results is not integral performance directly, rather, those results were motivated by robustness considerations.

1.1. Notation and approximation theoretic background

$\mathscr{W}$ denotes a parameter space, $\mathscr{X}$, $\mathscr{Z}$ denote the state space and error system state space, respectively, $\mathscr{O}$ denotes the output space; all are taken to be Euclidean spaces. Ω denotes the approximation domain, it is a subset of $\mathscr{O}=\mathbb{R}$. $L^2(\Omega)$ denotes the standard Lebesgue space over Ω, and the weighted inner product space $L^2(\Omega;w)$ has the inner product: $\langle f,g\rangle_{L^2(\Omega;w)}=\int_\Omega f(x)g(x)w(x)\,\mathrm{d}x$, where w is a measurable function. $C(\Omega)$ is the normed space of continuous functions on Ω, with the uniform norm. $C^k(\mathbb{R}^n,\mathbb{R}^m)$ is the space of k times differentiable functions mapping $\mathbb{R}^n$ to $\mathbb{R}^m$. The unit matrix will be denoted by I. If the eigenvalues of a matrix R are $\lambda_1,\dots,\lambda_n$, then $\bar\lambda(R)$, $\underline{\lambda}(R)$ are defined to be $\max_{1\leqslant i\leqslant n}|\lambda_i|$, $\min_{1\leqslant i\leqslant n}|\lambda_i|$, respectively. Norms for various spaces $\mathscr{F}$ will be denoted as $\|\cdot\|_{\mathscr{F}}$, for convenience $\|\cdot\|$ will mean $\|\cdot\|_2$ over the appropriate space, and if R is a positive-definite matrix, $\|x\|_R$ will denote the weighted norm $\sqrt{|x^{\mathrm{T}}Rx|}$ of vector x. $\partial\Omega$ denotes the topological boundary of $\Omega\subset\mathscr{X}$, Ω° the interior and $\bar\Omega$ denotes the closure. Ω^{c} denotes the complement of Ω. $m(\Omega)$ denotes the Lebesgue measure of Ω. e_i denotes the ith basis vector $(0,\dots,0,1,0,\dots,0)^{\mathrm{T}}$. For a function $V:\mathscr{X}\to\mathbb{R}$, $L(V,r)$ denotes the level set $\{x\in\mathscr{X}\colon V(x)\leqslant r\}$. $P_i:\mathbb{R}^n\to\mathbb{R}^i$ denotes the projection: $P_i(x_1,\dots,x_n)=(x_1,\dots,x_i)$. A system is denoted by Σ, a controller by Ξ, a system, controller interconnection is denoted by (Σ,Ξ), it is said to be well posed on $[0,T]$, if over the time interval $[0,T]$ all outputs and internal signals exist, and are bounded. Solutions to discontinuous differential equations are interpreted in the sense of Fillipov (1998).

We will be concerned throughout this paper by linear approximants of the form $\theta^{\mathrm{T}}\phi:\Omega\to\mathbb{R}$ $\theta\in\mathscr{W}=\mathbb{R}^p$, $\Omega\subset\mathscr{O}=\mathbb{R}$. $\mathscr{W}$ will be called the weight or parameter space and Ω is the approximation domain. As we are interested in multi-output approximation, we introduce the following notation. We define a model Φ as $\Phi=(\phi_1,\phi_2,\dots,\phi_n)^{\mathrm{T}}$, where $\phi_i:\Omega\to\mathscr{W}_i$, $\phi_i=(\varphi_{i1},\varphi_{i2},\dots,\varphi_{im_i})^{\mathrm{T}}$ where $\mathscr{W}_i=\mathbb{R}^{m_i}$ and where for convenience we assume that $\varphi_{ij}\in C^{n+3}(\mathscr{O},\mathbb{R})$. Note that for clarity, we are using a hierarchy of notation Φ,ϕ,φ to denote the model, model component, and basis function, respectively. Approximation theory typically considers families of such approximants, which we formalize as follows: Let $K(\Omega)\subset\mathscr{F}(\Omega)\subset C(\Omega)$. A $K(\Omega)$ dense linear model resolution schema is a sequence of the form: $\{\Phi^m\}_{m\in\mathbb{N}}$ where $\sup_{f\in K(\Omega)}\inf_{\theta_{f_i}\in\mathscr{W}_m}\|f_i-\theta_{f_i}^{\mathrm{T}}\phi_i^m\|_{\mathscr{F}(\Omega)}\to 0$ as $m\to\infty$ for $1\leqslant i\leqslant n$. The size of a model is the dimension of the weight space, $m=\sum_{i=1}^n|\mathscr{W}_i|=\sum_{i=1}^n m_i<\infty$. Typical examples of linear resolution schemas would be polynomials of increasing degree (Rivlin, 1969); mesh-based approximants such as splines on decreasing mesh sizes (Rivlin, 1969); or wavelets (Daubechies, 1992). Note that if $\mathscr{F}(\Omega)=C(\Omega)$ then the required density property can be achieved only over compact domains, unless $K(\Omega)$ is required to be excessively regular. A key problem is to determine bounds on the size of the model to achieve a specified approximation error ε. In order to do this we need to introduce further assumptions concerning the smoothness of the function to be approximated. The main 'meta-theorem' of approximation theory can be stated as follows (where typical smoothness classes for compact domains would be Lipschitz constraints, or bounds in Sobolev spaces, and a well-known example of a realisation of this meta-theorem is Jackson's Theorem (Rivlin, 1969):

Theorem 1.1. *Suppose a $K(\Omega) \subset \mathscr{F}(\Omega)$ dense linear model resolution schema is given. Let $\varepsilon > 0$ be given, and suppose $K(\Omega)$ is a smoothness class of functions $f : \Omega \to \mathbb{R}$. Then $\forall m > M(K(\Omega), \varepsilon)$ (where $M(K, \varepsilon)$ is given constructively) $\forall f \in K(\Omega)$, $\exists \theta \in \mathscr{W}_m$ such that $\|f - \theta^{\mathrm{T}}\phi^m\|_{\mathscr{F}(\Omega)} < \varepsilon$.*

2. Problem domain

Necessary and sufficient geometric conditions are known (Krstić, Kanellakopoulos, & Kokotović, 1995) for the existence of a global diffeomorphism ($S : \mathscr{X} \to \mathscr{X}$) which transforms an affine system Σ' into an output feedback normal form Σ:

$$\Sigma' : \dot{s} = k(s) + g(s)u, \quad y = h(s),$$

$$\Sigma_{\{f\}} : \dot{x}_i = x_{i+1} + f_i(y), \quad 1 \leqslant i \leqslant n - 1, \tag{1}$$

$$\dot{x}_n = u + f_n(y), \quad y = x_1,$$

whilst leaving y invariant. We will consider such systems and assume that the only signal which is available for measurement is $y \in \mathscr{O} = \mathbb{R}$; in particular the state vectors $x = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathscr{X} = \mathbb{R}^n$ and $s \in \mathbb{R}^n$ are assumed to be unavailable for measurement. f denotes the uncertain function $f = (f_1, \ldots, f_n)^{\mathrm{T}}$, and $f^0 = (f_1^0, \ldots, f_n^0)^{\mathrm{T}}$ represents the (known) nominal system. We assume that $f, f^0 \in C^{n+3}(\mathscr{O}, \mathscr{X}), C(\mathscr{O}, \mathscr{X})$, respectively. The control task is to stabilise y to a small neighbourhood of 0, $[-\sqrt{2}\eta, +\sqrt{2}\eta] = \Omega_0$, whilst keeping all signals bounded. It should also be observed that, for brevity, as in Teel (1993), we are considering a simpler normal form than Marino and Tomei (1993a) and Krstić and Kokotović (1996), as we are assuming that the system is of relative degree $\rho = n$. However, the designs given here can be extended to the case $\rho < n$ when the minimum phase assumptions of Marino and Tomei (1993a) and Krstić and Kokotović (1996) hold. The basic uncertainty set we consider is: $\Delta = \Delta(L^2(\Omega; w), f^0, \delta) = \{f \in C^{n+3}(\mathscr{O}, \mathscr{X}) \mid f - f^0 \in K, \|f_i - f_i^0\|_{L^2(\Omega; w)} \leqslant \delta_i, 1 \leqslant i \leqslant n\}$, where $\Omega \subset \mathscr{O}$ is typically compact. K denotes an approximation theoretic smoothness class, see above for details. It is important to observe that the spatial L^2 nature of the uncertainty model is very different to uncertainty models utilised to date in nonlinear control. Robust backstepping designs, e.g. Marino and Tomei (1993b), and older, simpler designs e.g. Corless and Leitmann (1981) utilise pointwise bounds on the nonlinearity. Similarly, the adaptive design of Jiang (1998) also utilises a pointwise bound (of unknown magnitude). To some extent these L^2 uncertainty models are well-tailored to identification data: often models can be obtained with MSE or l^2/L^2 error bounds. In contrast, it is hard to obtain good pointwise error bounds from identification data.

Even when the system is modelled physically, spatial integral descriptions of uncertainty can be appropriate. For example, consider the motion of a particle moving on the surface of an a-priori unknown 1D hilly landscape given by the function $l : \mathbb{R} \to \mathbb{R}$. Assuming l is smooth, let $s(t)$ denote the arc-length from the origin at time t, which is the measured output ($y = s$). The control is applied by a force tangential to the landscape, with the actuator dynamics modelled as a single integrator. Applying Newton's law, we thus have a system of the form of Eq. (1) with $n = 3$, $f_1 = f_3 = 0$ and $f_2(y) = -mg\cos(\tan^{-1}(\partial l/\partial y|_y))$; since, there is a smooth bijection between the arc-length position $s(t)$ and the horizontal position $y(t)$, f_2 is a function of the output only. Now consider an uncertainty set which comprises of a landscape of single 'bumps' at unknown locations, e.g. $\Delta_k = \{f = (f_1, f_2, f_3)^{\mathrm{T}} \mid f_1 = f_3 = 0,\ f_2(y) = -mg\cos(\tan^{-1}(\partial l/\partial y|_y)),\ l = \exp(-k\|x - z\|^2),\ z \in \mathbb{R}\}$. The steepness of the 'bumps' is indexed by $k \geqslant 0$. It can now be easily seen that as $k \to \infty$, $\|\Delta_k\|_{L^2} \to 0$, whilst $\|\Delta_k\|_{L^\infty} \to mg$. The L^2 description is more appropriate as it can capture the essence the uncertainty is small, but spatially uncertain, whereas the L^∞ description cannot reflect the spatial uncertainty. The essence of this example is that pointwise measures for this type of uncertainty can lead to descriptions which are needlessly conservative. Consequently, a control design based on the less conservative uncertainty model can be reasonably expected to have superior performance.

Performance will be measured in a worst-case LQ manner, penalising both the output and the control: $\mathscr{P} = \mathscr{P}(c_1, k, \eta) = \sup_{f \in \Delta} \sup_{\mathrm{solns}(\Sigma_f, \Xi)} \int_{T_2} c_1 y^2(t) + ku^2(t)\,\mathrm{d}t$, where $T_2 = \{t \geqslant 0 \mid y(t) \notin \Omega_0 = [-\sqrt{2}\eta, \sqrt{2}\eta]\}$.

3. Adaptive control design, stability and performance

The adaptive control methodology is based on that of Marino and Tomei (1993a) and Krstić and Kokotović (1996) and robust backstepping (Freeman & Kokotović, 1996; Krstić et al., 1995). We write the system in the form

$$\dot{x}_i = x_{i+1} + f_i^0(y) + \theta_i^{\mathrm{T}}\phi_i(y) + d_i^f(y), \quad 1 \leqslant i \leqslant n-1,$$

$$\dot{x}_n = u + f_n^0(y) + \theta_n^{\mathrm{T}}\phi_n(y) + d_n^f(y), \quad y = x_1, \tag{2}$$

$$d_i^f(y) = f_i(y) - f_i^0(y) - \theta_i^{\mathrm{T}}\phi_i(y).$$

When f is clear from the context, we often write d_i for d_i^f. As in French and Rogers (1998) it should be noted that even in the absence of disturbance terms, this system differs from the standard parametric normal form (Krstić et al., 1995; Marino & Tomei, 1993a) (although it is not more general) because the vectors $\theta_1, \ldots, \theta_n$ are distinct. d^f defines the vector $d^f = (d_1^f, \ldots, d_n^f)^{\mathrm{T}}$.

However, to define filters for the system, it is convenient to write the system in a similar form to Krstić and Kokotović (1996). We reparameterise the system as follows. Define $\theta' = (\theta_1', \theta_2', \ldots, \theta_m')^{\mathrm{T}} = (\theta_1^{\mathrm{T}} \mid \theta_2^{\mathrm{T}} \mid \cdots \mid \theta_n^{\mathrm{T}})^{\mathrm{T}} \in \mathscr{W} = \mathbb{R}^m$, $\phi_1' = (\phi_1^{\mathrm{T}} \mid 0 \mid \cdots \mid 0)^{\mathrm{T}}$, $\phi_2' = (0 \mid \phi_2^{\mathrm{T}} \mid \cdots \mid 0)^{\mathrm{T}}, \ldots, \phi_n' = (0 \mid \cdots \mid 0 \mid \phi_n^{\mathrm{T}})^{\mathrm{T}}$. $\Phi_j' = (\phi_{1j}', \phi_{2j}', \ldots, \phi_{nj}')^{\mathrm{T}}$. Note that by definition, $\Phi_j' : \mathscr{O} \to \mathscr{X}$ only has one non-zero entry. Then we can rewrite the system in the alternative form: $\dot{x} = Ax + e_n u + f^0(y) + \sum_{j=1}^m \theta_j' \Phi_j'(y) + d^f(y)$, where A is the matrix:

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

We now follow the definition of the filters for state estimation in Krstić and Kokotović (1996): A gain vector $\gamma = (\gamma_1, \ldots, \gamma_n)^{\mathrm{T}} \in \mathbb{R}^n$ is chosen such that $A_0 = A - e_1\gamma^{\mathrm{T}}$ is Hurwitz (such γ is said to be admissible). The nominal (ω), model (ζ) and control (v) filters are defined as follows:

$$\dot{\omega} = A_0\omega + \gamma y + f^0(y), \quad \omega(0) = 0, \ \omega \in \mathscr{X}, \tag{3}$$

$$\dot{\zeta}_j = A_0\zeta_j + \Phi_j'(y), \quad \zeta_j(0) = 0, \ \zeta_j \in \mathscr{X}, \ 1 \leqslant j \leqslant m, \tag{4}$$

$$\dot{v} = A_0 v + e_n u, \quad v(0) = 0, \ v \in \mathscr{X}. \tag{5}$$

For convenience, ζ denotes the vector $\zeta = (\zeta_1, \ldots, \zeta_m)^{\mathrm{T}}$. The state estimation error $\varepsilon \in \mathscr{X}$ is defined to be:

$$\varepsilon = x - \left(\omega + \sum_{j=1}^m \theta_j'\zeta_j + v\right). \tag{6}$$

An error system is recursively defined as the vector $z = (z_1, \ldots, z_n)^{\mathrm{T}} \in \mathscr{Z} = \mathbb{R}^n$: $z_1 = y$, $z_i = v_i - \alpha_{i-1}$, $2 \leqslant i \leqslant n$ where $\alpha_i = \alpha_i(y, v_1, \ldots, v_i, \omega, \zeta; \tilde{\theta}_k, \hat{\theta}_k', 1 \leqslant k \leqslant i)$ with $\tilde{\theta}_k \in \mathscr{W}_1$, $\hat{\theta}_k' = (\hat{\theta}_{1k}', \ldots, \hat{\theta}_{mk}')^{\mathrm{T}} \in \mathscr{W}$ denoting the parameter estimates of $\theta_1 \in \mathscr{W}_1$ and $\theta' = (\theta_1', \ldots, \theta_m')^{\mathrm{T}} \in \mathscr{W}$ at step k, $1 \leqslant k \leqslant n$, respectively.

The functions α_i, $1 \leqslant i \leqslant n$ are defined: $\alpha_1 = -\omega_2 - \sum_{j=1}^m \hat{\theta}_{j1}'\zeta_{j,2} - \tilde{\theta}_1^{\mathrm{T}}\phi_1(y) - f_1^0(y) - c_1 z_1 - (nl^2/3 + \kappa^2)z_1$, and for $2 \leqslant i \leqslant n$,

$$\begin{aligned}
\alpha_i = {} & \frac{\partial\alpha_{i-1}}{\partial y}\left(\omega_2 + \sum_{j=1}^m \hat{\theta}_{ji}'\zeta_{j,2} + v_2 + \tilde{\theta}_i^{\mathrm{T}}\phi_1(y) + f_1^0(y)\right) \\
& + \frac{\partial\alpha_{i-1}}{\partial\omega}(A_0\omega + \gamma y + f^0(y)) \\
& + \sum_{j=1}^m \frac{\partial\alpha_{i-1}}{\partial\zeta_j}(A_0\zeta_j + \Phi_j'(y)) \\
& + \sum_{k=1}^{i-1}\left(\frac{\partial\alpha_{i-1}}{\partial\hat{\theta}_k'}\hat{\beta}_k + \frac{\partial\alpha_{i-1}}{\partial\tilde{\theta}_k}\tilde{\beta}_k\right) \\
& + \sum_{j=1}^{i-1}\frac{\partial\alpha_{i-1}}{\partial v_j}(v_{j+1} - \gamma_j v_1) \\
& - \left(\frac{nl^2}{3} + \kappa^2\right)\left(\frac{\partial\alpha_{i-1}}{\partial y}\right)^2 z_i - c_i z_i - z_{i-1} + \gamma_i v_1,
\end{aligned} \tag{7}$$

where

$$\begin{aligned}
&\tilde{\beta}_1 = \alpha z_1\tilde{G}\phi_1(y), \quad \tilde{\beta}_1 \in \mathscr{W}_1, \\
&\tilde{\beta}_k = -\alpha z_k\frac{\partial\alpha_{k-1}}{\partial y}\tilde{G}\phi_1(y), \quad \tilde{\beta}_k \in \mathscr{W}_1, \quad 2 \leqslant k \leqslant n, \\
&\hat{\beta}_1 = \alpha z_1\hat{G}(\zeta_{1,2}, \ldots, \zeta_{m,2})^{\mathrm{T}}, \quad \hat{\beta}_1 \in \mathscr{W},
\end{aligned} \tag{8}$$

$$\hat{\beta}_k = -\alpha z_k\frac{\partial\alpha_{k-1}}{\partial y}\hat{G}(\zeta_{1,2}, \ldots, \zeta_{m,2})^{\mathrm{T}}, \quad \hat{\beta}_k \in \mathscr{W}, \quad 2 \leqslant k \leqslant n$$

and where $\hat{G}$, $\tilde{G}$ are defined from an adaptive structure $G = (G_1, \ldots, G_n)$ as $\tilde{G} = G_1$, $\hat{G} = \mathrm{diag}(G_1, \ldots, G_n)$, where G_i, $1 \leqslant i \leqslant n$ are positive-definite matrices. $\alpha > 0$ is the adaptive gain, $\kappa > 0$ is the robust gain and $l > 0$ is the state estimation robust gain. Note that by the differentiability assumptions on f^0, Φ, it follows that for $1 \leqslant i \leqslant n$, α_i is defined and at least C^1, hence locally Lipschitz.

The controller, $\Xi = \Xi(G, Q, \alpha, \Phi, \kappa, l, \gamma)$, (where $Q = \mathrm{diag}(c_1, \ldots, c_n)$) is defined by filters 3–5, and Eqs. (9) below:

$$\Xi : u = \alpha_n(y, v, \omega, \zeta_j, \ 1 \leqslant j \leqslant n, \tilde{\theta}_k, \hat{\theta}_k', \ 1 \leqslant k \leqslant n),$$

$$\dot{\tilde{\theta}}_i = D(\hat{\Omega}_0, z)\tilde{\beta}_i, \quad \tilde{\theta}_i(0) = 0, \quad \tilde{\theta}_i \in \mathscr{W}_1, \ 1 \leqslant i \leqslant n,$$

$$\dot{\hat{\theta}}_i' = D(\hat{\Omega}_0, z)\hat{\beta}_i, \quad \hat{\theta}_i(0) = 0, \quad \hat{\theta}_i' = (\hat{\theta}_{1i}', \ldots, \hat{\theta}_{mi}')^{\mathrm{T}} \in \mathscr{W}, \quad 1 \leqslant i \leqslant n, \tag{9}$$

where $D(\hat{\Omega}_0, z)$ denotes the dead-zone function: $D(\hat{\Omega}_0, z)=0$ if $z\in\hat{\Omega}_0$, $D(\hat{\Omega}_0, z)=1$ if $z\notin\hat{\Omega}_0$, where $\hat{\Omega}_0=\{z : z^{\mathrm{T}}z\leqslant 2\eta^2\}$.

Before giving the main theorem, we give some more notation. It will be convenient to partition $\hat{\theta}_i'\in\mathscr{W}$ as follows: $\hat{\theta}_i'=(\hat{\theta}_{1i},\hat{\theta}_{2i},\dots,\hat{\theta}_{mi})^{\mathrm{T}}=(\hat{\theta}_{1i}^{\mathrm{T}}\,|\,\hat{\theta}_{2i}^{\mathrm{T}}|;\cdots|\hat{\theta}_{ni}^{\mathrm{T}})^{\mathrm{T}}$ where $\hat{\theta}_{ki}\in\mathscr{W}_k$ for $1\leqslant i,k\leqslant n$. $\hat{\theta}_{ki}$ is then the adaptive estimate at step k of the parameter θ_i. Note that θ_1 plays a special role, it has two parameter estimates constructed for it at each step k, namely $\tilde{\theta}_k$, $\hat{\theta}_{1k}$. The parameter estimate vectors, $\tilde{\theta}_i$, $\hat{\theta}_i'$, $1\leqslant i\leqslant n$ are concatenated into the vectors $\tilde{\Theta}=(\tilde{\theta}_1,\dots,\tilde{\theta}_n)^{\mathrm{T}}\in\mathscr{W}_1^n$, $\hat{\Theta}'=(\hat{\theta}_1',\dots,\hat{\theta}_n')^{\mathrm{T}}\in\mathscr{W}^n$. We define a mapping $T:\mathscr{X}\times\mathbb{R}\times\mathscr{X}\times\mathscr{X}^m\times\mathscr{W}_1^n\times\mathscr{W}^n\to\mathscr{Z}\times\mathbb{R}\times\mathscr{X}\times\mathscr{X}^m\times\mathscr{W}_1^n\times\mathscr{W}^n$ by $T((y,v_2,\dots,v_n),v_1,\omega,\zeta,\tilde{\Theta},\hat{\Theta}')=(z,v_1,\omega,\zeta,\tilde{\Theta},\hat{\Theta}')$. Model error constants are defined: $\sup_{f\in\Delta}\|d_i^f\|_{L^2(\Omega\setminus\hat{\Omega}_0;w)}\leqslant q_i$, $1\leqslant i\leqslant n$, $\sup_{f\in\Delta}\sup_{y\in\Omega\setminus\hat{\Omega}_0}\|d^{f^{\mathrm{T}}}(y)P_0+P_0d^f(y)\|_2\leqslant g$, $\sup_{f\in\Delta}\|d^f(y)\|_{C(\Omega)}\leqslant s$, where P_0 is the solution to Lyapunov's equation $A_0^{\mathrm{T}}P_0+P_0A_0=-I$. Note that when Ω is compact or w is integrable, the boundedness of q_i, $1\leqslant i\leqslant n$ and g follows from the boundedness of s.[3] Finally, we require the admissibility definition.

Definition 3.1. A model $\Phi=(\phi_1,\dots,\phi_n)$ is $(Q,\Omega,\Omega_0,P_0,\kappa,l)$ admissible if: s, g, q_i for $1\leqslant i\leqslant n$ are finite, and $2\underline{\lambda}(Q)\eta^2>ns^2/4\kappa^2+g^2/l^2$, $2\underline{\lambda}(Q)\eta^2>ns^2/4\kappa^2+3\bar{\lambda}(P_0)g^2/4l^2$.

The main theorem is then as follows:

Theorem 3.2. *Let $\Omega\subset\mathscr{X}$ be a fixed closed set. Consider the system Σ_Δ given by Eq.* (1) *with functional uncertainty $\Delta\subset\Delta(L^2(\Omega\setminus\Omega_0;w),f^0,\delta)$ and initial condition $x_0\in\mathscr{X}$. Consider the performance measure $\mathscr{P}=\mathscr{P}(c_1,k,\eta)$ for positive diagonal Q, and $k>0$, $\eta>0$. Let $G=(G_1,\dots,G_n)$ where G_i, $1\leqslant i\leqslant n$ are positive-definite adaptive structure matrices. Implement the controller $\Xi(G,Q,\alpha,\Phi,\kappa,l,\gamma)$ where Φ is a finite dimensional model and $\alpha>0$. Suppose the filter gain γ is admissible, and $\kappa>0$, $l>0$. Suppose Φ is $(Q,\Omega,\Omega_0,P_0,\kappa,l)$ admissible, and let*

$$W_\alpha=\frac{1}{2}\max\{z_0^{\mathrm{T}}z_0,\,2\eta^2\}+\frac{1}{l^2}\max\{x_0^{\mathrm{T}}P_0x_0,3\bar{\lambda}(P_0)g^2/4\}+\frac{n}{\alpha}\frac{(\delta_1+q_1)^2}{\underline{\lambda}(G_1)\underline{\lambda}(R_1)}+\frac{n}{2\alpha}\sum_{i=2}^{n}\frac{(\delta_i+q_i)^2}{\underline{\lambda}(G_i)\underline{\lambda}(R_i)},$$

where $R_i=\langle\phi_i,\phi_j\rangle_{L^2(\Omega\setminus\Omega_0;w)}$ is the Gram matrix of the model component ϕ_i and P_0 is defined as above. Then for all adaption gains $\alpha>0$ and state estimation control gains $l>0$ such that:

$$[-\sqrt{2W_\alpha},\sqrt{2W_\alpha}]\subset\Omega^\circ \tag{10}$$

we have that (Σ_Δ,Ξ) is well posed and all outputs satisfy $y(t)\to\Omega_0$ as $t\to\infty$. Furthermore, we have the bound:

$$\begin{aligned}&\mathscr{P}(c_1,k,\eta)\\&\leqslant\left(1+\frac{p_1}{(2\underline{\lambda}(Q)\eta^2-p_1)}+\frac{p_2}{(2\underline{\lambda}(Q)\eta^2-p_2)}\right)(W_\alpha-\eta^2)\\&\quad+k\sup_{\eta^2\leqslant V_*\leqslant W_\alpha}\left(\frac{1}{2\underline{\lambda}(Q)\eta^2-p_1}\int_{V_*}^{W_\alpha}\tilde{u}_1^2(v)\,\mathrm{d}v\right.\\&\quad\left.+\frac{1}{2\underline{\lambda}(Q)\eta^2-p_2}\int_{\eta^2}^{V_*}\tilde{u}_2^2(v)\,\mathrm{d}v\right),\end{aligned}\tag{11}$$

where $\tilde{u}_1:\mathbb{R}\to\mathbb{R}$, $\tilde{u}_2:\mathbb{R}\to\mathbb{R}$, $p_1>0$, $p_2>0$ are defined:

$$p_1=\frac{ns^2}{4\kappa^2}+\frac{g^2}{l^2},\quad p_2=\frac{ns^2}{4\kappa^2}+\frac{3\bar{\lambda}(P_0)g^2}{4l^2},$$

$$\begin{aligned}&\tilde{u}_1(r)\\&=\sup\{|u(y,v,\zeta,\omega,\hat{\Theta}',\tilde{\Theta})|\in\mathbb{R}\,|\,(y,v,\zeta,\omega,\hat{\Theta}',\tilde{\Theta})\in Z,\\&\qquad\eta^2\leqslant V(z,\varepsilon,\hat{\Theta}',\tilde{\Theta})\leqslant r\},\end{aligned}$$

$$\begin{aligned}&\tilde{u}_2(r)\\&=\sup\left\{|u(y,v,\zeta,\omega,\hat{\Theta}',\tilde{\Theta})|\in\mathbb{R}\;|\;(y,v,\zeta,\omega,\hat{\Theta}',\tilde{\Theta})\in Z,\right.\\&\qquad\left.\varepsilon^{\mathrm{T}}P_0\varepsilon\leqslant\frac{3\bar{\lambda}(P_0)g^2}{4l^2},\;\eta^2\leqslant V(z,0,\hat{\Theta}',\tilde{\Theta})\leqslant r\right\}\end{aligned}$$

where

$$\begin{aligned}Z=\Big\{&(y,v,\zeta,\omega,\hat{\Theta}',\tilde{\Theta})\;|\;\zeta_j^{\mathrm{T}}P_0\zeta_j\leqslant 4\bar{\lambda}(P_0)^3\\&\sup_{y^2\leqslant 2W_\alpha}\|\Phi_j'(y)\|;\,\omega^{\mathrm{T}}P_0\omega\leqslant 4\bar{\lambda}(P_0)^3\\&\sup_{y^2\leqslant 2W_\alpha}\|\gamma y+f^0(y)\|;\;(y,v_2,\dots,v_n)=z;\\&v_1=y-\omega_1-\sum_{j=1}^{m}\theta_j'\zeta_{j,1}-\varepsilon_1;\eta^2\\&\leqslant V(z,\varepsilon,\hat{\Theta}',\tilde{\Theta})\leqslant r\Big\}.\end{aligned}\tag{12}$$

If $f^0(0)=0$ and $\phi(0)=0$, then $x_0=0$ implies $z_0=0$ irrespective of the uncertainty level δ (e.g. in the case of stabilising to an equilibrium point given $x_0=0$ and e.g. a polynomial basis). This can shown recursively from the definition of α_i (Eq. (7)) and the fact that the filters ω,ζ,v are initialised at 0. In this case, given Ω, then suitable adaption and state-estimation control gains can be computed for any uncertainty level δ to ensure condition (10).

[3] The reason for the variety of error constants is thus to minimise conservatism in the performance bounds, and for notational simplicity.

It is interesting to observe that orthogonal models have the property that the size of their basis can be increased without altering the control/adaption gains to maintain stability/uniform bound on the output cost. For many other typical approximants, such as Gaussian RBF models of Sanner and Slotine (1992), or B-splines defined on uniform lattices, we have that $\underline{\lambda}(R_i) \to 0$ as the resolution of the model is increased: to ensure stability/uniform bound on the output cost for such schemes it is necessary to select the adaption gain proportional to $1/\underline{\lambda}(R_i)$. It remains unclear how these scalings affect the control cost (it can easily be observed that selecting α proportional to $1/\underline{\lambda}(R_i)$ yields a uniformly bounded tracking error when the resolution is increased. In a case of matched uncertainty, these scalings have been investigated for the full LQ cost, e.g. it has also been shown (French, Szepesvari, & Rogers, 1999) that special constructions of the basis, adaption gains and adaption structure matrices G lead to uniformly bounded (state and control) performance. Similar results in the output feedback case remain the subject for future work.

4. Summary and discussion

The main contributions of the paper are as follows:

1. A rigorous dead-zone modified robust adaptive output feedback practical stabilisation design is given. The design achieves practical stability in the prescence of bounded disturbances.
2. Uncertainty is characterised by a weighted L^2 model about a nominal system, and upper bounds on worst-case control performance are obtained.

It should be noted that although a stable design is given for a class of affine systems, the uncertainty model is data given in the coordinates of the normal form. The extensions to tracking and to minimum phase systems are expected to be routine. Similar to French and Rogers (1998), the basic design is overparameterised. This is a drawback from a implementation viewpoint; however, it should be noted that there is no clear evidence as to the relative advantages/disadvantages of different parameterisations w.r.t. to non-singular transient performance. It is expected that this overparamaterisation can be removed by the interlacing design concept of Krstić et al. (1995). One of the interesting features of these results is the fact that the uncertainty is naturally expressed in L^2 as opposed to pointwise bounds as in the robust results of Marino and Tomei (1993b), or the adaptive results of Jiang (1998). From a modelling/identification perspective, these L^2 models may well be more realistic. In contrast, it is hard to obtain good pointwise error bounds from identification data. Clearly, this is a topic for future work.

Although, the bounds obtained are likely to be conservative, we contend that these results have utility beyond a first attempt at a constructive a-priori determined cost functional bound. The fundamental unanswered question concerning these approximate designs concerns the scaling of the performance when the resolution of the model is increased. We have shown in this paper that the output transient is uniformly bounded if the model basis is e.g. orthogonal or, if the adaption gain is taken to be proportional to the reciprocal of the minimum eigenvalue of the Gram matrix. As yet, for the output feedback case, there is no general construction of a model class whose (output and) control performance is uniformly bounded independently of the model resolution. However, by extending a recent construction (French et al., 1999) utilising multiresolution models for the matched (state feedback) case, and by using the bounds given in this paper, we expect that a suitable adaptive model basis can be constructed. This, however, also remains a topic for future work.

Appendix

Proposition A.1. *Consider a differential equation of the form*

$$\dot{x} = f(x,y,z), \quad x(0) = x_0,$$
$$\dot{y} = \begin{cases} 0 & \text{when } x \in \Omega_0, \\ p(x,y,z) & \text{when } x \notin \Omega_0, \end{cases} \quad y(0) = y_0, \tag{A.1}$$
$$\dot{z} = q(x,y,z), \quad z(0) = z_0,$$

where $x \in \mathscr{X}$, $y \in \mathscr{Y}$, $z \in \mathscr{Z}$, are finite dimensional, f, p, q are locally Lipschitz, and Ω_0 is of the form $\Omega_0 = \{x \in \mathscr{X} \mid x^{\mathrm{T}}Px \leqslant \eta^2\}$ for some positive definite P. Let $V : \mathscr{X} \times \mathscr{Y} \times \mathscr{Z} \to \mathbb{R}_+$ be defined: $V(x,y,z) = x^{\mathrm{T}}Px + g^2(y) + h^2(z)$ where $g \in C^1(\mathscr{Y})$, $h \in C^1(\mathscr{Z})$. Let $\varphi(t) = (x(t), y(t), z(t))$ be the absolutely continuous solution to Eq. (A.1) *defined over its maximal interval of existence $[0, t^*)$. Define $V(t) = V(x(t), y(t), z(t))$ and $W(t) = h^2(z(t))$, and let*:

$$\tau^* = \begin{cases} \inf\{0 \leqslant t < t^* \colon \dot{W}(t) > 0\} & \text{if } \exists t \in [0, t^*) \\ & \text{s.t. } \dot{W}(t) > 0, \\ t^* & \text{otherwise.} \end{cases} \tag{A.2}$$

Define $T_1 \subset \mathbb{R}_+$ as $T_1 = \{t \in [0, \tau^*) \mid x(t) \notin \Omega_0\}$. Suppose (1) $\limsup_{\|y\| \to \infty} |g(y)| = \infty$, (2) $\dot{V}(t) \leqslant -a < 0$ *for all $t \geqslant 0$ such that $t \in T_1$ and $\varphi(t) \in L$, where L is an open set containing $L(V, V_0)$,* (3) $D_-V(t) \leqslant 0$ *for all $t \geqslant 0$ such that $x(t) \in \partial\Omega_0$ and $D_-x(t)^{\mathrm{T}}Px(t) = 0$,* (4) *Let $T > 0$, then if $x, y \in L^\infty[0, T]$ are bounded then $z \in L^\infty[0, T]$, where $z(\cdot)$ is the solution of $\dot{z}(t) = q(x(t), y(t), z(t))$.*

Then $t^* > 0$ *and*: (1) *If* $t^* < \infty$ *then* $\tau^* < t^*$; (2) φ *is bounded on* $[0, \tau^*)$; (3) V *is decreasing on* $T_1 \cap [0, \tau^*)$; (4) $\varphi(t) \in L(V, V_0)$ *for all* $t \in [0, \tau^*)$ *where* $V_0 = \max(V(x_0, y_0, z_0), \max_{x \in \Omega_0} V(x, y_0, z_0))$. [4]

Proof. Let $\varphi(t)$ be an absolutely continuous local solution of (A.1) defined over its maximal interval of existence $[0, t^*)$. By e.g. (Filippov, 1988, Theorem 2, p. 78) either $t^* = \infty$, or $0 < t^* < \infty$ and $\limsup_{t \to t^{*-}} \|\varphi(t)\| = \infty$.

Let us establish 1. Assume that $t^* < \infty$. We claim that in order to prove $\tau^* < t^*$, it is sufficient to prove that

$$V(\varphi(t)) \leqslant V_0 \quad \forall t \in [0, \tau^*). \tag{A.3}$$

For, to derive a contradiction, assume that $\tau^* = t^* < \infty$. Then, from $\lim_{t \to t^{*-}} \|\varphi(t)\| = \infty$ it follows that also $\lim_{t \to t^{*-}} V(\varphi(t)) = \infty$. For if $\limsup_{t \to t^{*-}} \|(x(t), y(t))\| = \infty$ then by Condition 1. and since P is positive definite, $\lim_{t \to t^{*-}} V(\varphi(t)) = \infty$. If, on the other hand, $(x(t), y(t))$ remains bounded, then by Condition 4. $z(t)$ stays bounded and hence $\limsup_{t \to t^{*-}} \|\varphi(t)\| < \infty$, which is a contradiction. Therefore $\limsup_{t \to t^{*-}} \|(x(t), y(t))\| < \infty$ cannot hold and thus $\lim_{t \to t^{*-}} V(\varphi(t)) = \infty$. So (A.3) implies 1.

Now let us prove (A.3). If $x_0 \in \Omega_0$ then by letting $b = \min\{\inf\{t \in [0, \tau^*) : x(t) \in \partial\Omega_0\}, \tau^*\}$ and by the definition the dynamics (Eq. (A.1)), $V(\varphi(t)) \leqslant V(\varphi(b))$ as long as $t \in [0, b)$. So if we prove that for all $t \in [b, \tau^*)$ $V(\varphi(t)) \leqslant V(\varphi(b))$ then we will have $V(\varphi(t)) \leqslant V(\varphi(b)) \leqslant \max_{x \in \Omega_0} V(x, y_0, z_0) \leqslant V_0$ for all $t \in [0, \tau^*)$. Hence, by the time invariance of Eq. (A.1), we can assume without loss of generality $x_0 \notin \Omega_0$. So let us assume that $x_0 \notin \Omega_0$. Since φ, x, y, z are absolutely continuous on $[0, \tau^*)$, it follows that $V = V(t)$ is absolutely continuous on $[0, \tau^*)$. Hence (e.g. Rudin, 1987, Theorem 7.18), V is differentiable a.e. and $\dot{V} \in L^1[0, \tau^*)$. For a contradiction suppose $\tau \in [0, \tau^*)$ is such that (i) $\varphi(\tau) \in L \backslash L(V, V_0)$, and (ii) $\varphi(t) \in L$ $\forall t \in [0, \tau)$, then:

$$V(\tau) = V(0) + \int_0^\tau \dot{V}(t)\,\mathrm{d}t = V(0) + \int_{F_1} \dot{V}(t)\,\mathrm{d}t + \int_{F_2} \dot{V}(t)\,\mathrm{d}t + \int_{F_3} \dot{V}(t)\,\mathrm{d}t, \tag{A.4}$$

where $F_1 = T_1 \cap [0, \tau) = \{t \in [0, \tau) \mid x(t) \in \mathscr{X} \backslash \Omega_0\}$, $F_2 = \{t \in [0, \tau) \mid x(t) \in \Omega_0^\circ\}$ and $F_3 = \{t \in [0, \tau) \mid x(t) \in \partial\Omega_0\}$. F_1, F_2 are measurable (as Ω_0°, $\mathscr{X} \setminus \Omega_0$ are open and φ is continuous); F_3 is measurable as it is the complement of $F_1 \cup F_2$ in $[0, \tau^*)$. We are now going to estimate all the three integrals in (A.4). Let us assume $x(\tau) \notin \Omega_0^\circ$. Then:

(a) Since $\dot{V}(t) \leqslant -a \leqslant 0$ $\forall t \in F_1$ by Condition 2 and by (i) and (ii), it follows that $\int_{F_1} \dot{V}(t)\,\mathrm{d}t \leqslant 0$.

(b) Since $x(\tau) \notin \Omega_0^\circ$, we can write $F_2 = \bigcup_{a \in A} G_a = \bigcup_{n \geqslant 1} \bigcup_{a \in A_n} G_a$ where $G_a = (t_a^-, t_a^+)$ are maximal disjoint open intervals with $x(t_a^-)$, $x(t_a^+) \in \partial\Omega_0$; $A_n = \{a \in A \mid m(G_a) \geqslant 1/n\}$. As $m(F_2) < \infty$, the cardinality of each A_n is finite, hence by the dominated convergence theorem ($\dot{V} \in L^1[0, \tau^*)$): $\int_{F_2} \dot{V}(t)\,\mathrm{d}t = \lim_{n \to \infty} \int_{\bigcup_{a \in A_n} G_a} \dot{V}(t)\,\mathrm{d}t = \lim_{n \to \infty} \sum_{a \in A_n} V(t_a^+) - V(t_a^-) \leqslant 0$, since by definition of Ω_0 we have $x(t_a^-)^{\mathrm{T}} P x(t_a^-) = x(t_a^+)^{\mathrm{T}} P x(t_a^+) = \eta^2$, and $\dot{y}(t) = 0$ and $\dot{W}(t) \leqslant 0$ $\forall t \in G_a$ so $V(t_a^-) \geqslant V(t_a^+)$.

(c) We decompose $F_3 = F_4 \cup F_5$, where $F_4 = \{t \in F_3 \mid D_- x(t)^{\mathrm{T}} P x(t) = 0\}$, $F_5 = F_3 \setminus F_4$. (Note that $D_- x(t)^{\mathrm{T}} P x(t)$ is defined only a.e. (as $(\mathrm{d}/\mathrm{d}t) x^{\mathrm{T}} P x$ is defined a.e.) but is measurable a.e. so F_4 and F_5 are measurable). Write $T_1 = \bigcup_{b \in B} E_b$ where $E_b = (t_b^-, t_b^+)$ are maximal disjointed connected subsets of $\mathbb{R}$ (this can be done since $x(t)$ is continuous). If $t_0 \in F_5$ then there cannot exist an $\varepsilon > 0$ s.t. $\forall t \in (t_0 - \varepsilon, t_0]$ $t \in F_3$, since then (from the definition of Ω_0 and F_3) $D_- x^{\mathrm{T}} P x = 0$ would hold at t_0 which contradicts the definition of F_5. So if $t_0 \in F_5$ then $t_0 \notin F_3^\circ$, so $t_0 \in \partial F_3 \subset \bar{T}_1$. But $t_0 \notin T_1$ so t_0 must be an endpoint of E_b for some b. Since the E_b's are disjoint open intervals B must be countable and so $m(F_5) = 0$. Hence, $\int_{F_3} \dot{V}(t)\,\mathrm{d}t = \int_{F_4} \dot{V}(t)\,\mathrm{d}t$. But similarly to Case 1 above, we have $\dot{V}(t) \leqslant 0$ for all $t \in F_4$ by Conditions 3, so $\int_{F_4} \dot{V}(t)\,\mathrm{d}t \leqslant 0$, hence $\int_{F_3} \dot{V}(t)\,\mathrm{d}t \leqslant 0$. Now suppose $x(\tau) \in \Omega_0^\circ$. Let $\tau' = \sup\{t < \tau \mid x(t) \in \partial\Omega_0\}$. Then $x(\tau') \notin \Omega_0^\circ$ and thus $V(\tau') \leqslant V(0)$. Then by definition of Ω_0, $V(\tau) \leqslant V(\tau') \leqslant V(0)$ which finishes the proof that $V(\tau) \leqslant V(0)$ holds for all τ in $[0, \tau^*)$.

Thus, Cases (a)–(c) and Eq. (A.4), show that $V(\tau) \leqslant V(0)$ $\forall \tau \in [0, \tau^*)$, hence contradiction. This establishes 4 and hence 1. Similarly, we can establish that $\forall t_1, t_2 \in T_1$, $t_1 < t_2$ implies $V(\varphi(t_2)) \leqslant V(\varphi(t_1))$, thus establishing 3. Boundedness of z follows easily, since consequence 4 implies x, y will be bounded and so by Condition 4. z is bounded, hence establishing consequence 2, and hence completing the proof. □

Lemma A.2. *Consider the differential equation* $\dot{x} = f(x)$, *where* $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ *and* f *maps bounded sets to bounded sets. Let* $x(t)$ *be an absolutely continuous solution. Suppose* $x_1(t)$ *is bounded and differentiable on* $T_1 = x_1^{-1}(\Omega_0^c)$, *where* $\Omega_0 \subseteq \mathbb{R}^{n_1}$ *is compact. Further, let* $V \in C^1(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ *be a real-valued function and suppose* $v(t) = V(x(t))$ *satisfies* (1) $0 \leqslant V = V(x)$; (2) v *is non-increasing on* T_1; (3) $\dot{v}|_{T_1} < -a < 0$.

Then $m(T_1) \leqslant (1/a) v(0)$ *and* $\lim_{t \to \infty} d(x_1(t), \Omega_0) = 0$.

Proof. Initially, assume that $x_1(0) \notin \Omega_0^\circ$ and define $\dot{\hat{v}} = \dot{v}$ if $t \in T_1$ and 0, otherwise with $\hat{v}(0) = v(0)$.

[4] Note that if V is radially unbounded then this result is much simpler to state and prove, however, this does not suffice for the application required.

Then by Condition 2. $v|_{T_1} \leqslant \hat{v}|_{T_1}$. Now consider $\tau \in T_1$. Then $v(\tau) \leqslant \hat{v}(\tau) = v(0) + \int_0^\tau \dot{\hat{v}}(t)\,dt \leqslant v(0) - \int_{[0,\tau)\cap T_1} a\,dt = v(0) - am([0,\tau)\cap T_1)$ and thus $am([0,\tau)\cap T_1) \leqslant v(0) - v(\tau) \leqslant v(0)$, where the second inequality follows since by Condition 1. Now letting $\tau \to \sup T_1$, we obtain $m(T_1) \leqslant v(0)/a$ as required.

Now write $T_1 = \bigcup_{n\in N} A_n$, where $\{A_n\}$ is an ordered set of open intervals, where N is a countable set, and $m(A_n) \to 0$ as $n \to \sup N$ (this decomposition exists as T_1 is open). If N is finite then $\lim_{t\to\infty} d(x_1(t), \Omega_0) = 0$, so assume that N is countably infinite. To derive a contradiction assume that there exists an $\varepsilon > 0$ s.t. $\limsup_{t\to\infty} d(x_1(t), \Omega_0) > \varepsilon$. Then there exists an increasing sequence of naturals n_k s.t. for all k there exists $t \in A_{n_k} = (l_{n_k}, u_{n_k})$ s.t. $d(x_1(t), \Omega_0) \geqslant \varepsilon$. Define $t^*_{n_k} = \inf\{t \in A_{n_k} \mid d(x_1(t), \Omega_0) = \varepsilon/2\}, t'_{n_k} = \inf\{t \in A_{n_k} \mid d(x_1(t), \Omega_0) = \varepsilon\}$. By the mean-value theorem there exists a time $\hat{t}_{n_k} \in (t^*_{n_k}, t'_{n_k})$ s.t. $\|\dot{x}_1(\hat{t}_{n_k})\| = [\|x_1(t^*_{n_k}) - x_1(t'_{n_k})\|]/(t^*_{n_k} - t'_{n_k}) \geqslant \varepsilon/2m(A_{n_k})$ and therefore $\lim_{k\to\infty} \|\dot{x}_1(\hat{t}_{n_k})\| = \infty$. However, $W = \{x_1(\hat{t}_{n_k})\}_k$ is a bounded set (as $w \in W$ lies within ε of the compact set Ω), f maps bounded sets to bounded sets and so $\{\dot{x}_1(\hat{t}_{n_k})\}_k = f(W)$ is also bounded which is a contradiction. Therefore, we must have $\limsup_{t\to\infty} d(x_1(t), \Omega_0) = 0$.

Now, let us assume that $x_1(0) \in \Omega_0$. Let $t^* = \inf\{t : x_1(t) \in \partial\Omega_0\}$. Then $x_1(t^*) \notin \Omega_0^\circ$. Consider $y(t) = x(t + t^*)$, $t > 0$. Then, by the above argument $\limsup_{t\to\infty} d(y_1(t), \Omega_0) = 0$ and thus also $\limsup_{t\to\infty} d(x_1(t), \Omega_0) = 0$. Since $d(x_1(t), \Omega_0) \geqslant 0$, the result follows. □

Proposition A.3. *Suppose T is a C^1 mapping $T : \mathscr{X} \times \mathscr{W} \to \mathscr{Z} \times \mathscr{W}$ and let $x : \mathbb{R}_+ \to \mathscr{X}$, $y : \mathbb{R}_+ \to \mathscr{Y}$ be continuous signals. Define $z : \mathbb{R}_+ \to \mathscr{Z}$ by $z(t) = P_n T(x(t), y(t))$, and let $V : \mathscr{Z} \times \mathscr{W} \to \mathbb{R}$ be defined by: $V(z, y) = \frac{1}{2}z^{\mathrm{T}}z + \frac{1}{g}g^2(y)$ where $g \in C^1(\mathscr{Y}, \mathbb{R})$. Suppose (1) $\dot{V}(z(t), y(t)) \leqslant -a < 0\ \forall t \in T_1$, and (2) V is decreasing on T_1, where $T_1 = \{t \geqslant 0 \mid z(t) \notin \Omega'_0\}$ and $\Omega'_0 = \{z \in \mathscr{Z} \mid z^{\mathrm{T}}z \leqslant 2\eta^2\}$. If $u : \mathscr{X} \times \mathscr{W} \to \mathbb{R}$ and $\tilde{u} : \mathbb{R} \to \mathbb{R}$ satisfy: $u^2(x(t), y(t)) \leqslant \tilde{u}^2(V(T(x(t), y(t))))\ \forall t \in T_1$, then, $\int_{T_1} u^2(x(t), y(t))\,dt \leqslant (1/a)\int_{\eta^2}^{V(z(0),y(0))} \tilde{u}^2(v)\,dv$.*

Proof. As P_n, T, x, y are continuous, z is continuous, hence $T_1 = z^{-1}(\mathscr{Z}\backslash\Omega'_0)$ is measurable since $\mathscr{Z}\backslash\Omega'_0$ is open. Consider the change of variables $v(t) = V(z(t), y(t)) = V_t$, then: $\int_{T_1} u^2(x(t), y(t))\,dt \leqslant \int_{T_1} \tilde{u}^2(V(T(x(t), y(t))))\,dt \leqslant \int_{v(T_1)} \tilde{u}^2(v)|dt/dv|\,dv \leqslant (1/a)\int_{\eta^2}^{V_0} \tilde{u}^2(v)\,dv$ where the change of variables is justified since over T_1, v is decreasing; $v(T_1)$ is measurable by the measurability of V and T_1; and the final inequality follows from the inclusion: $v(T_1) \subset [\eta^2, V(z(0), y(0))]$. □

Lemma A.4. *Consider the following system: $\dot{x} = Ax + \mathrm{d}(t)\ x \in \mathbb{R}^n$, $x(0) = x_0$, where A is a Hurwitz matrix and let P be the solution to the Lyapunov equation $A^{\mathrm{T}}P + PA = -I$. If solutions are defined on $[0, \tau]$ then:*

(1) $\forall t \in [0, \tau]\ x_0^{\mathrm{T}}Px_0 \leqslant 4\bar{\lambda}(P)^3 \sup_{s\in[0,\tau]} \| d(s) \|^2 \Rightarrow x^{\mathrm{T}}(t)Px(t) \leqslant 4\bar{\lambda}(P)^3 \sup_{s\in[0,\tau]} \| d(s)\|^2$.
(2) *If $\sup_{t\geqslant 0} \| d(t)^{\mathrm{T}}P + P\,d(t)\|_2 \leqslant g$ and if L_0 is defined by: $L_0 = \{x \in \mathscr{X} \mid x^{\mathrm{T}}Px \leqslant \bar{\lambda}(P)g^2\}$, then $x(t) \to L_0$ as $t \to \infty$. Further, if $\tau = \inf\{t \geqslant 0 \mid x^{\mathrm{T}}(t)Px(t) \leqslant \bar{\lambda}(P)g^2\}$ then $V(t) = x(t)^{\mathrm{T}}Px(t)$ is monotonically decreasing on $[0, \tau]$.*

Proof. The proof is easily obtained by considering the Lyapunov function $V(x) = x^{\mathrm{T}}Px$. □

Proof of Theorem 3.2. Some standard algebraic manipulations give the system in the error system coordinates when $z \in \mathscr{Z} \setminus \hat{\Omega}_0 : \dot{z}_1 = -c_1z_1 + z_2 + \sum_{j=1}^m (\theta'_j - \hat{\theta}_{j1})\zeta_{j,2} + (\theta_1 - \tilde{\theta}_1)^{\mathrm{T}}\phi_1(y) + (d_1(y) - \kappa^2 z_1) + (\varepsilon_2 - nl^2/3z_1)$. By letting $z_{n+1} = 0$ for $2 \leqslant i \leqslant n$ we have

$$\dot{z}_i = z_{i+1} - c_iz_i - z_{i-1} - \frac{\partial\alpha_{i-1}}{\partial y}\sum_{j=1}^m (\theta'_j - \hat{\theta}_{ji})\zeta_{j,2} - \frac{\partial\alpha_{i-1}}{\partial y}(\theta_1 - \tilde{\theta}_i)^{\mathrm{T}}\phi_1(y) + \frac{\partial\alpha_{i-1}}{\partial y}\left(-\varepsilon_2 - \frac{nl^2}{3}\frac{\partial\alpha_{i-1}}{\partial y}z_i\right) + \frac{\partial\alpha_{i-1}}{\partial y}\left(-d_1(y) - \kappa^2\frac{\partial\alpha_{i-1}}{\partial y}z_i\right). \tag{A.5}$$

It is also straightforward to compute the dynamics of the state estimation error, ε:

$$\dot{\varepsilon} = A_0\varepsilon + d(y). \tag{A.6}$$

We can now use Proposition A.1 to show the existence and boundedness of solutions in the maximal interval $[0, b]$ satisfying $\forall t \in [0, b]\ \dot{W}(t) \leqslant 0$ (for an exact definition see Proposition A.1). For this we identify x of the proposition with z, y with the adaptive estimator parameter, and z with the rest of the state variables. Then h is identified with $1/l^2(\varepsilon^{\mathrm{T}}P_0\varepsilon)$, and then Condition 1 is satisfied. Since $V(z, \varepsilon, \hat{\Theta}, \tilde{\Theta}) = \frac{1}{2}z^{\mathrm{T}}z + 1/l^2\varepsilon^{\mathrm{T}}P_0\varepsilon + (1/2\alpha)\sum_{i=1}^n ((\theta_1 - \tilde{\theta}_i)^{\mathrm{T}}\tilde{G}^{-1}(\theta_1 - \tilde{\theta}_i) + (\theta' - \hat{\theta}'_i)^{\mathrm{T}}\hat{G}^{-1}(\theta' - \hat{\theta}'_i))$, we can compute the following inequality $\forall y(t) \in \Omega\backslash\Omega_0$ and thus which also holds $\forall\varphi(t) \in L = \{(z, \varepsilon, \hat{\Theta}, \tilde{\Theta}) \mid z_1(t) \in \Omega^\circ\}$ such that $t \in T_1$ (note that $L(V, V_0) \subset L$):

$$\dot{V}(z, \varepsilon, \hat{\Theta}, \tilde{\Theta}) \leqslant -(2\underline{\lambda}(Q)\eta^2 - p_1), \tag{A.7}$$

where the inequalities follow from several applications of Young's inequality $ab - \frac{1}{4}b^2 \leqslant a^2$. Therefore Condition 2 holds. Consider $\frac{1}{2}z^{\mathrm{T}}z = \eta^2$, $D_-\frac{1}{2}z^{\mathrm{T}}z = 0$. Then similarly to inequality (A.7) we obtain

$$0 = D_{-}\tfrac{1}{2}z^{\mathrm{T}}z = -z^{\mathrm{T}}Qz + z_1(\theta_1 - \tilde{\theta}_1)^{\mathrm{T}}\phi_1 + \sum_{j=1}^{m} z_1(\theta'_j - \hat{\theta}'_{j1})\zeta_{j,2} + z_1(d_1 - \kappa^2 z_1) + z_1\left(\varepsilon_2 - \frac{nl^2}{3}z_1\right) + \sum_{i=2}^{n}\left(-z_i\frac{\partial\alpha_{i-1}}{\partial y}(\theta_1 - \tilde{\theta}_i)^{\mathrm{T}}\phi_1 - z_i\frac{\partial\alpha_{i-1}}{\partial y}\sum_{j=1}^{m}(\theta'_j - \hat{\theta}'_{ji})\zeta_{j,2} + z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-\varepsilon_2 - \frac{nl^2}{3}z_i\frac{\partial\alpha_{i-1}}{\partial y}\right) + z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-d_1 - \kappa^2 z_i\frac{\partial\alpha_{i-1}}{\partial y}\right)\right) \quad (A.8)$$

so, by definition of the solution at the discontinuity, for some $\lambda = \lambda(t) \in [0,1]$, we have

$$D_{-}V = -\lambda\left(\frac{1}{l^2}(-\varepsilon^{\mathrm{T}}\varepsilon + d^{\mathrm{T}}P_0\varepsilon + \varepsilon^{\mathrm{T}}P_0 d) - z^{\mathrm{T}}Qz + z_1(d_1 - \kappa^2 z_1) + z_1\left(\varepsilon_2 - \frac{nl^2}{3}z_1\right) + \sum_{i=2}^{n}\left(z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-\varepsilon_2 - \frac{nl^2}{3}z_i\frac{\partial\alpha_{i-1}}{\partial y}\right) + z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-d_1 - \kappa^2 z_i\frac{\partial\alpha_{i-1}}{\partial y}\right)\right)\right) \leqslant -\lambda(2\underline{\lambda}(Q)\eta^2 - p_1) \quad (A.9)$$

so $\dot{V} \leqslant 0$. This establishes Condition 3. Finally, we show that Condition 4, which requires that if z and $(\hat{\Theta}, \tilde{\Theta})$ are bounded then the signals (ω, ζ, v and x) also remain bounded. Explicit bounds on ω, ζ, and the boundedness of v and x are shown as follows: Since $y = z_1$, it follows that $y(t)$ is bounded by some bound B (in fact, B can be chosen to be equal to $\sqrt{2V_0}$ since $V(t)$ decreases outside of the dead zone, and y is constant inside the dead zone and therefore $(1/2)y^2(t) \leqslant V(t) \leqslant V(0)$ for all $t > 0$). By the definition of the filter 3, and Lemma A.4 we have:

$$\omega^{\mathrm{T}}P_0\omega \leqslant 4\bar{\lambda}(P_0)^3 \sup_{t\geqslant 0}\|\gamma y(t) + f^0(y(t))\| \leqslant \sup_{|y|\leqslant B}\|\gamma y + f^0(y)\|. \quad (A.10)$$

Similarly from the definition of the filter 4 we have

$$\dot{\zeta}_j = A_0\zeta_j + \Phi'_j(y), \quad (A.11)$$

so by Lemma A.4 we have for each $1 \leqslant j \leqslant m$: $\zeta_j^{\mathrm{T}}P_0\zeta_j \leqslant 4\bar{\lambda}(P_0)^3 \sup_{|y|\leqslant B}\|\Phi'_j(y)\|$. We now show v is bounded. Since $x_1 = y$, Eq. (6) implies:

$$\varepsilon_1 = y - \left(\omega_1 + \sum_{j=1}^{m}\theta'_j\zeta_{j,1} + v_1\right), \quad (A.12)$$

from which the boundedness of v_1 follows from the boundedness of $\theta', \zeta_1, \omega_1, y, \varepsilon_1$. If $v_1, \ldots, v_i$ are bounded, then α_i is bounded, and hence by the boundedness and definition of z it follows that v_{i+1} is bounded. Hence, v is bounded. Boundedness of u follows from the boundedness of α_n. Boundedness of x follows from Eq. (6). This establishes Condition 4. We now apply Proposition A.1 to show that solutions exist whilst $\dot{W} \leqslant 0$, $\varphi(t) \in L(V, V_0)$, and inequality 5 holds $\forall t \in [0,b) \cap T_1$.

Now, let $\tau = \inf\{t \geqslant 0 \mid W(t) \leqslant \bar{\lambda}(P)g^2\}$. We claim that $b \geqslant \tau$. Indeed, by Lemma A.4 and Eq. (A.6), $\dot{W}(t) \leqslant 0$ for all $t \in [0,\tau]$. We now consider the system on $[\tau, \infty)$. We use Proposition A.1 once again, but now to deduce that the solutions can be continued to infinity. Write $U(z, \varepsilon, \hat{\Theta}, \tilde{\Theta}) = V(z, 0, \hat{\Theta}, \tilde{\Theta})$, and identify U with V of Proposition A.1 (now $h \equiv 0$). Conditions 1 and 4 follow as previously. By Lemma A.4, we know that $\varepsilon(t)^{\mathrm{T}}P_0\varepsilon(t) \leqslant \bar{\lambda}(P)g^2$ for all $t \in [\tau, \infty)$. Hence, similarly to the derivation of inequality (5) we have $\forall y(t) \in \Omega \setminus \Omega_0$,

$$\dot{U}(z, \varepsilon, \hat{\Theta}, \tilde{\Theta}) = \dot{V}(z, 0, \hat{\Theta}, \tilde{\Theta}) \leqslant -z^{\mathrm{T}}Qz + ns^2/4\kappa^2 + 3\varepsilon_2^2/4l^2 \leqslant -z^{\mathrm{T}}Qz + ns^2/4\kappa^2 + 3\bar{\lambda}(P_0)g^2/4l^2 \leqslant -(2\underline{\lambda}(Q)\eta^2 - p_2). \quad (A.13)$$

since $y(t) \in \Omega \setminus \Omega_0$ implies $\varphi(t) \in L$ (note that $(U, U_\tau) \subset L(U, V_\tau) \subset L(V, V_\tau) \subset L(V, V_0) \subset L$. This establishes Condition 2.

To show Condition 3 of Proposition A.1 holds, consider $\frac{1}{2}z^{\mathrm{T}}z = 0$, $D_{-}\frac{1}{2}z^{\mathrm{T}}z = \eta^2$. Then similarly to 5, using Eq. (A.8), and by definition of the solution at the discontinuity, for some $\lambda = \lambda(t) \in [0,1]$, we have

$$D_{-}V = -\lambda\left(-z^{\mathrm{T}}Qz + z_1(d_1 - \kappa^2 z_1) + z_1\left(\varepsilon_2 - \frac{nl^2}{3}z_1\right) + \sum_{i=2}^{n}\left(z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-\varepsilon_2 - \frac{nl^2}{3}z_i\frac{\partial\alpha_{i-1}}{\partial y}\right) + z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-d_1 - \kappa^2 z_i\frac{\partial\alpha_{i-1}}{\partial y}\right)\right)\right) \leqslant -\lambda(2\underline{\lambda}(Q)\eta^2 - p_2) \quad (A.14)$$

so $\dot{V} \leqslant 0$ as required. This finishes the proof of the existence of bounded solutions on $[0, \infty)$, the boundedness of $\varphi(t)$ by $L(V, V_0)$ and similarly to the previous case, equations A.10, A.11, A.12 also hold on the interval $[\tau, \infty)$. We further have explicit bounds on $\dot{V}$ and $\dot{U}$ on $T_1 \cap [0,\tau)$, $T_1 \cap [\tau, \infty)$, respectively (inequalities A.7, A.13). Lemma A.2 yields the convergence of z to $\hat{\Omega}_0$, hence the convergence of y to Ω_0. Now we bound the state performance. Firstly, we have: $\int_{T_1} z(t)^{\mathrm{T}}Qz(t)\,\mathrm{d}t = \int_{T_1\cap[0,\tau]} -\dot{V}\,\mathrm{d}t + \int_{T_1\cap[0,\tau]} D_1(t)\,\mathrm{d}t +$

$\int_{T_1\cap[\tau,\infty)} -\dot{U}\,\mathrm{d}t + \int_{T_1\cap[\tau,\infty)} D_2(t)\,\mathrm{d}t$, where

$$D_1(t) = \sum_{i=2}^{n}\left(z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-\varepsilon_2 - \frac{nl^2}{3}z_i\frac{\partial\alpha_{i-1}}{\partial y}\right)\right.$$
$$\left.+ z_i\frac{\partial\alpha_{i-1}}{\partial y}\left(-d_1 - \kappa^2 z_i\frac{\partial\alpha_{i-1}}{\partial y}\right)\right)$$
$$+ z_1(d_1 - \kappa^2 z_1) + z_1\left(\varepsilon_2 - \frac{nl^2}{3}z_1\right)$$
$$+ \frac{1}{l^2}(-\varepsilon^{\mathrm{T}}\varepsilon + d^{\mathrm{T}}P_0\varepsilon + \varepsilon^{\mathrm{T}}P_0 d),$$
$$D_2(t) = D_1(t) - \frac{1}{l^2}(-\varepsilon^{\mathrm{T}}\varepsilon + d^{\mathrm{T}}P_0\varepsilon + \varepsilon^{\mathrm{T}}P_0 d), \qquad (\mathrm{A.15})$$

The first term is bounded by a MCT argument: We write $T_1 \cap [0,\tau) = \bigcup_{c\in C} E_c$ where $E_c = (t_c^-, t_c^+)$ are maximal disjointed connected subsets of $\mathbb{R}_+$ (this can be done since $z(t)$ is continuous), and define $C_n = \{c \in C \mid m(E_c) \geqslant 1/n\}$. Since: $\int_{\bigcup_{c\in C_n} E_c} -\dot{V}(t)\,\mathrm{d}t = \sum_{c\in C_n} V(t_c^-) - V(t_c^+) \leqslant V_0 - V_\tau$, by taking the limit as $n\to\infty$, and applying the monotone covergence theorem we can bound the first term: $\int_{T_1\cap[0,\tau]} -\dot{V}\,\mathrm{d}t = V(0) - V(\tau)$. Similarly, the third term is bounded: $\int_{T_1\cap[\tau,\infty)} -\dot{U}\,\mathrm{d}t \leqslant U(\tau) - \eta^2$. Since $\dot{V} \leqslant -2\underline{\lambda}(Q)\eta^2 + p_1\ \forall t \in T_1\cap[0,\tau]$, and since V is decreasing on T_1, we have

$$m(T_1\cap[0,\tau]) \leqslant \frac{\sup_{t\in[0,\tau]} V(t) - \inf_{t\in[0,\tau]} V(t)}{\inf_{t\in T_1}|\dot{V}(t)|}$$
$$\leqslant \frac{V(0) - V(\tau)}{2\underline{\lambda}(Q)\eta^2 - p_1} < \infty.$$

Bounds on $\int_0^\tau D_1(t)\,\mathrm{d}t$ are then given by

$$\int_{T_1\cap[0,\tau]} |D_1(t)|\,\mathrm{d}t \leqslant \tau\|D_1(\cdot)\|_{L^\infty} \leqslant \frac{p_1(V(0) - V(\tau))}{(2\underline{\lambda}(Q)\eta^2 - p_1)},$$
$$\int_{T_1\cap[0,\tau]} |D_2(t)|\,\mathrm{d}t \leqslant \frac{p_2(U(\tau) - \eta^2)}{(2\underline{\lambda}(Q)\eta^2 - p_2)}$$

since $\|D_1(\cdot)\|_{L^\infty} < p_1$, $\|D_2(\cdot)\|_{L^\infty} < p_2$ similarly to inequality 5. Hence the above inequalities and the inequalities $V(\tau) \geqslant \eta^2$, $V(0) \geqslant U(\tau)$ show that

$$\int_{T_1} z(t)^{\mathrm{T}}Qz(t)\,\mathrm{d}t$$
$$\leqslant \left(1 + \frac{p_1}{(2\underline{\lambda}(Q)\eta^2 - p_1)} + \frac{p_2}{(2\underline{\lambda}(Q)\eta^2 - p_2)}\right)(V(0) - \eta^2).$$

The control performance effort is bounded as follows. Consider $t \in [0,\tau]\cap T_1$ and define $r = V(z(t), \varepsilon(t), \hat{\Theta}'(t), \tilde{\Theta}(t))$. Then $(y(t), v(t), \zeta(t), \omega(t), \hat{\Theta}'(t), \tilde{\Theta}(t))$ lies in the set Z defined by Eq. (12). Now taking $(y, v_2, \ldots, v_n)$, $(\varepsilon, \hat{\Theta}', \tilde{\Theta})$ to have the role of x and y in Proposition A.3 we have the inequality: $u^2(t) \leqslant \tilde{u}_1^2(V(z, \varepsilon(t), \hat{\Theta}'(t), \tilde{\Theta}(t)))$, and hence by Proposition A.3,

$$\int_{T_1\cap[0,\tau]} u^2(t)\,\mathrm{d}t \leqslant \frac{1}{2\underline{\lambda}(Q)\eta^2 - p_1}\int_{V(\tau)}^{V(0)} \tilde{u}_1^2(v)\,\mathrm{d}v,$$

since $\dot{V} \leqslant -(2\underline{\lambda}(Q)\eta^2 - p_1)$ for all $t \in T_1\cap[0,\tau]$. Similarly consider $t \in [\tau,\infty)\cap T_1$ and define $r = U(z(t), \hat{\Theta}'(t), \tilde{\Theta}(t))$, then $(y(t), v(t), \zeta(t), \omega(t), \hat{\Theta}'(t), \tilde{\Theta}(t))$ lies in the set:

$$\left\{(y(t), v(t), \zeta(t), \omega(t), \hat{\Theta}'(t), \tilde{\Theta}(t)) \in Z \mid \varepsilon^{\mathrm{T}}P_0\varepsilon \leqslant \frac{3\bar{\lambda}(P_0)g^2}{4l^2}\right\}.$$

Now taking $(y, v_2, \ldots, v_n)$, $(\hat{\Theta}', \tilde{\Theta})$ to have the role of x and y in Proposition A.1 and taking $W = 0$ we have the inequality: $u^2(t) \leqslant \tilde{u}_1^2(V(z, \varepsilon(t), \hat{\Theta}'(t), \tilde{\Theta}(t)))$, and hence by Proposition A.3

$$\int_{T_1\cap[\tau,\infty)} u^2(t)\,\mathrm{d}t \leqslant \frac{1}{2\underline{\lambda}(Q)\eta^2 - p_2}\int_{\eta^2}^{V(\tau)} \tilde{u}_2^2(v)\,\mathrm{d}v.$$

We can then establish

$$\int_{T_1} u^2(t)\,\mathrm{d}t \leqslant \sup_{\eta^2 \leqslant V_* \leqslant V_0}\left(\frac{1}{2\underline{\lambda}(Q)\eta^2 - p_1}\int_{V_*}^{V_0} \tilde{u}_1^2(v)\,\mathrm{d}v + \frac{1}{2\underline{\lambda}(Q)\eta^2 - p_2}\int_{\eta^2}^{V_*} \tilde{u}_2^2(v)\,\mathrm{d}v\right).$$

The result now follows once we have shown that $V_0 \leqslant W_\alpha$. But since $\varepsilon^{\mathrm{T}}(0)P_0\varepsilon(0) = x_0^{\mathrm{T}}P_0x_0$, it follows that

$$V_0 = \frac{1}{2}z_0^{\mathrm{T}}z_0 + \frac{1}{l^2}x_0^{\mathrm{T}}P_0x_0 + \frac{1}{2\alpha}\sum_{i=1}^{n}\frac{(\delta_1 + q_1)^2}{\underline{\lambda}(G_1)\underline{\lambda}(R_1)}$$
$$+ \frac{n}{2\alpha}\sum_{i=1}^{n}\frac{(\delta_i + q_i)^2}{\underline{\lambda}(G_1)\underline{\lambda}(R_i)} = W_\alpha$$

since e.g. $(\theta_1 - \tilde{\theta}_i(0))^{\mathrm{T}}G_1^{-1}(\theta_1 - \tilde{\theta}_i(0)) = \theta_1^{\mathrm{T}}G_1^{-1}\theta_1 \leqslant 1/\underline{\lambda}(G_1)\underline{\lambda}(R_1)\theta_1^{\mathrm{T}}R_1\theta_1$ and

$$\begin{aligned}\theta_1^{\mathrm{T}}R_1\theta_1 &= \langle\theta_1^{\mathrm{T}}\phi_1, \theta_1^{\mathrm{T}}\phi_1\rangle_{L^2(\Omega\backslash\Omega_0;w)}\\ &= \|f - f^0 - (f - f^0 - \theta_1^{\mathrm{T}}\phi_1)\|^2_{L^2(\Omega\backslash\Omega_0;w)}\\ &\leqslant (\|f - f^0\|_{L^2(\Omega\backslash\Omega_0;w)} + \|f - f^0 - \theta_1^{\mathrm{T}}\phi_1\|_{L^2(\Omega\backslash\Omega_0;w)})^2\\ &= (\delta_1 + q_1)^2. \qquad (\mathrm{A.16})\end{aligned}$$

Thus, the result follows. □

References

Corless, M. J., & Leitmann, G. (1981). Continuous state feedback guaranteeing uniform ultimate boundedness for uncertain dynamic systems. *IEEE Transactions on Automatic Control*, *26*, 1139–1144.

Daubechies, I. (1992). *Ten lectures on wavelets.* Regional conference series on a. maths. Philadelphia: CBMS-NSF.

Filippov, A. F. (1988). *Differential equations with discontinuous right-hand sides* (1st Ed.). Netherlands: Kluwer Academic Publishers.

Freeman, R. A., & Kokotović, P. V. (1996). *Robust nonlinear control design* (1st ed.). Systems and control: Foundations and applications. Boston: Birkhauser.

French, M. (1998). *Adaptive control of functionally uncertain systems*. Ph.D. thesis, University of Southampton, UK.

French, M., & Rogers, E. (1998). Control and state performance measures in unmatched approximate adaptive nonlinear control. *Proceedings of the NOLCOS '98*, Vol. 3, Enschede, NL, July (pp. 679–684).

French, M., Szepesvari, C., & Rogers, E. (1999). An asymptotic scaling analysis of LQ performance for an approximate adaptive control design. *Mathematics of Control Signals and Systems*, in press.

French, M., Szepesvari, C., & Rogers, E. (2000). Uncertainty, performance and model dependency in approximate adaptive nonlinear control. *IEEE Transactions on Automatic Control*, *45*(2), 353–358.

Ikhouane, F., & Krstić, M. (1998). Robustness of the tuning functions adaptive backstepping design for linear systems. *IEEE Transactions on Automatic Control*, *43*(3), 431–437.

Jiang, Z. P. (1998). A note on robust adaptive output feedback control. *Proceedings of the NOLCOS '98*, IFAC (pp. 19–24).

Krstić, M., Kanellakopoulos, I., & Kokotović, P. V. (1995). *Nonlinear and adaptive control design* (1st ed.). New York: Wiley.

Krstić, M., & Kokotović, P. V. (1996). Adaptive nonlinear output-feedback schemes with Marino–Tomei controller. *IEEE Transactions on Automatic Control*, *41*(2), 274–280.

Li, Z. H., & Krstić, M. (1997). Optimal design of adaptive tracking controllers for non-linear systems. *Automatica*, *33*(8), 1459–1473.

Marino, R., & Tomei, P. (1993a). Global adaptive output-feedback control of nonlinear systems, part I: Linear paramaterization. *IEEE Transactions on Automatic Control*, *38*(1), 17–32.

Marino, R., & Tomei, P. (1993b). Robust stabilization of feedback linearizable time-varying uncertain nonlinear systems. *Automatica*, *29*(1), 181–189.

Marino, R., & Tomei, P. (1997). Adaptive output feedback control with disturbance attenuation for a class of nonlinear systems. *Proceedings of the European control conference.*

Narendra, K. S., & Annaswamy, A. M. (1990). *Stable adaptive control*. Information and System Sciences Series. Englewood Cliffs, NJ: Prentice-Hall.

Rivlin, T. J. (1969). *An introduction to the approximation of functions* (1st ed.). Waltham, MA: Bloisdell.

Rudin, W. (1987). *Real and complex analysis* (3rd Ed.) *International edition, mathematics series*. New York: McGraw-Hill.

Sanner, R. M., & Slotine, J. J. E. (1992). Gaussian networks for direct adaptive control. *IEEE Transactions on Neural Networks*, *3*(6), 837–863.

Teel, A. (1993). Adaptive tracking with robust stability. *Proceedings of the IEEE conference on decision and control* (pp. 570–575).

Yao, B., & Tomizuka, M. (1997). Adaptive robust control of SISO nonlinear systems in a semi-strict feedback form. *Automatica*, *33*(5), 893–901.

Csaba Szepesvári was born in 1969. He received his Ph.D. degree in statistics from the "Jozsef Attila" University, Szeged, in 1999. He is the holder of the Pro-Sciencia Golden Legature in Computer Science. After finishing his Ph.D. studies he worked as a researcher and lecturer in the Research Group on Artificial Intelligence at the University of Szeged, where he worked on reinforcement learning, machine learning and adaptive control. He joined Mindmaker Ltd. in October 1997, where he is presently the Vice President of Research.

His major research interests cover various aspects of artificial intelligence, such as machine learning, natural language processing and adaptive and intelligent control. He is currently an Editorial Board Member of AI Communications.

Eric Rogers was born in 1956 near Dungannon in Northern Ireland. He read mechanical engineering as an undergraduate in The Queen's University of Belfast, UK, and was subsequently awarded his Ph.D. by The University of Sheffield, UK, for a thesis in the area of multidimensional control systems theory. He has been with The University of Southampton, UK, since 1990 where he is currently Professor of Control Systems Theory and Design in the Department of Electronics and Computer Science. His major research interests currently include multidimensional systems theory and its applications (with particular emphasis on the behavioural approach and systems with repetitive dynamics), iterative learning control, flow control, and the control of spacecraft clusters. He is currently the editor of The International Journal of Control, an associate editor of Multidimensional Systems and Signal Processing, and has acted as a consultant to numerous government agencies and industries.

Mark French read mathematics at Oxford University, and obtained his Ph.D. from the University of Southampton in 1998. He is currently a lecturer in the Department of Electronics and Computer Science, University of Southampton. His research interests are in the area of nonlinear and adaptive control. He is currently serving as an associate editor of the International Journal of Control.