

Uncertainty, Performance, and Model Dependency in Approximate Adaptive Nonlinear Control

M. French *

E. Rogers ‡

Cs. Szepesvári †

IEEE Transactions on Automatic Control

To appear.



**University
of Southampton**

Abstract

We consider systems satisfying a matching condition which are functionally known up to weighted L^2 and L^∞ measures of uncertainty. A modified LQ measure of control and state transient performance is given, and the performance of a class of approximate model based adaptive controllers is studied. An upper performance bound is derived in terms of the uncertainty models (stability and the state transient bounds require only the L^2 uncertainty model; control effort bounds require both L^2 and L^∞ uncertainty models), and various structural properties of the model basis. Sufficient conditions are given to ensure that the performance is bounded independently of the model basis size.

*M.French@ecs.soton.ac.uk, Department of Electronics and Computer Science, University of Southampton, UK.

†szepes@sol.cc.u-szeged.hu, Research Group on Artificial Intelligence, Hungarian Academy of Sciences – JATE, Szeged, Hungary

‡E.Rogers@ecs.soton.ac.uk, Department of Electronics and Computer Science, University of Southampton, UK.

1 Introduction

Many approximate model based schemes have been presented (often in the neural network literature) for the adaptive stabilisation and tracking of the integrator chain:

$$\Sigma_f : \dot{x}_i = x_{i+1} \quad 1 \leq i \leq n-1 \quad (1)$$

$$\dot{x}_n = f(x) + u, \quad (2)$$

where typically the state vector $x = (x_1, \dots, x_n)^T \in \mathcal{X} = \mathbb{R}^n$ is available for measurement. However, despite the simplicity of the systems under consideration, little attention has been paid to performance and uncertainty aspects of the designs – eg. for which functional uncertainties are the designs stable? Can non-asymptotic performance, ie. transient performance, be estimated a-priori? Can two schema using the same adaptive laws but different model representations be compared analytically? (for example adaptive laws for linear parametrisations can be applied to eg. B-spline models, radial basis function models, polynomial approximants: the same adaptive laws are being used, but with different model representations). Does the number of parameters used in the model detrimentally (or otherwise) affect the performance?

To date bounds on performance measures for adaptive controllers which include control effort terms have remained elusive in the adaptive control literature. This paper presents the first bound on L^2 control and state performance measures. The only related literature is [6] where ideas from inverse optimality are used. In that paper, a special adaptive controller is constructed for strict feedback systems for which a ‘meaningful’ cost functional can be exhibited to which the controller can be shown to be optimal. However, the cost functionals constructed depend in a complicated fashion on the state, control and parameter estimate; and the interpretation of such performance measures is not clear. In this paper we fix the performance measure a-priori, and then produce a bound. Thus these results *are* interpretable although not optimal. It should be noted that there is a richer literature which concerns only the transient state performance, see eg. [5], or for results in the presence of disturbances see eg. [12].

The early schemes presented by [9] and [11] form the basis for much of the later work in approximate adaptive control. These schemes incorporate a sliding mode control to return the system to the region where the model approximates the system well; and thus global convergence results can be given. Semiglobal convergence results can be found in [1] and [4]. Here global results are sacrificed, but bounds on the allowable parameter and state initial conditions are given to guarantee stability. In all these results the allowable system functional uncertainty is defined via the parameteric uncertainty in the model, and not vice versa; to compare the performance of different adaptive designs, then it is necessary to formulate the system uncertainty independently of the model chosen. In this paper we formulate functional models of uncertainty and an L^2 -type cost functional. A semiglobal and a global result are given; in both cases the worst case cost is bounded in terms of the uncertainty models and the approximate model basis. Finally we discuss how these costs scale when the

resolution of the adaptive model basis is increased.

Notation. The weight space and the state space are denoted by $\mathcal{W} = \mathbb{R}^m$, $\mathcal{X} = \mathbb{R}^n$ respectively. $L^p(\Omega)$ denotes the standard Lebesgue space over Ω . $C(\Omega)$ is the normed space of continuous functions on Ω , with the uniform norm. General inner product spaces are denoted by $\mathcal{H}$, and the inner product is denoted by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. The weighted inner product space $L^2(\Omega; w)$ has the inner product: $\langle f, g \rangle = \int_{\Omega} f(x)g(x)w(x) dx$. The weighted L^∞ space $L^\infty(\Omega; w_\infty)$ has norm $\|f\|_{L^\infty(\Omega; w_\infty)} = \|f(\cdot)w_\infty(\|\cdot\|)\|_{L^\infty(\Omega)}$; note that the weight w_∞ is (non-standardly) defined as a decreasing function $w: \mathbb{R} \rightarrow \mathbb{R}$. Norms for various spaces $\mathcal{F}$ will be denoted as $\|\cdot\|_{\mathcal{F}}$, 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 $\sqrt{x^T R x}$ of vector x . The unit matrix will be denoted by I . The maximum and minimum eigenvalues of a matrix R denoted by $\bar{\lambda}(R), \underline{\lambda}(R)$ respectively. $\partial\Omega$ denotes the topological boundary of $\Omega \subset \mathcal{X}$, $\Omega^\bullet$ the interior and $\bar{\Omega}$ denotes the closure. $m(\Omega)$ denotes the Lebesgue measure of Ω .

2 System Uncertainty and Performance Measures

We consider systems in the normal form of equations 1, 2 where the state vector $x \in \mathcal{X}$ is available for measurement and the task is to stabilise x to a small neighbourhood of 0. We will assume the function f is known only to within some uncertainty set Δ , such an uncertain system is denoted Σ_Δ . f will typically be constrained a-priori to lie within a fixed smoothness class K ,¹ and the general uncertainty set we consider is:

$$\Delta(\mathcal{F}, \delta) = \{f \in K \mid \|f\|_{\mathcal{F}} \leq \delta\} \quad (3)$$

where $\mathcal{F}$ will typically taken to be an inner product function space $\mathcal{H}$ ($L^2(\Omega), L^2(\mathcal{X}), L^2(\mathcal{X}; w)$ etc.). In order to bound control effort terms in our performance measure we will need additional L^∞ uncertainty information, we will thus also be interested in the uncertainty set $\Delta(L^\infty(\mathcal{X}; w_\infty), \delta_\infty)$.

The notion of well posedness of the system controller pair is given by:

Definition 2.1. A system Σ_f , controller Ξ interconnection (Σ_f, Ξ) is said to be well posed if the system state x , the control u and the internal signals of the controller are defined, bounded, and a.e. continuous for all $t \geq 0$.² The interconnection (Σ_Δ, Ξ) is said to be well posed if (Σ_f, Ξ) is well posed for all $f \in \Delta$.

Typically approximate adaptive control schemes achieve a convergence to some residual set Ω_∞ containing the origin; asymptotic convergence is prevented by the ‘disturbance’ term which arises from the ‘inherent approximation’ error.

¹The reason for a smoothness requirement is so that standard approximation theory can be used to estimate suitable sizes of the model from particular smoothness classes.

²Note there is no requirement for the uniqueness of solutions

Natural performance measures such as linearly weighted L^2 norms of the state vector will therefore not be defined. The performance measure we consider in this paper is therefore modified to compensate:

$$\mathcal{P} = \mathcal{P}(Q, k; \Omega_0) = \sup_{f \in \Delta} \mathcal{P}_f(Q, k, \Omega_0) = \sup_{f \in \Delta} \sup_{\text{solns } (\Sigma_f, \Xi)} \int_{T_1} x^T(t) Q x(t) + k u^2(t) dt, \quad (4)$$

where Q is a positive definite, symmetric matrix; $(x(t))$ denotes a solution of the (well posed) system (Σ_f, Ξ) , (Ξ denotes the controller); $u(t)$ is a control implemented by the controller Ξ , and T_1 is defined:

$$T_1 = \{t \geq 0 \mid x_f(t) \notin \Omega_0\}, \quad (5)$$

where $\Omega_\infty \subset \Omega_\bullet$. We call $\Omega_\bullet$ the error set; it is taken to be a closed neighbourhood of the origin. The second supremum in Equation 4 is taken over all solutions of the closed loop system (Σ_f, Ξ) , and is required because of a possible loss of uniqueness of solutions, since, due to the use of a dead zone in the adaptive law, our differential equations have discontinuous right hand sides.³

2.1 Adaptive Control Solution

The key idea in model-based adaptive neuro-control is to replace the functional uncertainty ($f \in \Delta$) with a parametric uncertainty ($f \approx \theta^T \phi$, $\theta \in \Delta' \subset \mathcal{W}$, where ϕ is a known basis), and then use parametric adaptive control ideas to estimate the parameters $\theta \in \mathcal{W}$ and implement the control. Thus we first rewrite the equation 2 in the form:

$$\dot{x}_n = \theta_f^T \phi(x) + u + d_f(x) \quad (6)$$

where $d(x) = d_f(x) = f(x) - \theta_f^T \phi(x)$ and where $\theta_f \in \mathcal{W} = \mathbb{R}^m$ is fixed given f . Throughout we assume that the ‘ideal’ model $\theta_f^T \phi(x)$ is accurate enough that various measures of the approximation error (d_f) are small. The measures of the magnitude of the approximation error are given by:

$$g = g(\Omega, \Omega_0) = \sup_{f \in \Delta} \sup_{x \in \Omega \setminus \Omega_0} \frac{|x^T \mathbf{b} d_f(x)|}{\|x\|^2}, \quad (7)$$

$$q = q(\Omega, \Omega_0) = \sup_{f \in \Delta} \|d_f\|_{\mathcal{H}(\Omega \setminus \Omega_0)}, \quad s = s(\Omega, \Omega_0) = \sup_{f \in \Delta} \|d_f\|_{C^0(\Omega \setminus \Omega_0)}, \quad (8)$$

where the approximation region $\Omega \subset \mathcal{X}$ is not necessarily a compact set, and $\mathbf{b}$ is a weighting vector. Standard approximation theory provides uniform and $\mathcal{H}$ bounds for d for different approximants, and for different smoothness classes K . These uniform bounds can in turn be used to estimate the ‘size’ of approximant required to achieve specified disturbance bounds (s, q, g) as above.

The control is taken to be:

$$u(x) = -\hat{\theta}^T \phi(x) - a^T x, \quad (9)$$

³We adopt the notion of a solution as in the theory of differential inclusions [2]. It was noted in [7] that the rigorous treatment of dead-zones must require the use of such theory, and that most treatments have ignored these issues.

where $\hat{\theta}$ is an adaptive estimate and $a = [a_1, \dots, a_n]^T$ is chosen such that the matrix:

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_1 & -a_2 & -a_3 & \dots & -a_n \end{bmatrix} \quad (10)$$

is Hurwitz. The x subsystem is then written in the form:

$$\dot{x} = Ax + ((\theta_f - \hat{\theta})^T \phi(x) + d(x))e_n, \quad x(0) = x_0 \quad (11)$$

where $e_n = (0, \dots, 0, 1)^T$. The dynamics of the estimate $\hat{\theta}$ of θ_f are given by an adaptive law of the form:

$$\dot{\hat{\theta}} = \alpha x^T b D(\Omega_\bullet, x) G \phi(x), \quad \hat{\theta}(0) = 0 \quad (12)$$

where G is a positive definite and symmetric, adaptive structure matrix, b is a weighting vector, α is a scalar (the adaption gain) and D is the dead-zone function, defined by:

$$D(\Omega_0, x) = \begin{cases} 0, & \text{when } x \in \Omega_0, \\ 1 & \text{when } x \notin \Omega_\bullet. \end{cases} \quad (13)$$

Equations 9,12 define the controller $\Xi(G, \alpha, \phi, \Omega_0, A, b)$. The closed loop interconnection (Σ_f, Ξ) is defined by equations 11, 12.

The following assumptions are assumed to hold throughout this paper. First we let P be the positive definite, symmetric solution of the Lyapunov equation $A^T P + P A = -Q$, we take $b = (P^T + P)e_n$, and $\Omega_0 = \{x \in \mathcal{X} \mid x^T P x \leq \eta^2\}$ for some $\eta > 0$. To ensure well posedness of (Σ_Δ, Ξ) we make the technical assumptions that $f \in \Delta \subset K$ is contained in the set of locally Lipschitz continuous functions and that $\theta^T \phi$ is defined and locally Lipschitz continuous in $x \forall \theta \in \mathcal{W}$. Finally we assume that $\phi_i \in \mathcal{H}$.

3 Performance Bound

Let us first introduce two functions of the model basis ϕ , which measure structural properties of the basis. We define $\mu: \mathbb{R} \rightarrow \mathbb{R}$ by:

$$\mu(r) = \sup_{\eta \leq \|x\|_\bullet \leq r} \|\phi(x)\|_2, \quad (14)$$

this captures the pointwise spatial behaviour of the basis. Secondly, we let R denote the Gram matrix of the basis:

$$R = \{r_{ij}\}_{1 \leq i, j \leq m} = \{\langle \phi_i, \phi_j \rangle_{\mathcal{H}(\Omega \setminus \Omega_0)}\}_{1 \leq i, j \leq m} \quad (15)$$

The Gram matrix enables us to move between the $\mathcal{H}$ functional uncertainty and the induced parametric uncertainty in the model; $\underline{\lambda}(R)$ is a critical value for the stability analysis.

We now give the central results, these bound the worst case performance measure $\mathcal{P}$ for the system (Σ_Δ, Ξ) in terms of the uncertainty models and the function μ and the matrix R .

Theorem 3.1. *Let $\Omega \subset \mathcal{X}$ be a fixed closed set. Consider the system Σ_Δ with functional uncertainty $\Delta \subset \Delta(\mathcal{H}(\Omega), \delta) \cap \Delta(L^\infty(\Omega; w_\infty), \delta_\infty)$ and suppose $x_0 \in \Omega^\bullet$. Consider the performance measure $\mathcal{P} = \mathcal{P}(Q, k, \Omega_0)$. Implement the controller $\Xi(G, \alpha, \phi, \Omega_0, A, b)$ where b, A, P, Q, Ω_0 are as Section 2 and ϕ is a finite dimensional basis. Suppose i) $\Omega_0 \subset \Omega$, ii) $g(\Omega, \Omega_0) < \underline{\lambda}(Q)$, iii) $q(\Omega, \Omega_0), s(\Omega, \Omega_0) < \infty$, and let*

$$W_\alpha = \max\{x_\bullet^T P x_\bullet, \eta^2\} + \frac{1}{2\alpha} \frac{(\delta + q(\Omega, \Omega_0))^2}{\underline{\lambda}(G)\underline{\lambda}(R)}, \quad (16)$$

where R is the $\mathcal{H}(\Omega \setminus \Omega_0)$ Gram matrix of the model ϕ . Then for all adaption gains $\alpha > 0$ such that iv) $\{x \in \mathcal{X} \mid x^T P x \leq W_\alpha\} \subset \Omega^\circ$, we have: (Σ_Δ, Ξ) is well posed, $x(t) \rightarrow \Omega_0$; $x(t) \in \Omega$ for all $t \geq 0$, and:

$$\mathcal{P}(Q, k, \Omega_0) \leq \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g} (W_\alpha - \eta^2) + k \frac{\bar{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)} C_\alpha \quad (17)$$

where

$$C_\alpha = \int_{\eta^2}^{W_\alpha} \left(\frac{\|e_n^T A\| \sqrt{v}}{\sqrt{\underline{\lambda}(P)}} + \sqrt{2\alpha \bar{\lambda}(G) v} \mu(\sqrt{v}) + \frac{\delta_\infty}{w_\infty \left(\sqrt{\frac{v}{\underline{\lambda}(P)}} \right)} + s(\Omega, \Omega_0) \right)^2 dv. \quad (18)$$

Proof. Define $V: \mathcal{X} \times \mathcal{W} \rightarrow \mathbb{R}$ by:

$$V(x, \hat{\theta}) = x^T P x + \frac{1}{2\alpha} (\theta_f - \hat{\theta})^T G^{-1} (\theta_f - \hat{\theta}), \quad (19)$$

and by a gentle abuse of notation, let $V(t) = V(x(t), \hat{\theta}(t))$. Then $V_0 \leq W_\alpha$, since:

$$(\theta_f - \hat{\theta}(0))^T G^{-1} (\theta_f - \hat{\theta}(0)) = \theta_f^T G^{-1} \theta_f \leq \frac{1}{\underline{\lambda}(G)\underline{\lambda}(R)} \theta_f^T R \theta_f = \frac{1}{\underline{\lambda}(G)\underline{\lambda}(R)} \|\theta_f^T \phi\|_{\mathcal{H}}^2 \leq \frac{(\delta + q)^2}{\underline{\lambda}(G)\underline{\lambda}(R)}. \quad (20)$$

Now suppose $x_0 \in \Omega^\bullet \setminus \Omega_0$. By the local Lipschitz assumptions an absolutely continuous local solution $\varphi(t) = (x(t), \hat{\theta}(t))$ exists. Let $T' = [\bullet, t^*)$ be its maximal interval of existence. Now when $x(t) \in \Omega^\circ \setminus \Omega_0$,

$$\dot{V}(x(t), \hat{\theta}(t)) = x^T P \dot{x} + \dot{x}^T P x + \frac{1}{2\alpha} (\theta_f - \hat{\theta})^T G^{-1} \dot{\hat{\theta}} + \frac{1}{2\alpha} \dot{\hat{\theta}}^T G^{-1} (\theta_f - \hat{\theta})^T \quad (21)$$

$$= x^T P \left(Ax + \{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} e_n \right) + \left(Ax + \{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} e_n \right)^T P x \quad (22)$$

$$+ \frac{1}{2\alpha} (\theta_f - \hat{\theta})^T G^{-1} \dot{\hat{\theta}} + \frac{1}{2\alpha} \dot{\hat{\theta}}^T G^{-1} (\theta_f - \hat{\theta})^T \quad (23)$$

$$= -x^T Q x + x^T P \left(\{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} e_n \right) + \left(\{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} e_n \right)^T P x \\ - x^T b (\theta_f - \hat{\theta})^T \phi(x) \quad (24)$$

$$= -x^T Q x + x^T b d(x) \quad (25)$$

$$\leq -(\underline{\lambda}(Q) - g) \|x\|^2. \quad (26)$$

We now rigorously establish that that absolutely continuous solutions $(x(t), \hat{\theta}(t))$ exist for all $t \geq 0$, and that $(x(t), \hat{\theta}(t))$ is bounded by the set $L = \{(x, \hat{\theta}) \mid V(x, \hat{\theta}) \leq V_0\}$. Define $T'_1 = \{t \in [0, t^*) \mid x(t) \notin \Omega^\circ \setminus \Omega_0\}$. Since $V(t) \geq 0$ and by inequality 26, $\dot{V} \leq -(\lambda(Q) - g)\eta^2$, it follows that $m(T'_1) < \infty$ (measurability of T'_1 follows similarly to that of T_1). Suppose $\tau \in [0, t^*)$ is such that a) $x(\tau) \in \Omega^\circ \setminus \{x \mid x^T P x \leq V_0\}$, and b) $x(t) \in \Omega^\circ \forall t \in [0, \tau]$. Since $\varphi, x, \hat{\theta}$ are absolutely continuous on $[0, t^*)$, it follows that V is absolutely continuous on $[0, t^*)$. Hence (eg [8] theorem 7.18), V is differentiable a.e., $\dot{V} \in L^1[0, t^*)$, and:

$$V(\tau) = V(0) + \int_0^\tau \dot{V}(t) dt = V(0) + \int_{F_1} \dot{V}(t) dt + \int_{F_2} \dot{V}(t) dt + \int_{F_3} \dot{V}(t) dt, \quad (27)$$

where $F_1 = T'_1 = \{t \in [0, t^*) \mid x(t) \in \mathcal{X} \setminus \Omega_0\}$, $F_2 = \{t \in [0, t^*) \mid x(t) \in \Omega_0^\circ\}$ and $F_3 = \{t \in [0, t^*) \mid x(t) \in \partial\Omega_0\}$. F_1, F_2 are measurable (as $\Omega_0^\circ, \mathcal{X} \setminus \Omega_\bullet$ are open and φ is continuous); F_3 is measurable as it is the complement of $F_1 \cup F_2$ in $[0, t^*)$. We are going to estimate all the three integrals in 27. Since $\dot{V}(t) \leq 0 \forall t \in F_1$, (as inequality 26 holds, by iv) and by the fact that $V_0 \leq W_\bullet$), it follows that $\int_{F_1} \dot{V}(t) dt \leq 0$. Assume $x(\tau) \notin \Omega_0^\circ$, then we can write $F_2 = \cup_{\bullet \in A} G_\bullet = \cup_{n \geq 1} \cup_{\bullet \in A_n} G_\bullet$ where $G_\bullet = (t_\bullet^-, t_\bullet^+)$ are maximal disjoint open intervals with $x(t_\bullet^-), x(t_\bullet^+) \in \partial\Omega_0$; $A_n = \{\bullet \in A \mid m(G_\bullet) \geq 1/n\}$. As $m(F_2) < \infty$, the cardinality of each A_n is finite, hence by the dominated convergence theorem ($\dot{V} \in L^1[0, t^*)$):

$$\int_{F_2} \dot{V}(t) dt = \lim_{n \rightarrow \infty} \int_{\cup_{\bullet \in A_n} G_\bullet} \dot{V}(t) dt = \lim_{n \rightarrow \infty} \sum_{\bullet \in G_n} V(t_\bullet^-) - V(t_\bullet^+) = 0, \quad (28)$$

since by definition of the dead-zone $x(t_\bullet^-)^T P x(t_\bullet^-) = x(t_\bullet^+)^T P x(t_\bullet^+) = \eta^2$, and $\dot{\hat{\theta}}(t) = 0 \forall t \in G_\bullet$ so $V(t_\bullet^-) = V(t_\bullet^+)$. We decompose $F_3 = F_4 \cup F_5$, where $F_4 = \{t \in F_3 \mid D_- x(t)^T P x(t) = 0\}$, $F_5 = F_3 \setminus F_4$. (Note that $D_- x(t)^T P x(t)$ is defined only a.e. (as $\frac{d}{dt} x^T P x$ is defined a.e.) but is measurable a.e. so F_4 and F_5 are measurable). Write $T'_1 = \cup_{p \in P} E_p$ where $E_p = (t_p^-, t_p^+)$ are maximal disjoint connected subsets of $\mathbb{R}$ (this can be done since $x(t)$ is continuous). If $t_0 \in F_5$ then $x(t)$ cannot slide along $\partial\Omega_0$ locally from below t_0 since then (from the definition of Ω_0) $D_- x^T P x = 0$ 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_p for some p . Since the E_p 's are disjoint open intervals P must be countable and so $m(F_5) = 0$. Now $\forall t \in F_4$, $\dot{\hat{\theta}} = \nu_\bullet x^T b G^{-1} \phi(x)$ for some $0 \leq \nu \leq 1$, $\nu = \nu(t)$, and

$$0 = \frac{d}{dt} x^T P x = x^T P \left(Ax + \{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} e_n \right) + \left(Ax + \{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} e_n \right)^T P x \quad (29)$$

$$= -x^T Q x + \{(\theta_f - \hat{\theta})^T \phi(x) + d(x)\} x^T b, \quad (30)$$

so $x^T b (\theta_f - \hat{\theta})^T \phi(x) = x^T Q x - x^T b d(x)$, hence (see 24)

$$\dot{V} = -\nu x^T b (\theta_f - \hat{\theta})^T \phi(x) = -\nu (x^T Q x - x^T b d(x)) \leq -\nu (\lambda(Q) - g) \eta^2 \leq 0. \quad (31)$$

Hence similarly to F_1 , $\int_{F_3} \dot{V}(t) dt = \int_{F_4} \dot{V}(t) dt \leq 0$. Thus we have shown that $V(\tau) \leq V(0)$, thus contradicting the existence of τ . Hence, $\forall t \in [0, t^*)$, $V(t) \leq V(0) = V_0$, thus $\varphi(t)$ is bounded by L on $[0, t^*)$. Since L is bounded, it follows that the solutions can be extended, hence $t^* = \infty$ [2].

The well posedness of (Σ, Ξ) now follows from the observation that u can be bounded as a function of φ – eg see equation 38 later. It also follows that x never leaves Ω since $\|x(t)\|_P \leq V(x(t), \hat{\theta}(t)) \leq V(x(\bullet), \bullet)$ for all $t \geq 0$ and therefore by condition iv) we then have that $x \in \Omega$. Thus, we have also shown that inequality 26 holds $\forall t \in T_1 = T'_1$.

Now we bound $\mathcal{P}$. By 25:

$$\mathcal{P}_f = \int_{T_1} x^T(t) Q x(t) + k u^2(t) dt = \int_{T_1} -\dot{V} dt + \int_{T_1} x^T b d(x) dt + k \int_{T_1} u^2(t) dt. \quad (32)$$

Let us consider the first term. We write $T_1 = \cup_{p \in P} E_p$ where $E_p = (t_p^-, t_p^+)$ are maximal disjointed connected subsets of $\mathbb{R}$ (this can be done since $x(t)$ is continuous), and define $P_n = \{p \in P \mid m(E_p) \geq 1/n\}$. Since:

$$\int_{\cup_{p \in P_n} E_p} -\dot{V}(t) dt = \sum_{p \in P_n} V(t_p^-) - V(t_p^+) \leq V_0 - \eta^2, \quad (33)$$

by taking the limit as $n \rightarrow \infty$, and applying the monotone convergence theorem we can bound the first term of 32 by $V_0 - \eta^2$.

Let the sum of the first two terms in 32 be denoted by $\mathcal{P}'$. Then since $V_0 = V(0)$, and $|x^T b d(x)| \leq g \|x\|^2 \leq g \|x\|_Q^2 / \underline{\lambda}(Q)$, we have:

$$\mathcal{P}' \leq (V_0 - \eta^2) + \int_{T_1} x^T b d(x) dt \leq (V_0 - \eta^2) + \frac{g}{\underline{\lambda}(Q)} \int_{T_1} \|x\|_Q^2 dt \leq (V_0 - \eta^2) + \frac{g}{\underline{\lambda}(Q)} \mathcal{P}'. \quad (34)$$

Rearranging the terms yields

$$\mathcal{P}' \leq \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g} (V_0 - \eta^2), \quad (35)$$

so we are left with estimating the third integral. The idea is to change the integration over T_1 to an integration over $[\eta^2, V_0]$ as a function of $V(t)$. Let $0 < \gamma < 1$, and define T^j as the time set where:

$$V_0 - j\gamma\eta^2 \leq V(t) \leq V_0 - (j-1)\gamma\eta^2, \quad j = 1, 2, \dots, J_\gamma, \quad (36)$$

where J_γ is defined to be the least integer larger than $\frac{V_0 - \eta^2}{\gamma\eta^2}$. Then $T_1 \subseteq \cup_{j=1}^{J_\gamma} T^j$ since $x \notin \Omega_\bullet$ implies $V(t) > \eta^2$. First, we would like to bound $m(T_1 \cap T^j)$. By inequality 26, and since V is decreasing on T_1 , we have:

$$\begin{aligned} m(T_1 \cap T^j) &= \int_{T_1 \cap T^j} 1 dt \leq \frac{\bar{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)} \int_{T_1 \cap T^j} -\dot{V} dt \\ &= \frac{\bar{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)} \left((V_0 - (j-1)\gamma\eta^2) - (V_0 - j\gamma\eta^2) \right) = \gamma \frac{\bar{\lambda}(P)}{\underline{\lambda}(Q) - g}. \end{aligned} \quad (37)$$

Now assume that $t \in T_1 \cap T^j$, and define $c_j = \sqrt{V_0 - (j-1)\gamma\eta^2}$. From the definition of u :

$$u^2(t) \leq (|e_n^T A x(t)| + |\hat{\theta}^T \phi(x(t))|)^2, \quad (38)$$

$$\begin{aligned} &\leq \left(\|e_n^T A\| \frac{c_j}{\underline{\lambda}(P)^{1/2}} + |(\hat{\theta} - \theta_f)^T \phi(x(t))| + |f(x(t)) - d(x(t))| \right)^2, \\ &\leq \left(\|e_n^T A\| \frac{c_j}{\underline{\lambda}(P)^{1/2}} + \mu(\|x(t)\|_P) \sqrt{2\alpha} \left(\frac{\frac{1}{2\alpha} \|\hat{\theta} - \theta_f\|_{G^{-1}}^2}{\underline{\lambda}(G^{-1})} \right)^{\frac{1}{2}} + s(\Omega, \Omega_0) + \frac{\delta_\infty}{w_\infty(\frac{\|x(t)\|_P}{\sqrt{\underline{\lambda}(P)}})} \right)^2, \\ &\leq \left(\|e_n^T A\| \frac{c_j}{\underline{\lambda}(P)^{1/2}} + \mu(c_j) c_j \sqrt{2\alpha \bar{\lambda}(G)} + s(\Omega, \Omega_0) + \frac{\delta_\infty}{w_\infty(\frac{c_j}{\sqrt{\underline{\lambda}(P)}})} \right)^2. \end{aligned} \quad (39)$$

which is independent of t , and therefore by 37

$$\begin{aligned} \int_{T_1} u^2(t) dt &= \sum_{j=1}^{J_\gamma} \int_{T_1 \cap T^j} u^2(t) dt \\ &\leq \sum_{j=1}^{J_\gamma} \frac{\bar{\lambda}(P)}{\bar{\lambda}(Q) - g} \left(\frac{\|e_n^T A\| c_j}{\sqrt{\bar{\lambda}(P)}} + \mu(c_j) c_j \sqrt{2\alpha \bar{\lambda}(G)} + s(\Omega, \Omega_0) + \frac{\delta_\infty}{w_\infty(\frac{c_j}{\sqrt{\bar{\lambda}(P)}})} \right)^2. \end{aligned}$$

Now considering the summation as a Riemann sum, and letting $\gamma \rightarrow 0$, we obtain:

$$\begin{aligned} \int_{T_1} u^2(t) dt &\leq \frac{\bar{\lambda}(P)}{\bar{\lambda}(Q) - g} \int_0^{\frac{V_0 - \eta^2}{\eta^2}} \left(\frac{\|e_n^T A\| \sqrt{V_0 - s\eta^2}}{\sqrt{\bar{\lambda}(P)}} + \mu(\sqrt{V_0 - s\eta^2}) \sqrt{2\alpha \bar{\lambda}(G)} (V_0 - s\eta^2) \right. \\ &\quad \left. + s(\Omega, \Omega_0) + \frac{\delta_\infty}{w_\infty(\sqrt{\frac{V_0 - s\eta^2}{\bar{\lambda}(P)}})} \right)^2 ds \\ &= \frac{\bar{\lambda}(P)}{\eta^2(\bar{\lambda}(Q) - g)} \int_{\eta^2}^{V_0} \left(\frac{\|e_n^T A\| \sqrt{v}}{\sqrt{\bar{\lambda}(P)}} + \mu(\sqrt{v}) \sqrt{2\alpha \bar{\lambda}(G)} v + s(\Omega, \Omega_0) + \frac{\delta_\infty}{w_\infty(\sqrt{\frac{v}{\bar{\lambda}(P)}})} \right)^2 dv \end{aligned}$$

Finally, inequality 20 completes the bound. Finally we show that $x(t) \rightarrow \Omega_0$. Since $m(T_1) < \infty$ (eg. from equation 37) and u is bounded, the definition of Σ_f implies $\dot{x}$ is bounded and thus $x(t)$ is uniformly continuous. If $m = \infty$, then since $m(T_1) < \infty$, we get that $\lim_{i \rightarrow \infty} (t_{2i+1} - t_{2i}) = 0$, and so, by the uniform continuity of $x(t)$, we have that $x(t) \rightarrow \Omega_0$; the assertion is trivial if $m < \infty$. This completes the proof when $x_0 \in \Omega \setminus \Omega_0$. If $x_0 \in \Omega_0$ then either x remains in Ω_0 forever, in which case $T_1 = \emptyset$, $\mathcal{P}_f = 0$, all signals are defined and bounded and the conclusions of the theorem hold. Otherwise let τ be the initial time instant when x leaves Ω_0 (Ω_0 is closed and $x(t)$ is continuous so τ exists). Then $V(\tau) = V(x(\tau), \hat{\theta}(\tau)) = V(x(\tau), \hat{\theta}(\bullet))$ since $\hat{\theta}$ is constant while $x \in \Omega_0$, and thus $V(\tau) = x^T(\tau) P x(\tau) + (1/2\alpha) \theta_f^T G^{-1} \theta_f = \eta^2 + (1/2\alpha) \theta_f^T G^{-1} \theta_f = \eta^2 + V(0, 0) \leq W_\alpha$ and so $x(\tau) \in \Omega$. So, by applying the above arguments with $x_\bullet = x(\tau)$ we obtain the required result as $V_0 = V(x(\tau), 0)$, thus completing the proof. $\square$

Theorem 3.1 is given weight by the fact that finite dimensional models satisfying ii) and iii) always exist for compact Ω , and since $x_0 \in \Omega^\circ$, it follows that for all sufficiently large adaption gains α , that condition iv) holds. But this comes at a price. If $k > 0$ then the upper bound on the performance growth is $O(\alpha)$. The increase of $\mathcal{P}_f$ is solely due to the increased control effort, if $k = 0$ then there is no increase, for $\mathcal{P}_f = \text{const.} + O(1/\alpha)$. It should also be observed that the L^∞ uncertainty model is only required to bound the control effort; it is neither required to bound the state transient, or to guarantee stability.

To apply Theorem 3.1 for stabilisation on compacta (ic. when Ω is compact), it is necessary to know the uncertainty level δ , in order to ensure, via a suitable selection of the adaption gain α , that the transient state vector lies within the model's approximation range. This is an atypical assumption in adaptive control; universality to the uncertainty level is a desirable feature of an adaptive design. In order to apply Theorem 3.1 to systems with arbitrary uncertainty levels, it is clearly necessary to take Ω to be global. This then calls for infinite dimensional models since finite dimensional models cannot suffice for global approximation.

There are two further issues that must be considered when utilizing infinite dimensional models. The first is the restriction that only a finite number of parameter updates can be updated at any given time, so that the controller is physically realisable. We satisfy this requirement by restricting our attention to Semiglobal Finite Dimensional (SFD) models, which have the property that for any compact set $\Omega \subset \mathcal{X}$ only a finite number basis functions have supports intersecting Ω . A widely known example of an SFD model is a B-spline basis.

The second issue concerns the rate of growth of the uncertainty. Spatial growth information is built into the basis function system in the parametric adaptive case, simply because the basis functions themselves reflect the nonlinearity of the system. In the case of approximate models, the basis functions do not generally have the order of growth of the system uncertainty, and hence the adaptive law may not be ‘fast’ enough⁴, as illustrated by considering the system $\Sigma : \dot{x} = f(x) + u$, where $f(x) = \sum_{i=\bullet}^{\infty} \theta_i^* \phi_i(x)$ where the uncertainty is in the parameters θ_i^* . Suppose further that P1) $0 \leq \phi_i(x) \leq 1$, where P2) $\text{supp } \phi_i = [i/2, i/2 + 1]$ and P3) $\phi_i(x) + \phi_{i-1}(x) \geq 1/2$ and take $\theta_i^* = 2 + (i + 1)^2$.⁵ If we consider the adaptive controller $\Xi(I, \alpha, \phi, [-1, 1], -1, 1)$ it is straightforward to show (see [3]), that for any $\alpha > 0$ and for any large initial condition $x(0) = x_{\bullet} > 0$ we have $x(t) \rightarrow \infty$ as $t \rightarrow \infty$. This is despite the fact that the model’s domain is global (it can even exactly reconstruct the nonlinearity globally!). This example highlights that it does not suffice simply to ensure that the model’s approximation domain is global in order to obtain global results; it is also necessary that the adaptive updates reflect the uncertainty growth in some manner.

In order to state the global result one further piece of notation is required: $\phi|_{\Omega}$ is the model basis which contains only the basis functions of ϕ whose supports intersect Ω . $G|_{\Omega}$ is the corresponding restriction of G .

Corollary 3.1.1. *Consider the system Σ_{Δ} with functional uncertainty $\Delta \subset \Delta(L^2(\mathcal{X}; w), \delta) \cap \Delta(L^{\infty}(\Omega; w_{\infty}), \delta_{\infty})$ where $\delta > 0$ is unknown. Suppose $x_0 \in \mathcal{X}$ is the initial condition. Consider the performance measure $\mathcal{P}(Q, k, \Omega_0)$. Partition an SFD model ϕ as $[\phi^1 | \phi^2 | \phi^3 | \dots]$, and partition θ equivalently: $\theta = [\theta^1 | \theta^2 | \theta^3 | \dots]$, where for each i θ^i is a finite-dimensional vector. Let R_i be the $L^2(\mathcal{X}, w)$ Gram matrix for the model ϕ^i , let $\{a_i\}_{i \geq 1}$ be any sequence such that $\sum_{i=1}^n 1/a_i < \infty$, define*

$$G_i = \frac{a_i}{\underline{\lambda}(R_i)} I_{\dim(\theta^i) \times \dim(\theta^i)}, \quad (40)$$

and let $G^ = \text{diag}[G_1, G_2, \dots]$. Suppose i) $\Omega_0 \subset \Omega$, ii) $g(\Omega, \Omega_{\bullet}) < \underline{\lambda}(Q)$, and iii) $\mathbf{q}(\Omega, \Omega_0), s(\Omega, \Omega_0) < \infty$. Implement the controller $\Xi(G^*, \alpha, \phi, \Omega_0, A, b)$ where $\alpha > \bullet$. Then the interconnection (Σ_{Δ}, Ξ) is well posed, $x(t) \rightarrow \Omega_0$ as $t \rightarrow \infty$ and*

⁴This is analogous to the problem that occurs with normalised estimation laws vs. Lyapunov update laws with finite dimensional uncertainty.

See eg. [5], where it is shown that the linear growth assumption of eg [10] is needed for the adaptive law to be ‘fast’ enough; in contrast with Lyapunov update laws which are automatically ‘fast’ enough.

⁵These assumptions can be realised by eg. a first order B-spline.

only a finite number of parameters $\hat{\theta}_i$ are adapted. Now define $\Omega = \{x \in \mathcal{X} \mid x^T P x \leq W_\alpha\}$, where

$$W_\alpha = \max\{x_0^T P x_0, \eta^2\} + \frac{(\delta + q)^2}{2\alpha}. \quad (41)$$

Then $x(t) \in \Omega$ for all $t \geq 0$ and:

$$\mathcal{P}(Q, k, \Omega_0) \leq \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g} (W_\alpha - \eta^2) + k \frac{\bar{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)} C_\alpha \quad (42)$$

where

$$C_\alpha = \int_{\eta^2}^{W_\alpha} \left(\frac{\|e_n^T A\| \sqrt{v}}{\sqrt{\underline{\lambda}(P)}} + \sqrt{2\alpha \bar{\lambda}(G|_\Omega) \mu(\sqrt{v}) \sqrt{v}} + \frac{\delta_\infty}{w_\infty \left(\sqrt{\frac{v}{\underline{\lambda}(P)}} \right)} + s(\Omega, \Omega_0) \right)^2 dv. \quad (43)$$

Proof. First we obtain the estimate:

$$\theta_f^T G^{-1} \theta_f = \sum_{i=1}^{\infty} \theta_f^{iT} G_i^{-1} \theta_f^i \leq \sum_{i=1}^{\infty} \frac{\bar{\lambda}(G_i^{-1})}{\underline{\lambda}(R_i)} \theta_f^{iT} R_i \theta_f^i \leq \sum_{i=1}^{\infty} \frac{(\delta + q)^2}{a_i} = (\delta + q)^2. \quad (44)$$

Now consider the restricted system where only the parameters $\hat{\theta}_i$ for which $\text{supp } \phi_i \cap \Omega \neq \emptyset$ are adapted. Since the model is SFD, only a finite number of such parameters exist and thus Theorem 3.1 applies. By the choice of Ω , condition iv) of Theorem 3.1 is satisfied and thus Theorem 3.1 yields that this system is stable and its performance is bounded by 42. Now observe that since by Theorem 3.1 the state vector of the system does not leave Ω the unrestricted system exactly simulates the restricted one, which yield the theorem. $\square$

4 Model Dependency

The key question which distinguishes approximate adaptive designs from their exact parametric counterparts concerns the effect on the system performance by the choice of basis functions and the dimension of the model basis. Here we briefly discuss the effect of the resolution of the model on the performance of the controller. Using models of overly high resolutions is common in applications: for example the bounds provided by approximation theory on the required model dimension are usually not tight, and the smoothness of the system functions is hard to estimate a-priori.

Under the conditions of Theorem 3.1 it follows that the sufficient conditions:

$$\liminf_{m \rightarrow \infty} \underline{\lambda}(R_m) \alpha_m > 0, \quad \limsup_{m \rightarrow \infty} \alpha_m^{\frac{1}{2}} \mu_{\phi_m}(\sqrt{W_{\alpha_m}}) < \infty \quad (45)$$

ensure a uniform bound:

$$\limsup_{m \rightarrow \infty} \mathcal{P}_{(\Sigma_\Delta, \Xi_m)}(Q, k, \Omega_0) < \infty \quad (46)$$

on the performance of the controllers $\Xi_m = \Xi(I, \alpha_m, \phi^m, \Omega_\bullet, A, b)$, where $\{\phi^m\}_{m \geq 1}$ is a basis of increasing dimension. The first condition ensures stability and uniform boundedness of the transient state performance, the second term ensures that the control performance is bounded uniformly. If the model basis is orthogonal then the dimension of the model can

be increased arbitrarily without affecting either the stability conditions i) – iv), or the state performance transient (the case $k = 0$), since $\underline{\lambda}(R_m) = 1$.

5 Summary and Further Research

We have given LQ -type performance bounds for a class of approximate model based adaptive designs. The results can be easily extended to multi-input systems with matched uncertainty and other simple variants on the control designs. Due to the dependance of the proofs largely only on the existence of quadratic Lyapunov functions, extensions are also possible to the case of unmatched uncertainty via backstepping designs, see [3].

References

- [1] F-C. Chen and C-C. Liu. Adaptively controlling nonlinear continuous time systems using multilayer neural networks. *IEEE Trans. on Automatic Control*, 39(6):1306–1310, 1994.
- [2] A.F. Filippov. *Differential Equations with Discontinuous Righthand sides*. Kluwer Academic Publishers, Netherlands, 1st edition, 1988.
- [3] M. French. *Adaptive Control of Functionally Uncertain Systems*. PhD thesis, University of Southampton, 1998.
- [4] M. French and E. Rogers. Approximate models for adaptive feedback linearization. *Int. J. Control*, 68(6):1305–1321, 1997.
- [5] M. Krstić, I. Kanellakopoulos, and P.V. Kokotović. *Nonlinear and Adaptive Control Design*. Wiley, New York, 1st edition, 1995.
- [6] Z.H. Li and M. Krstić. Optimal design of adaptive tracking controllers for non-linear systems. *Automatica*, 33(8):1459–1473, 1997.
- [7] M.M. Polycarpou and P.A. Ioannou. On the existence and uniqueness of solutions in adaptive control systems. *IEEE Trans. on Automatic Control*, 38(3):474–479, 1993.
- [8] W. Rudin. *Real and Complex Analysis*. International edition, mathematics series. McGraw-Hill, 3rd edition, 1987.
- [9] R.M. Sanner and J.J.E. Slotine. Gaussian networks for direct adaptive control. *IEEE Trans. on Neural Networks*, 3(6):837–863, 1992.
- [10] S. Sastry and A. Isidori. Adaptive control of linearizable systems. *IEEE Trans. on Automatic Control*, 34(11):1123–1131, 1989.

- [11] E. Tzirkel-Hancock. *Stable Control of Nonlinear Systems Using Neural Networks*. PhD thesis, Department of Engineering, University of Cambridge, 1992.
- [12] B. Yao and M. Tomizuka. Adaptive robust control of SISO nonlinear systems in a semi-strict feedback form. *Automatica*, 33(5):893–901, 1997.