Proceedings of the 36th
Conference on Decision & Control
San Diego, California USA • December 1997



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

M. French [1] Cs. Szepesvári [2] E. Rogers [3]


### Abstract

We consider systems satisfying a matching condition which are functionally known up to a $L^2$ measure of uncertainty. A modified $L^2$ performance measure is given, and the performance of a class of model based adaptive controllers is studied. An upper performance bound is derived in terms of the uncertainty measure and measures of the approximation error of the model. Asymptotic analyses of the bounds under increasing model size are undertaken, and sufficient conditions are given on the model that ensure the performance bounds are bounded independent of the model size.


## 1 Introduction

Despite the simplicity of the systems under consideration in neuro-control, little attention has been paid to performance and uncertainty aspects of the various solutions – eg. for which functional uncertainties are the designs stable; can transient performance be estimated a-priori? In this paper we consider a limited class of systems (feedback linearisable and satisfying a matching condition). We provide upper bounds on $L^2$ performance measures of both the state vector and the control. $L^2$ measures of uncertainty are considered: these allow us to completely specify the uncertainty under consideration independantly of the model chosen for the adaptive design, and these measures will be shown to be natural for obtaining conditions on the uncertainty for convergence of the state vector to residual sets; stability in the large; and also for bounding the state vector part of performance measures. $L^\infty$ measures will be used to bound the control effort terms in the performance measures. The first major result shows that if sufficiently high adaption gains are utilised then a sufficiently large model sufffices for semi-global stabilisation. This design requires knowledge of the $L^2$ uncertainty level; conditions for stability can be obtained and the state performance measure can be explicitly bounded. With $L^\infty$ information, the control effort performance measure can also be bounded. However if the uncertainty level is unknown the state vector transient cannot be bounded a-priori, and we must consider global models and corresponding global uncertainty measures incoporating the uncertainty growth. An example is constructed that shows that if the uncertainty growth is not known then stability cannot be guaranteed. The second major result shows that if the uncertainty growth is known, then a class of physically realisable, but non-finite dimensional (structurally adaptive) models suffices. This is proved using weighted $L^2$ descriptions of the uncertainty, and the state performance measures are bounded; under further (weighted) $L^\infty$ assumptions on the uncertainty, the control effort can also be bounded. Finally we investigate the asymptotic properties of the performance bounds, and the dependency of the performance on the model resolution structure.

**Notation** $\mathcal{W}$, $\mathcal{X}$ denote the weight space, and state space respectively, both are taken to be Euclidean spaces. $L^p(\Omega)$ denotes the standard Lebesgue space over $\Omega$; $l^p$ the standard sequence space. $C(\Omega)$ is the normed space of continuous functions on $\Omega$, with the uniform norm. General inner product spaces are denoted by $\mathcal{H}$, and the inner product is denoted by $\langle .,. \rangle_{\mathcal{H}}$. 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 unit matrix will be denoted by $I$. If the eigenvalues of a matrix $R$ are $\lambda_1, \ldots, \lambda_n$, then $\overline{\lambda}(R), \underline{\lambda}(R)$ are defined to be $\max_{1\leq i \leq n}|\lambda_i|$, $\min_{1 \leq i \leq n}|\lambda_i|$ respectively. The weighted $L^\infty$ space $L^\infty(\Omega; w)$ has norm $\|f\|_{L^\infty(\Omega,w)} = \|f(\cdot)w(\|\cdot\|_P)\|_{L^\infty(\Omega)}$; note that the weight $w$ is (non-standardly) defined as a function $w\colon \mathbb{R} \to \mathbb{R}$. 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 $x$. $\partial\Omega$ denotes the topolog-

[1] M.French@ecs.soton.ac.uk, Department of Electronics and Computer Science, University of Southampton, UK.
[2] szepes@sol.cc.u-szeged.hu, Department of Adaptive Systems and Research Group on Artificial Intelligence, Hungarian Academy of Sciences – JATE, Szeged, Hungary
[3] E.Rogers@ecs.soton.ac.uk, Department of Electronics and Computer Science, University of Southampton, UK.

ical boundary of $\Omega \subset \mathcal{X}$, $m(\Omega)$ denotes the Lebesgue measure of $\Omega$. For any real-valued function $h$ defined over $\Omega$ let $\overline{h}(r) = \sup_{\eta \leq \|x\|_P \leq r} \frac{|h(x)|}{\|x\|_P}$.

## 2 System Uncertainty, Performance Measures and Adaptive Control

We consider systems in the normal form:

$$\begin{aligned} \dot{x}_i &= x_{i+1} \quad 1 \leq i \leq n-1 \\ \dot{x}_n &= f(x) + u \end{aligned} \tag{1}$$

where the state vector $x = (x_1, x_2, \ldots, x_n)^T \in \mathcal{X} = \mathbb{R}^n$ is available for measurement and the task is to let $x \to 0$. Such a system is denoted $\Sigma_f$. We assume the function $f$ is known only to within some uncertainty set $\Delta$. Such an uncertain system is denoted $\Sigma_\Delta$. Thus the assumption that the state vector of the normal form is available for measurement is particularly strong for uncertain systems; in particular it is much stronger than the knowledge that an affine system has relative degree $n$ and the original state vector is available for measurement. Note that the results in this paper which apply to the normal form 1, and the majority of existing neuro-control results should not properly be considered as the control of affine systems; rather they are the control of systems of a specialised structure (satisfying a matching condition) with respect to a measurable state vector.
Our basic measure of uncertainty of $f$ will be with the $L^2$ norm, but $f$ will be constrained a-priori to lie within a fixed smoothness classes $K$.[1] The general uncertainty set we consider is:

$$\Delta = \Delta(\mathcal{F}, \delta) = \{f \in K \mid \|f\|_{\mathcal{F}} \leq \delta\} \tag{2}$$

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 $L^\infty$ uncertainty information, we will thus also be interested in the uncertainty set $\Delta(L^\infty(\mathcal{X}; w), \delta_\infty)$. These measures of uncertainty are the most natural for our problem.
Typically adaptive neuro-control schemes achieve a convergence to some residual set $\Omega_\infty$ containing the origin; asymptotic convergence is prevented by the 'disturbance' term which arises from the 'inherent approximation' error. Natural performance measures such as 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:

$$\mathcal{P}(Q, k, \Omega_0) = \sup_{f \in \Delta} \mathcal{P}_f = \sup_{f \in \Delta} \int_{T_1} x_f^T(t) Q x_f(t) + k u_f^2(t)\, dt, \tag{3}$$

where $Q$ is a positive definite matrix; $x_f(t)$ is a solution of the (well posed) [2] system $(\Sigma_f, \Xi)$, ($\Xi$ denotes the controller); $u_f(t)$ is the control implememted by the controller $\Xi$ for system $f$; $T_1$ is defined:

$$T_1 = \{t \geq 0 \mid x_f(t) \notin \Omega_0\}, \tag{4}$$

where $\Omega_\infty \subset \Omega_0$. We call $\Omega_0$ the error set – it is taken to be a neighbourhood of the origin. A good interpretable choice for $\Omega_0$ would be the set $\{x \in \mathcal{X} \mid x^T Q x \leq \eta^2\}$ for some choice of $\eta$ (typically $\eta$ would be small).
We now describe the adaptive control methodology. The idea in model-based adaptive neuro-control is to replace the functional uncertainty ($f \in \Delta$) with a parametric uncertainty ($f \approx \theta^T \phi$, $\theta \in \Delta'$), and then use parametric adaptive control ideas to estimate the parameters $\theta \in \mathcal{W}$ and implement the control. Thus we first rewrite the system $\Sigma_f$ in the form:

$$\begin{aligned} \dot{x}_i &= x_{i+1} \quad 1 \leq i \leq n-1 \\ \dot{x}_n &= \theta_f^T \phi(x) + u + d_f(x) \end{aligned} \tag{5}$$

where

$$d_f(x) = f(x) - \theta_f^T \phi(x), \tag{6}$$

where $x_i \in \mathbb{R}$, $x \in \mathcal{X} = \mathbb{R}^n$ and $\theta_f \in \mathcal{W} = \mathbb{R}^m$ is fixed given $f$. If $f$ is fixed then, for convenience, we will write $d$ instead of $d_f$. The Gram matrix $\langle \phi_i, \phi_j \rangle_{\mathcal{H}(\Omega \setminus \Omega_0)}$ of $\phi : \mathcal{X} \to \mathbb{R}^m$ will be denoted by $R$. Throughout we assume that the model $\theta_f^T \phi(x)$ is accurate enough that various measures of the approximation error are small. The size of the disturbance affects the performance of the system in two manners. Firstly it determines the stability and the size of $\Omega_\infty$ ($g$ below); secondly it affects the rate of convergence to $\Omega_0$, thus affecting the performance measure $\mathcal{P}$ ($q$,$s$ below). The measures of the magnitude of the disturbance are given by:

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

$$q(\Omega, \Omega_0) = \sup_{f \in \Delta(\mathcal{H}, \delta)} \|d_f\|_{\mathcal{H}(\Omega \setminus \Omega_0)} \tag{8}$$

[1] The 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.

[2] To ensure well-posedness of the interconnection $(\Sigma_\Delta, \Xi)$ we assume that $f$ is locally Lipschitz, a similar requirement is placed on $\Xi$ a.e.

and

$$s(\Omega,\Omega_0) = \sup_{f\in\Delta(\mathcal{H},\delta)} \|d_f\|_{C^0(\Omega\setminus\Omega_0)}, \qquad (9)$$

where the approximation region $\Omega \subset \mathcal{X}$ is not neccessarily a compact set, and $b$ is a weighting vector. If $\Omega$ is compact then a finite dimensional model (ie. $\dim(\mathcal{W}) < \infty$) suffices, however, if $\Omega$ is not compact, then generally an infinite model must be considered (however, as we shall see, some such models do lead to physically realisable controllers). Standard approximation theory provides uniform and $L^2$ bounds for $d$ for different approximants, and different smoothness classes $K$. These uniform bounds can in turn be used to estimate the 'size' of approximant required to achieve the specified disturbance bounds as above.
The control is taken to be:

$$u(x) = -\hat{\theta}^T\phi(x) - n(s)x_1, \qquad (10)$$

where $n(s) = s^n + \alpha_{n-1}s^{n-1} + \ldots + \alpha_1$ is a Hurwitz polynomial. The dynamics of the estimate $\hat{\theta}$ of $\theta_f$ is then given by an adaptive law of the form:

$$\dot{\hat{\theta}} = \alpha\, x^T\, b\, D(\Omega_0, x)\, G\phi(x), \quad \hat{\theta}(0) = 0 \qquad (11)$$

where $G$ is an invertible, positive definite, adaptive structure matrix, $b$ is the weighting vector, $\alpha$ is a scalar (the adaption gain) and $D$ is the dead-zone function, defined to be the characteristic function of $\mathcal{X} \setminus \Omega_0$. We further assume that $\phi_i \in \mathcal{H}$, and to ensure well posedness, we assume that the model $\theta^T\phi$ is defined and locally Lipschitz continuous in $x$ $\forall\theta \in \mathcal{W}$. Equations 10,11 define the controller $\Xi(\mathcal{H}, G, \alpha, \phi, \Omega, \Omega_0, A, b)$. The closed loop interconnection $(\Sigma_f, \Xi)$ can now be written as follows:

$$\begin{aligned} \dot{x} &= Ax + ((\theta_f - \hat{\theta})^T\phi(x) + d(x))e_n && (12)\\ \dot{\hat{\theta}} &= \alpha\, x^T\, b\, D(\Omega_0, x)\, G\phi(x), && (13) \end{aligned}$$

where $e_n = (0, \ldots, 0, 1)^T$ and

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -\alpha_1 & -\alpha_2 & -\alpha_3 & \ldots & -1 \end{bmatrix}. \qquad (14)$$

We assume throught that $P$ is the solution of the Lyapunov equation $A^TP + PA = -Q$, $b = (P^T + P)e_n$, and $\eta$ is a constant s.t. $\{x \in \mathcal{X} \mid x^TPx \le \eta^2\} \subset \Omega_0$.

## 3 Performance Bound

One further structural parameter is introduced, $\mu(\phi)$, which is defined: $\mu(\phi) = \sup_{x\in\Omega\setminus\Omega_0}\|\phi(x)\|$, (this is required to bound the control effort terms). We need one definition before we can give the main results.

**Definition 3.1** *An approximation region $\Omega$ for the model $\phi$ is said to be $(Q, b, \Omega_0)$* admissable *if: i) $\Omega_0 \subset \Omega$, ii) $g(\Omega, \Omega_0) < \underline{\lambda}(Q)$, iii) $q(\Omega, \Omega_0), s(\Omega, \Omega_0)$ are finite.*

**Theorem 3.1** *Let $\Omega \subset \mathcal{X}$ be a closed set. Consider the system $\Sigma_\Delta$ with functional uncertainty $\Delta \subset \Delta(\mathcal{H}(\Omega), \delta)$ and initial condition $x_0 \in \Omega^\circ$. Consider the performance measure $\mathcal{P} = \mathcal{P}(Q, k, \Omega_0)$. Implement the controller $\Xi(\mathcal{H}, G, \alpha, \phi, \Omega, \Omega_0, A, b)$ where $\phi$ is a finite dimensional model. Suppose $\Omega$ is $(Q, b, \Omega_0)$ admissable, and*

$$W_\alpha = \max\{x_0^TPx_0, \eta^2\} + \frac{1}{2\alpha}\frac{(\delta + q(\Omega, \Omega_0))^2}{\underline{\lambda}(G)\underline{\lambda}(R)}, \qquad (15)$$

*where $R = \langle\phi_i, \phi_j\rangle_{\mathcal{H}(\Omega\setminus\Omega_0)}$ is the Gram matrix of the model $\phi$.*
*Then for all adaption gains $\alpha$ such that $\{x \in \mathcal{X} \mid x^TPx \le W_\alpha\} \subset \Omega$, $x(t) \to \Omega_0$, and $x(t) \in \Omega$ for all $t \ge 0$. If $k = 0$ the performance measure is bounded by:*

$$\mathcal{P}(Q, 0, \Omega_0) \le \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g}(W_\alpha - \eta^2) \qquad (16)$$

*If $k > 0$, and $\Delta \subset \Delta(L^\infty(\Omega; w_\infty), \delta_\infty)$ then*

$$\mathcal{P}(Q, k, \Omega_0) \le \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g}(W_\alpha - \eta^2) + k\frac{\underline{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)}F_\alpha \qquad (17)$$

*where*

$$\begin{aligned} E_\alpha &= \frac{\|e_n^TA\|}{\sqrt{\underline{\lambda}(P)}} + \mu(\phi)(\alpha\overline{\lambda}(G))^{1/2}, \\ F_\alpha &= \int_{\eta^2}^{W_\alpha}\left(E_\alpha + \frac{\delta_\infty w_\infty^{-1}(\sqrt{v}) + s(\Omega, \Omega_0)}{\eta}\right)^2 v\,dv. \end{aligned}$$

**Proof:** (Sketch). Consider the Lyapunov function $V(x, \hat{\theta}) = x^TPx + \frac{1}{2\alpha}(\theta_f - \hat{\theta})G^{-1}(\theta_f - \hat{\theta})$. Invariance of $\Omega$ wrt. $x$, and the $V(x_0, 0)$ level set of $V$ follow from the estimate $\dot{V} \le -(\underline{\lambda}(Q) - g)\|x\|^2$ for $x(t) \in \Omega\setminus\Omega_0$ and $V(x_0, 0) \le W_\alpha$. Convergence to $\Omega_0$ follows from a uniform continuity argument. State performance is estimated by $\int_{T_1} = \int_{T_1} -\dot{V} + x^Tbd(x) \le$

$V(x_0,0)-\eta^2+\frac{g}{\underline{\lambda}(Q)}\int_{T_1} x^TQx$. Control performance is more delicate, $\int_{T_1} u^2$ is bounded by an approximation and transformation procedure which allows us to change th integral from one over $T_1$ to one over $[\eta^2, W_\alpha]$. A full proof can be found in [3]. ■

**Remark.** Theorem 3.1 is given weight by the fact that finite dimensional admissable models always exist for compact $\Omega$, and since $x_0 \in \Omega^\circ$, it follows that for all sufficiently large adaption gains $\alpha$, the condition **??** holds. Also note that whilst the control effort part of the performance bound has been given in terms of an integral, this integral can be computed a-priori from the uncertainty set assumptions.
If the system is of the form $\Sigma_{f+c}$, where

$$c \in \mathcal{C} = \{c: \Omega \to \mathbb{R} \mid x^Tbc(x) \leq 0\ \forall x \in \Omega\}, \quad (18)$$

then the above proof remains valid. Similarly, the uncertainty need only be measured on the "unstable" part of $f$, likewise the disturbance, eg.

$$d_f(x) = \mathrm{dist}\{\theta_f^T\phi(x), \mathcal{C}\}, \quad (19)$$

$$\Delta = \{f - c \in K \mid \|f-c\|_{\mathcal{H}} \leq \delta,\ c \in \mathcal{C}\}. \quad (20)$$

**Example.** (Legendre polynomials). Consider the system $\dot{x} = f(x) + u$, where $x, u \in \mathbb{R}$ and $x_0 = 1/2$. Let $\Omega = [-1,1]$, $K = \mathrm{Lip}_1 1$ and suppose we have the uncertainty $\Delta = \Delta(L^2(\Omega),1)$. We measure the performance with $\mathcal{P}(Q = I, 0, \Omega_0 = [-0.1, 0.1])$. $0 < \alpha, M)$. Take the control gain to be $A = -1$. Solving the Lyapunov equation we obtain $P = 1/2$ and so $b = 1$. Take $\phi$ as the union of two polynomial approximants of degree $n$ with ranges $[-1,-0.1]$, $[0.1, 1]$ respectively. Jackson's theorem [?] implies $s \leq 6(0.9)^2/n$ from which we can estimate also that $q \leq 9\sqrt{2}s/10$, $g \leq 243/5n$. For stability we require that $g < 1$, but for a reasonable performance we enforce $g \leq 1/2$, so we take. $n \geq 98$. In this case $s \leq 1/20$ , so $W_\alpha = 1/8 + (1+\sqrt{2}/20)^2/2\alpha$; for stability we also require $W_\alpha \leq 1$, so we choose $\alpha \geq \frac{8}{14}(1+\sqrt{2}/20)^2$ Take the polynomial model to be the Legendre basis, so $R = I$, $\underline{\lambda}(R) = 1$. Applying theorem 3.1 the performance $\mathcal{P} = \mathcal{P}(I, 0, [-0.1, 0.1])$ is bounded by:

$$\mathcal{P} \leq \frac{99n}{100n - 243} \quad n \geq 98. \quad (21)$$

As $\alpha \to \infty$ we can also show that the upper bound converges to 23/200.
If a modification of the control law is required, eg. if $g \neq 1$, so we implement the control $u(x) = \frac{1}{g(x)}(-\hat{\theta}^T\phi(x) - n(s)x_1)$, or we know a nominal system $f_0$ so we want to implement the control $u(x) = -\hat{\theta}^T\phi(x) - n(s)x_1 - f_0(x)$, then the entire control effort term would have to be estimated by the integral

$$\frac{\overline{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)} \int_{\eta^2}^{W_\alpha} \tilde{U}^2(\sqrt{v})v\,dv \quad (22)$$

where $\tilde{U}(r) = \sup_{\eta \leq \sqrt{V(x,\hat{\theta})} \leq r} \frac{U(x,\hat{\theta})}{\sqrt{V(x,\hat{\theta})}}$.

### 3.1 Global Stabilization

To apply theorem 3.1 for stabilisation on compacta, we need to know the uncertainty level in order to design the controller (so that we can ensure that the transient state vector lies within the model's range). This is an atypical requirement in parametric adaptive control, we would like to have controllers which are universal to the uncertainty level. According to theorem 3.1 for this we need $\Omega$ to become global. This calls for infinite dimensional models since we cannot hope that finite models will suffice for global approximation. But then one must be careful in choosing $G$ since in general (for good approximation) we will have that $\theta_f^{(i)} \to \infty$, $|i| \to \infty$ and so $V(x(0), 0)$ may become infinite. The following example shows an unstable case when $G$ is defined inappropriately:
**Example.** Consider the system $\Sigma$ : $\dot{x} = f(x) + u$, where $f(x) = \sum_{i=-\infty}^{\infty} \theta_i^* \phi_i(x)$ and the uncertainty is in the paramaters $\theta_i^*$. Suppose further that P1) $0 \leq \phi_i(x) \leq 1$, where P2) $\mathrm{supp}\,\phi_i = [i/2, i/2+1]$ and P3) $\phi_i(x) + \phi_{i-1}(x) \geq 1/2$ (note that here $\theta^*$ denotes the exact parameter vector). [3] Now consider the following adaptive controller:

$$\dot{\hat{\theta}}_i = \alpha x \phi_i(x), \quad i \in \mathbf{Z}, \quad (23)$$

$$u = -\sum_{i=-\infty}^{\infty} \hat{\theta}_i \phi_i(x) - x \quad (24)$$

with $\alpha > 0$ and the initial condition $\hat{\theta}_i(0) = 0$.
Take $\theta_i^* = 2 + (i+1)n(i)$, where $n(i)$ is an arbitrary increasing function with $n(i) \to \infty$ as $i \to \infty$. We show in [3] that for *any* $\alpha > 0$ there exist a number $M = M(\alpha)$ such that if $x(0) \geq M$ then $x(t) \to \infty$, i.e., the system cannot be stabilized on $\mathbb{R}$ by uniformly increasing the adaption rate.
By considering only a finite sum of (compactly supported) basis functions in the model, we can interpret the above example as: no finite model of this type suffices for stabilisation. Note that by a similar construction, it is possible to construct an unstable system for any choice of the controller paramaters: $\alpha, G, A$ etc.
Note that for the function $f$ of the above example

[3] 1st order B-splines satisfy P1-P3

$||f||_2 = \infty$ so $\Delta \not\subset: L^2(\Omega)$, ie., the set $\Delta(L^2(\mathcal{X}), \delta)$ is clearly too restrictive. Since it is natural to remain in an $L^2$ setting we suugest the consideration of uncertainty sets of form $\Delta(L^2(\mathcal{X}; w), \delta)$. The weighting function $w$ allows us capture the growth of the system uncertainty, whilst maintaining a bounded uncertainty in an appropriate space for analysis. For example, if $f(x) = O(x^n)$ then $w: \mathcal{X} \to \mathbb{R}$ can be taken to be $w(x) = ||x||^{-n-1}$. The key to dealing with global uncertainties is to increase the rate of adaption appropriately since Theorem 3.1 requires that $(\theta_f - \hat{\theta})^T G^{-1}(\theta_f - \hat{\theta})$ is defined and thus by choosing $G$ (the adaption gain matrix) appropriately we should be able to deal with a wide class of uncertainties. This then gives us a systematic method for choosing the adaption gain, structure $\alpha, G$ for stability and performance.

Note that the model in example 3.1 is *semiglobally finite dimensional* (SFD), ie., for all compact sets $\Omega \subset \mathcal{X}$, the cardinality of $\{i \in \mathbb{N} \mid \operatorname{supp} \varphi_i \cap \Omega \neq \emptyset\}$ is finite. Provided tha the system is stable such models are *physically realizable.*

A construction of a suitable $G^*$ for the SFD model $\phi = (\varphi_1, \varphi_2, \ldots, \varphi_i, \ldots)$ is as given in [2]: Partition $\phi$ by $[\phi^1|\phi^2|\phi^3|\ldots]$, and partition $\theta$ equivalently: $\theta = [\theta^1|\theta^2|\theta^3|\ldots]$, where for each $i$ $\theta^i$ is a finite-dimensional vector. Let $R_i$ be the global $w$ weighted Gram matrix for the model $\phi^i$, and define:

$$G_i = \frac{a_i}{\underline{\lambda}(R_i)} I_{\dim(\theta^i)\times\dim(\theta^i)}, \tag{25}$$

where $\{a_i\}$ is any positive sequence for which $\sum_{i=1}^{\infty} 1/a_i = 1$. Let $G^*$ be the diagonal matrix:

$$G^* = \operatorname{diag}(G_1, G_2, \ldots). \tag{26}$$

**Theorem 3.2** *Consider the system $\Sigma_\Delta$ with functional uncertainty $\Delta \subset \Delta(L^2(\mathcal{X}; w))$ and initial condition $x_0 \in \mathcal{X}$. Consider the performance measure $\mathcal{P}(Q, k, \Omega_0)$. Suppose $\phi$ is globally $(Q, b, \Omega_0)$ admissable. Let $G^*$ be given by 26 and implement the controller $\Xi(L^2(\mathcal{X}; w), G^*, \alpha, \phi, \mathcal{X}, \Omega_0, A, b)$ where $\alpha > 0$ and $\phi$ is an SFD model. Then $x(t) \to \Omega_0$ as $t \to \infty$. If $\Delta \subset \Delta(L^2(\mathcal{X}; w), \delta)$, then define $\Omega = \{x \in \mathcal{X} \mid x^T P x \leq W_\alpha\}$ where $W_\alpha = \max\{x_0^T P x_0, \eta^2\} + (\delta + q)^2/(2\alpha)$. Then $x(t) \in \Omega$ for all $t \geq 0$ and only a finite number of paramaters are adapted, and*

$$\mathcal{P}(Q, 0, \Omega_0) \leq \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g}(W_\alpha - \eta^2). \tag{27}$$

*If $k > 0$, and $\Delta \subset \Delta(L^\infty(\Omega; w_\infty), \delta_\infty)$ then*

$$\mathcal{P}(Q, k, \Omega_0) \leq \frac{\underline{\lambda}(Q)}{\underline{\lambda}(Q) - g}(W_\alpha - \eta^2) + k\frac{\overline{\lambda}(P)}{\eta^2(\underline{\lambda}(Q) - g)} F_\alpha \tag{28}$$

*where*

$$E_\alpha = \frac{||e_n^T A||}{\sqrt{\underline{\lambda}(P)}} + \mu(\phi)(\alpha\overline{\lambda}(G))^{1/2},$$

$$F_\alpha = \int_{\eta^2}^{W_\alpha} \left(E_\alpha + \frac{\delta_\infty w_\infty^{-1}(\sqrt{v}) + s(\Omega, \Omega_0)}{\eta}\right)^2 v\,dv.$$

**Proof:** The proof is similar to that of theorem 3.1, with a more localised estimate of $W_\alpha$. ■

The key point about of theorem 3.2 is that in contrast to 3.1 we do not need to know a-priori the uncertainty level ($\delta$), or the initial condition ($x_0$), to design a stable controller, ie. the control design is universal to both, just the uncertainty growth characterised by $w$. (Of course the performance bounds are still given in terms of the (unknown) uncertainty levels.) This contrasts to robust control schemes, where the uncertainty level is needed to design a stable controllers. The following example exploits this fact to give a performance bound for non-a-priori known bounded uncertainty: one which has an uncertainty level defined by a distribution. There is no possible robust controller for such an uncertain system.

**Example.** (Spline basis). Consider the system $\dot{x} = f(x) + u$, where $x, u \in \mathbb{R}$ and $x_0 \geq 2$, and suppose we have the $L^2$ uncertainty $\int_{-\infty}^{\infty} f^2(x)\frac{1}{x^2}\,dx < \delta^2$, where the smoothness class is given by $K = Lip_1 1$ and $\delta^2$ is given by an exponential distribution with parameter $\lambda$. Let $\Omega_0 = [1, 1]$, and $Q = I$. Let the model $\phi$ be given by first order B-splines, defined on a uniform lattice of width $1/2$, and take $G = \operatorname{diag}(1/2, 1/8, 1/8, 1/16, 1/16, \ldots)$, where the basis functions are ordered:

$$\phi = [\phi_0, \phi_{-1}, \phi_1, \phi_{-2}, \phi_2, \ldots] \tag{29}$$

where $\operatorname{supp} \phi_i = [(i-1)/2, (i+1)/2]$. Take the control gain to be $A = -1$. Solving the Lyapunov equation we have $P = 1/2$, so $b = 1$. It is easily shown from the smoothness class $K = Lip_1 1$ that $s(\mathbb{R}, \Omega_0) \leq 1/2$, and so $q(\mathbb{R}, \Omega_0) \leq 1$, $g(\mathbb{R}, \Omega_0) \leq 1/2$. Then we have

$$\begin{aligned}
\mathrm{E}[\mathcal{P}(Q, 0, \Omega_0)] &\leq \mathrm{E}\left[\frac{1}{1 - 1/2}\left(\frac{1}{2}x_0^2 - 1 + \frac{(\delta + q)^2}{2\alpha}\right)\right] \\
&= x_0^2 - 1/2 + \frac{1}{\alpha}\mathrm{E}[(\delta + q)^2] \\
&= x_0^2 - 1/2 + \frac{1}{\alpha}\Big(\frac{1}{\lambda} + q^2 + q\sqrt{\frac{\pi}{\lambda}}\Big). \quad (30)
\end{aligned}$$

It is of interest to investigate the asymptotics of the performance as we increase the size/resolution of the model. Typically, as the resolution of the approximant increases, $q, g, s \to 0$. However without additional assumptions, it may be the case that $\underline{\lambda}(R) \to 0$, which will mean the asymptotic bound is useless.

**Definition 3.2** *A linear model resolution schema is a set of the form:*

$$\{\phi_m : \Omega \to I\!R^{i_m} \mid m \in I\!N\} \tag{31}$$

*where $\{i_m\}_m$ is a strictly increasing sequence in $I\!N$ and where*

$$\sup_{f \in C(\Omega)} \|f - \theta_f^T \phi_m\|_{C^0(\Omega)} \to 0 \tag{32}$$

*as $m \to \infty$. The size of a model is the dimension of the weight space, $\dim(\mathcal{W}_m) = i_m$. A linear model resolution control schema is a set of controllers $\Xi_m$ s.t. $\Xi_m = (\mathcal{H}, G_m, \alpha_m, \phi_m, \Omega_0, A, b)$.*

*A linear model resolution control schema is said to be performance $\mathcal{P}(Q, k, \Omega_0)$ resolution structurally scaleable (prss) if for fixed initial conditions $x_0$, and fixed uncertainty $\Delta(\mathcal{H}(\Omega), \delta)$, there exists a uniform performance bound $\mathcal{P}_\infty = \mathcal{P}_\infty(Q, k, \Omega_0) < \infty$ s.t.*

$$\liminf_{m \to \infty} \mathcal{P}_m(Q, k, \Omega_0) \le \mathcal{P}_\infty(Q, k, \Omega_0), \tag{33}$$

*where $\mathcal{P}_m(Q, k, \Omega_0)$ is the performance of $(\Sigma_\Delta, \Xi_m)$.*

**Corollary 3.1**
*Consider the system $\Sigma$ with functional uncertainty $\Delta(L^2(\Omega), \delta)$ and initial condition $x_0$. Implement the controller $\Xi(\mathcal{H}, I, \alpha_m, \phi_m, \Omega_0, A, b)$. Sufficient conditions for a linear model resolution control schema to be performance $\mathcal{P}(Q, 0, \Omega_0)$ resolution structurally scaleable are*

$$\{x \in \mathcal{X} \mid x^T P x \le \max\{x_0^T P x_0, \eta^2\} + \delta^2/\lambda\} \subset \Omega \tag{34}$$

*and $\liminf_{m \to \infty} \underline{\lambda}(R_m)\alpha_m = \lambda > 0$, where $R_m$ is the matrix of the model $\phi_m$ and $\alpha_m$ is the adaption gain. A uniform $\mathcal{P}(Q, 0, \Omega_0)$ performance bound is then given by:*

$$\mathcal{P}_\infty = \max\{x_0^T P x_0, \eta^2\} - \eta^2 + \frac{\delta^2}{\lambda}, \tag{35}$$

*where $\alpha = \liminf_{m \to \infty} \alpha_m$. If additionally $\limsup_{m \to \infty} \alpha_m^{\frac{1}{2}} \mu(\phi_m) \le \mu < \infty$ then this ensures $\mathcal{P}(Q, k, \Omega_0)$ resolution scaleability for $k > 0$.*

**Proof:** This follows from definition 3.2 and theorem 3.1. ∎

**Example.** Let $\psi_m$ be an orthonormal model resolution schema. Then consider

$$\Xi(\mathcal{H}, I, \alpha_m = \mu(\psi_m)^2, \phi_m = \frac{1}{\mu(\psi_m)}\psi_m, \Omega_0, A, b). \tag{36}$$

Then $\underline{\lambda}(R_m(\phi_m))\alpha_m = 1$, $\alpha_m^{1/2}\mu(\phi_m) = 1$ and the sufficient conditions are met for small enough uncertainty levels $\delta$ given an approximation region $\Omega$.

**Example.** Is is shown in [3] that the Gaussian networks of [4] do not generally satisfy the above sufficient condition and indeed are not prss (stability is ensured with a sliding mode) when $G = I$ and $\alpha$ is constant.

## 4 Further Research

Further work will extend the examples 3.1, 3.1 into a general instability result and a general lower bound. This hopefully will lead to necessary and sufficient conditions for prss for example. It is straightforward to generalise these results to more general multiple input / output systems with matched uncertainties. Importantly, these results can be extended to the non-matched case via the backstepping procedure. This is possible due to the fact the backstepping designs also construct quadratic Lyapunov functions. However there are extra difficulties in extending these results in this way due to the coordinate transform; which mean that these extensions are not straightforward. Further details can be found in the forthcoming paper [2]. Further applications of this line of research can be found in the thesis [1].

## References


[1] M. French. *Adaptive control of functionally uncertain systems.* PhD thesis, University of Southampton, 1997. Under preparation.

[2] M. French and E. Rogers. Uncertainty, performance in unmatched approximate adaptive nonlinear control. 1997. Submitted.

[3] M. French, C. Szepesvari, and E. Rogers. Uncertainty, performance and model dependancy in approximate adaptive nonlinear control. 1997. To be submitted.

[4] R.M. Sanner and JJ.E. Slotine. Gaussian networks for direct adaptive control. *IEEE Trans. on Neural Networks*, 3(6):837–863, 1992.