Sufficient Conditions for Robust Performance of Adaptive Controllers with General Uncertainty Structure*

JAMES M. KRAUSE,†‖ GUNTER STEIN†‡ and PRAMOD P. KHARGONEKAR†§

A robust gain scheduled controller combined with a parameter estimator produces a robust adaptive controller for a broad class of deterministic but uncertain systems, provided the estimator is efficient in a new, deterministic sense.

Abstract—Sufficient conditions are given under which an adaptive control system is robustly stable and achieves a guaranteed robust asymptotic performance level equal to that of the robust controller given perfect parameter information. The conditions are general in several respects. For example, structured non-parametric uncertainty (e.g. block diagonal) is allowed, as well as exogenous noise inputs. In addition, the structure of the parametric uncertainty is very general, and even allows for parameters which scale the uncertainty magnitudes. This allows one to identify the size of the non-parametric uncertainty and to schedule the controller based on this size. Finally, the robust gain scheduled controller is largely unrestricted. Identification mechanisms which are proven to satisfy the sufficient conditions are not given here and, for the general problem, have not yet been developed. However, an example of such a mechanism for a subclass of systems does exist and is referenced. For the general problem, this paper provides properties to be sought in the development of robust identification laws for robust adaptive control.

1. INTRODUCTION

ROBUST CONTROL seeks stability and performance guarantees in the presence of modeling errors. The modeling errors are typically characterized as real parameter uncertainties and/or dynamical uncertainties with frequency domain magnitude bounds (examples of the latter include Bode, 1945; Horowitz, 1963; Doyle and Stein, 1981; Lehtomaki, 1981).

Adaptive control is the combination of an adjustable controller and a plant parameter estimation process which drives the controller adjustment. In effect, adaptive control is designed to overcome parameter uncertainty. Stable approaches have been obtained (e.g. Narendra et al., 1980; Morse, 1980; Goodwin et al., 1980).

Robust adaptive control is adaptive control in the presence of non-parametric modeling uncertainty. It is a challenge receiving much attention, and various robust stability results have been obtained (e.g. Kreisselmeier and Anderson, 1986; Narendra and Annaswamy, 1987; Krause et al., 1989; Ortega et al., 1985; Ioannou and Tsakalis, 1986).

Of course, these various results have limitations (as all results do). Among the practically significant limitations are the structural assumptions regarding the parametric and non-parametric uncertainty. This is in contrast to the generality in uncertainty structure one can find in robust multivariable control theory (e.g. Doyle, 1985; Doyle et al., 1982).

Another limitation is an absence of an on-line identification of the non-parametric uncertainty magnitude. This magnitude is relevant to the assignment of a controller to maximize the robust performance of the system. Prior bounds can be used in place of identified bounds in determining the robust controller assignment, at the cost of potential conservatism.

Recent research directions may eventually address these limitations. Separate from the

† Honeywell Systems and Research Center, Honeywell MN65-2500, 3660 Technology Drive, Minneapolis, MN 55418, U.S.A.
‡ Department of Electrical Engineering, Massachusetts Institute of Technology, MA, U.S.A.
§ Department of Electrical Engineering, University of Michigan, MI, U.S.A.
‖ Author to whom all correspondence should be addressed.

problem of adaptive control, progress has been made in structured uncertainty model invalidation (Smith and Doyle, 1989; Smith, 1989). This problem is a precursor to robust parameter identification with structured non-parametric uncertainty. Progress has been made in consistent parameter set estimation (Younce, 1989; Younce and Rohrs, 1990; Kosut et al., 1990), which creates a potential for adaptive algorithms which take advantage of the additional information provided by the membership set of the parameters rather than a single element of the set. Likewise, progress has been made in on-line modeling error characterization (LaMaire and Valavani, 1987; Helmicki et al., 1990a, b; Goodwin and Salgado, 1989; Younce, 1989; Kosut, 1987, 1988), which creates a potential for adaptive algorithms which take advantage of the particular value of the modeling error rather than assuming the worst case.

Progress in these areas could lead to robust stability and performance guarantees in an adaptive control context. This paper contributes to this development by providing sufficient conditions under which an adaptive system yields robust stability and asymptotic performance guarantees. Structured uncertainty is allowed, and adaptive systems with and without on-line modeling error characterization are addressed. The adaptive controller robust stability and asymptotic performance guarantees are as strong as that of the robust gain scheduled controller given perfect plant parameter information.

Abstractly, the conditions are:

Condition 1. Uncertainty bounds

The unknown parameters must lie in some known bounded set. Energy bounds (L^2 with exponential weighting) are required for exogenous disturbances. Energy-gain bounds (induced norms) are required for non-parametric dynamical modeling errors. Some additional uncertainty information, such as uncertainty structure, is allowed but not required.

Condition 2. Lipschitzian gain schedule

The controller gains must be a locally Lipschitzian function of the plant parameter estimates. This eliminates some switching control schemes, but still admits an interesting class of gain schedules.

Condition 3. Robust performance of the gain scheduled controller

In this paper, the "gain scheduled controller" is the controller which would result from perfect plant parameter information, using the designed gain schedule. There is a family of possible gain scheduled controllers since there is a set of possible true plant parameters. We require that for any value of the true plant parameter, the corresponding ideal gain scheduled controller must provide robust performance. This is a natural minimum requirement when one takes the philosophical perspective that an adaptive control system consists of a gain scheduled controller and an identification mechanism. This perspective is not universal in adaptive control theory; it is perfectly legitimate to regard an adaptive controller as a single nonlinear control mechanism.

Condition 4. Non-degeneracy

The "robustly performant" gain scheduled system must also have bounded gain in response to certain added inputs. Moreover, the performance level must not degrade discontinuously for arbitrarily small errors in the controller parameters. These are properties which are normally expected of a robust control system, and are not an issue in practice. They are explicitly included here merely because of the necessity of precluding degenerate counter-examples.

Condition 5. Asymptotic efficiency

The parameter estimate must remain bounded, and be asymptotically efficient in a newly-defined deterministic sense. Effectively, the estimate must in the limit as time approaches infinity become consistent (modulo decaying transients) with recent measurements and prior system information.

Even more abstractly, this paper shows that robust control plus asymptotically efficient estimation yields robust adaptive control. In this sense, this paper provides a separation principle for adaptive control which puts theoretical rigor behind the common engineering heuristic decomposition of an adaptive controller into a scheduled controller and an identification process. Note that robust control plus parameter convergence to the true value would also provide stability and performance, but does not serve to separate the problem since the convergence depends on excitation, which in turn depends on the adjustable controller parameter behavior, which in turn depends on the parameter estimate behavior. The use of efficiency in this paper, rather than parameter convergence, allows true separation; efficiency is a property of the estimation algorithm, not a joint property of the algorithm and excitation. As such, estimation efficiency allows us to conclude adaptive controller stability and performance without

excitation assumptions and without parameter convergence to the true value.

Overall, the treatment of this paper is fairly philosophical. The benefit of this is generality; robust parameter adjustment mechanisms may now be developed under various uncertainty scenarios. Application of the robust parameter adjustment mechanisms in an adaptive controller will automatically lead to robust asymptotic performance if the relatively simple abstract conditions are satisfied.

A concrete identification mechanism which also provides a guarantee of robust asymptotic performance has been given in Krause *et al.* (1989), although that paper and this paper are entirely different. Krause *et al.* (1989) addresses special uncertainty structure, while this paper allows a very broad class of uncertainties (e.g. block diagonal and of arbitrary interconnection to the known dynamics). Moreover, Krause *et al.* (1989) provides a particular complete recursive robust adaptive algorithm, while this paper provides generally applicable sufficient conditions for robustness of an adaptive system.

Another paper which provides robust asymptotic performance results for adaptive controllers is Poolla and Shamma (1990). There the notion of "asymptotic" is distinct from the notion used here; the distinction resembles the distinction between signals in L^2 ($\approx$ their paper) and bounded signals which go to zero ($\approx$ this paper). Moreover, Poolla and Shamma (1990) employ a switching control scheme which is distinct from the continuous adaptation of Krause *et al.* (1989) and the general identification processes of this paper. Needless to say, these fundamental problem distinctions lead to highly divergent formal details. Still, Poolla and Shamma (1990), Krause *et al.* (1989) and this paper share a common theme: robust asymptotic performance of adaptive controllers from an operator theoretic perspective.

2. PRELIMINARIES

A. *Mathematical notation and terminology*

Throughout the development, σ is a particular fixed and known nonnegative number. Superscript T denotes transposition.

Consider a function $x:\mathbb{R}^+\to\mathbb{R}^N$. Define the seminorm

$$\|x\|^{\sigma,t} := \left[\int_0^t e^{-2\sigma(t-\tau)}x^T(\tau)x(\tau)\,d\tau\right]^{1/2}, \quad (1)$$

where the superscript σ is omitted when $\sigma = 0$, and the superscript t is omitted when $t = \infty$. When $\|x\|^{\sigma,t}$ exists for all finite t, x is said to be

TABLE 1. NORM NOTATION SUMMARY

$\|\cdot\|$	Euclidean or Frobenius Norm
$\|\cdot\|^{\sigma,t}$	Signal norm (equation (1))
$\|\cdot\|^{\sigma}$	Signal norm with $t = \infty$
$\|\cdot\|^{0,t}$	Signal norm with $\sigma = 0$
$\|\cdot\|$	Signal norm with $t = \infty$ and $\sigma = 0$
$\|\cdot\|_{i2}^{\sigma}$	$L^{2,\sigma}$-induced operator norm
$\|\cdot\|_{i2}$	L^2-induced operator norm

in $L^{2e,\sigma}$. When $\|x\|^{\sigma,t}$ is uniformly bounded over all $t > 0$, x is said to be in $L^{2,\sigma}$. When $\sigma = 0$, we omit it from the superscript, and use the more common notation $x \in L^{2e}$ or $x \in L^2$.

For a real vector θ, $|\theta|$ denotes the Euclidean norm.

Table 1 summarizes the norm notation used in this paper. Note that signal norms take the form "$\|\cdot\|^{\text{norm parameters}}$", while operator norms take the form "$\|\cdot\|_{\text{type of norm}}^{\text{norm parameters}}$".

A set of operators Δ will be said to be a convex set of operators if and only if for any $\Delta_1 \in \Delta$ and $\Delta_2 \in \Delta$ and any scalar $\rho \in [0, 1]$, it follows that $(\rho\Delta_1 + (1 - \rho)\Delta_2) \in \Delta$.

When X is a set and γ is a scalar, the scaled set (denoted γX) is defined to be

$$\gamma X := \{\gamma x : x \in X\}.$$

This definition applies to sets of operators as well as sets of signals, taking γ to be either the cascade of a scaling operation with the operation x, or a direct scaling of the signal x, as dictated by the context.

For $X \subseteq \mathbb{R}^n$ a function $F:X\to\mathbb{R}^m$ will be said to be locally Lipschitzian on X if and only if for each compact set Y in X there exists a finite constant $k = k(Y)$ such that for any $x_1 \in Y$ and $x_2 \in Y$,

$$|F(x_1) - F(x_2)| < k\,|x_1 - x_2|.$$

B. *System assumptions*

Assumption 1. The system is assumed to take the form of Fig. 1.
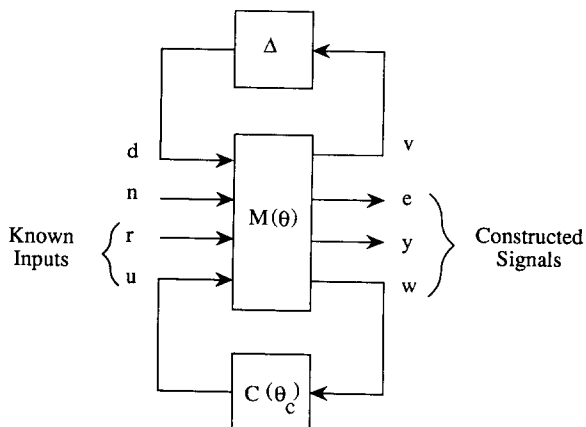

FIG. 1. General system structure.

Assumption 2. In the figure, $M(\theta)$ is a linear time invariant system which depends on an unknown parameter vector $\theta \in \mathbb{R}^N$, but is otherwise known. M includes much of the plant description as well as much of the controller. This style of representation affords considerable flexibility in terms of plant parameterizations and controller structures (Doyle, 1985; Doyle *et al.*, 1982).

Assumption 3. We will assume prior knowledge of θ of the form $\theta \in \Theta_P$ where Θ_P is a known compact set.

Assumption 4. External to M in Fig. 1 is a non-parametric uncertainty Δ. We will assume that $\Delta \in \mathbf{\Delta}$, where $\mathbf{\Delta}$ is a known convex set of operators containing the zero operator. Moreover, we assume that Δ has been normalized by the choice of M such that $\|\Delta\|_{i2}^\sigma \leq 1$ for all $\Delta \in \mathbf{\Delta}$. Note that the assumptions on Δ allow structured uncertainty (e.g. block diagonal). The signals v and d are internal signals of the plant.

Assumption 5. Also external to M in Fig. 1 is the block $C(\theta_c)$. C is a real matrix, which depends on the adjustable controller gains $\theta_c(t): \mathbb{R}^+ \to \mathbb{R}^M$. We will assume only that C is a locally Lipschitzian function of θ_c on $\mathbb{R}^M$.

Assumption 6. The signals $w: \mathbb{R} \to \mathbb{R}^{n_w}$ and $u: \mathbb{R} \to \mathbb{R}^{n_u}$ are internal signals of the controller.

Assumption 7. The input vector $n: \mathbb{R} \to \mathbb{R}^{n_n}$ of Fig. 1 is a noise input. We will assume that $n \in \mathbf{N}$, where $\mathbf{N}$ is a convex set of noise signals containing zero, and satisfying $\|n\|^{\sigma,t} \leq 1$ for all $n \in \mathbf{N}$.

Assumption 8. The input vector $r: \mathbb{R} \to \mathbb{R}^{n_r}$ of Fig. 1 is the known exogenous input to the system, and is assumed to lie in $L^\infty \cap L^{\sigma,t}$.

Assumption 9. The output vector $e: \mathbb{R} \to \mathbb{R}^{n_e}$ of Fig. 1 contains the performance variables. These are assumed to be constructable from available measurements and/or known inputs (the definition of M includes the construction). Later, robust performance will involve keeping e small in some sense. A typical example of a performance variable is the command tracking error.

Assumption 10. The output vector $y: \mathbb{R} \to \mathbb{R}^{n_y}$ of Fig. 1 represents other measurements not already accounted for by e and w, but potentially relevant to the on-line parameter identification task. The vector dimension of y can be zero (y absent).

Assumption 11. As time evolves, the known inputs to M are r and u, and the constructed (known) output signals are e, y and w. Hereafter we will refer to these as the known inputs and constructed outputs, respectively.

Assumption 12. In addition to the separate assumptions on Δ, M, and C above, we also assume that the feedback interconnection of Fig. 1 is well-posed (see Willems, 1971) for each $\theta \in \Theta_P$ and each $\Delta \in \mathbf{\Delta}$. That is, taking $n(t)$, $r(t)$, and $\theta_c(t)$ as exogenous varying inputs to the system, and taking $\theta \in \Theta_P$ and $\Delta \in \mathbf{\Delta}$ to be fixed, we assume that: (i) for all finite time and as long as the inputs are bounded, the solution to the system of equations represented by Fig. 1 exists and is unique, where "solution" includes all signals shown in the figure as well as the internal states of $M(\theta)$, and (ii) for any finite interval of time, this solution is locally Lipschitz-continuous as a function of n, r, and θ_c.

We will use the shorthand $S(\theta_0, \theta_{c0})$ to denote the system of Fig. 1 with the plant parameter vector taking on a particular specified value θ_0 and the controller parameter vector taking on a particular specified value θ_{c0}. In this notation, we have suppressed the dependence of the system on Δ; S actually denotes a family of systems parameterized by Δ alone.

C. *Initial conditions*

Since the subsystems M and Δ of Fig. 1 are dynamical, and since this paper measures and analyzes the time response on the half line ($t \geq 0$), the issue of initial conditions is important.

We will represent the initial condition of Δ as an added response to the output. Initial conditions arise a variety of ways; our characterization is consistent with the notion of past excitation causing initial conditions. That is, we suppose Δ satisfies

$$\sqrt{\int_{-\infty}^0 e^{-2\sigma(-\tau)} d^T(\tau) d(\tau) d\tau}$$
$$\leq \sqrt{\int_{-\infty}^0 e^{-2\sigma(-\tau)} v^T(\tau) v(\tau) d\tau} =: \bar{d}_x(0) < \infty, \quad (2)$$

and

$$\|d\|^{t,\sigma} \leq \sqrt{\int_{-\infty}^t e^{-2\sigma(t-\tau)} v^T(\tau) v(\tau) d\tau}$$
$$\leq \sqrt{\int_{-\infty}^0 e^{2\sigma\tau} v^T(\tau) v(\tau) d\tau}$$
$$+ \sqrt{\int_0^t e^{-2\sigma(t-\tau)} v^T(\tau) v(\tau) d\tau}$$
$$= \bar{d}_x(0) e^{-\sigma t} + \|v\|^{t,\sigma}. \quad (3)$$

Thus there exists a partitioning of the signal d into

$$d(t) = d_x(t) + d_f(t), \quad t \geq 0, \quad (4)$$

such that

$$\|d_x\|^{t,\sigma} \leq \bar{d}_x(0)e^{-\sigma t} \quad (5a)$$

$$\|d_f\|^{t,\sigma} \leq \|v\|^{t,\sigma}. \quad (5b)$$

Note that we have not applied superposition in the treatment of the initial condition of Δ.

Without loss of generality, $M(\theta)$ can be written as the result of closing a (multivariable) loop around another matrix transfer function $M_1(\theta)$, with $M_1(\theta)$ uniformly stable over all $\theta \in \Theta_P$ (see Fig. 2). Applying superposition, we represent the response of $M_1(\theta)$ as an open-loop initial condition response (v_x, e_x, y_x, w_x, and z_x) and a force response. By the open-loop stability of M_1, it is known that the $\| \cdot \|^{\sigma,t}$ norm of v_x, e_x, y_x, w_x, and z_x is bounded and decays to zero as $t \to \infty$.

By augmenting the unknown exogenous signal n with the additional signals d_x, v_x, e_x, y_x, w_x, and d_x, we can remove the effects of the initial conditions from the internal workings of M. In effect, by expanding the vector n and its set N with additional bounded and transient exogenous disturbances, we may assume, without loss of generality, that M and Δ are initially at rest.

Assumption 13. Hereafter, we assume that n incorporates the effects of initial conditions in this way.

D. *Notation reference*

TABLE 2. SYMBOLS

θ	Unknown system parameters
$\hat{\theta}(t)$	Estimate of θ
$\theta_c(t)$	Controller parameters
n	Unknown exogenous inputs
r	Known exogenous inputs
e	Performance variables
y	Miscellaneous measurements
Δ	Unmodeled dynamics
$\mathbf{\Delta}$	Known membership set for Δ
N	Known membership set for n
Θ_P	Prior known membership set for θ
$\Theta(t)$	Membership set at time t

3. CONDITIONS FOR ROBUST ASYMPTOTIC PERFORMANCE

This section develops conditions under which the robust asymptotic performance of an adaptive system equals that of a gain scheduled robust controller designed with perfect parameter information.

First, the robust performance of the gain scheduled system is defined. Next, the concept of "consistent parameter set" is introduced and
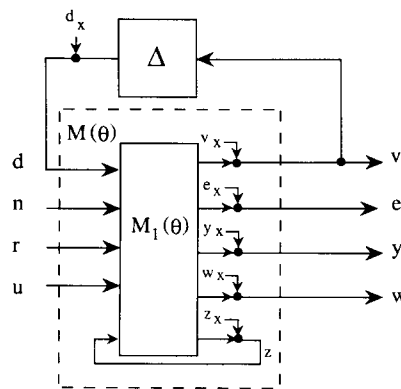

FIG. 2. Nonzero initial condition effects.

shown to have certain asymptotic properties. Finally, modest constraints on the gain scheduled system and an "efficiency" constraint on the identification process are shown to imply robust stability and asymptotic performance of the adaptive system.

A. *Scheduled system robust performance*

In this section, the robust performance of the gain scheduled system is defined. The definition used here deliberately resembles that of Doyle *et al.* (1982) because of its practical utility (Enns, 1987; Enns *et al.*, 1987; Enns and Bugajski, 1987). However, we deviate slightly with our definition. Most importantly, (1) exponentially weighted signal norms are used here to allow a more isolated description of asymptotic performance, and (2) the "known" part of the system is parameterized by the unknown plant parameter vector, requiring that a family of systems be designed and analyzed.

The identification process will be defined shortly. For now, let Θ_R be a prior known subset of $\mathbb{R}^N$ which covers the range of the identification process, so $\hat{\theta}(t) \in \Theta_R$ $\forall t$ is known *a priori*. This allows estimators with parameter limits, such as switching estimators which confine $\hat{\theta}(t)$ to a pre-defined finite set of points. On the other hand, one may take the set Θ_R to be all of $\mathbb{R}^N$.

Definition 1. Gain schedule. Considering each value $\theta \in \Theta_P$ separately, one can assign controller gains. Let $f : \Theta_P \to \mathbb{R}^M$ represent this assignment, and let f be extended to the rest of Θ_R, so that $f : \Theta_R \to \mathbb{R}^M$. We refer to f as a gain schedule. When $\theta_c = f(\theta)$ and $\theta \in \Theta_P$, the resulting controller is said to be a gain scheduled controller. For any value of $\theta \in \Theta_P$, the system of Fig. 1 with the parameter vector taking on the value θ and with θ_c taking on the value $\theta_c = f(\theta)$ is a gain scheduled system. That is, for any particular $\theta \in \Theta_P$, $S(\theta, f(\theta))$ is a gain scheduled system.

Assumption 14. We assume that there exists an open set Θ_0 containing Θ_P such that f is locally Lipschitzian on $\Theta_0 \cap \Theta_R$.

Remark. Assumption 14 is structured such that the theorems to follow, apply to estimators which adjust parameters smoothly as well as estimators which switch between distinct candidate parameter values, or combinations of the two. Note in particular that assumption 14 is satisfied for the special case in which Θ_R is $\mathbb{R}^N$ and f is locally Lipschitzian on an open set containing Θ_P. Assumption 14 is also clearly satisfied for the special case in which Θ_R is a finite set of points and f is Lipschitzian on Θ_R.

Definition 2. Graph of the gain schedule. For each value of $\theta \in \Theta_P$, there corresponds a gain scheduled system. The collection of all pairs consisting of $\theta \in \Theta_P$ and the associated gain scheduled system is the graph of the gain schedule. That is, $\{S(\theta, f(\theta)),\ \theta \in \Theta_P\}$ is the graph of the gain schedule.

The significance of the gain schedule graph is as follows. Prior analysis of the particular gain scheduled system corresponding to the true value of θ is not possible, since the particular value of θ is not known. Nonetheless, one can determine the performance guarantees for the gain schedule graph, since the graph is known *a priori*. This performance level, then, is the most favorable level one can hope to guarantee *a priori* with an adaptive system.

Of course, the performance of a particular gain scheduled system (corresponding to a particular value of θ) could exceed a prior guarantee valid for the entire graph (namely the worst case over all $\theta \in \Theta_P$ of the gain scheduled system performance). In this case the adaptive system may also exceed the prior performance guarantee.

Consequently, various forms of p rformance are of interest. These are now defined.

Definition 3. Performance measures. For each $\theta \in \Theta_P$, the unforced robust performance measure of the gain scheduled system is a real scalar $\alpha_1(\theta)$ defined by

$$\alpha_1(\theta) := \sup [\|e\|^{\sigma,t} : \|r\|^{\sigma,t} = 0,\ n \in \mathbf{N},$$

$$\Delta \in \Delta,\ \theta_c = f(\theta)].$$

The forced robust performance measure of the gain scheduled system is a real scalar $\alpha_2(\theta)$ defined by

$$\alpha_2(\theta) := \sup \left[\frac{\|e\|^{\sigma,t} - \alpha_1(\theta)}{\|r\|^{\sigma,t}} : \|r\|^{\sigma,t} \neq 0, \right.$$

$$\left. n \in \mathbf{N},\ \Delta \in \Delta,\ \theta_c = f(\theta) \right].$$

Definition 4. Robust performance levels. The performance level of the gain scheduled system is a real scalar function $J_1(t, r, \theta)$ defined by

$$J_1(t, r, \theta) = \alpha_1(\theta) + \alpha_2(\theta) \|r\|^{\sigma,t}.$$

The performance bound for bounded r norm of the gain scheduled system is a real scalar function $J_2(\bar{r}, \theta)$ defined by

$$J_2(\bar{r}, \theta) = \alpha_1(\theta) + \alpha_2(\theta)\bar{r}.$$

The performance guarantees over the entire graph of the gain schedule are clearly

$$J_1(t, r, \theta) \leq \bar{J}_1(t, r) := \sup_{\theta \in \Theta_P} J_1(t, r, \theta), \quad (6)$$

$$J_2(\bar{r}, \theta) \leq \bar{J}_2(\bar{r}) := \sup_{\theta \in \Theta_P} J_2(\bar{r}, \theta). \quad (7)$$

Remark. J_1 is useful as a nonconservative combination of α_1 and α_2. The J_2 alternative is conservative for some values of the function r. It is nonetheless useful in that it represents a time-invariant tradeoff between noise response and command response. Later, this allows a time-invariant partial ordering of the gain scheduled systems, so some points on the gain schedule graph can be viewed as "better" than others.

To facilitate a comparison after initial adaptation transients have subsided, we define asymptotic robust performance.

Definition 5. Asymptotic robust performance. For a particular value of the input function $r : \mathbb{R} \to \mathbb{R}^{nr}$, an asymptotic robust performance level of $J_1(t, r, \theta)$ [respectively, $J_2(\bar{r}, \theta)$] is achieved if and only if there exists a bounded $\varepsilon(t)$ with a limit of zero as $t \to \infty$ such that

$$\|e\|^{\sigma,t} \leq J_1(t, r, \theta) + \varepsilon(t) \quad \forall t \geq 0$$

(respectively, $\|e\|^{\sigma,t} \leq J_2(\bar{r}, \theta) + \varepsilon(t) \quad \forall t \geq 0$).

Note that, because of the inequality above, a system can be said to achieve many asymptotic performance levels.

We now proceed to the technical detail of non-degeneracy—a detail which is reminiscent of internal stability.

Definiton 6. Non-degeneracy. The robustly performant gain scheduled system $(\theta_c = f(\theta),\ \theta \in \Theta_p)$ is said to be non-degenerate if, and only if: (i) the effect on performance of perturbations added to the outputs of all blocks of Fig. 1 (as shown in Fig. 3) is bounded, i.e. for the graph of gain scheduled systems,

$$\|e\|^{\sigma,t} - \alpha_2(\theta) \|r\|^{\sigma,t} - \alpha_1(\theta)$$

$$\leq c(\|d_d\|^{\sigma,t} + \|n_d\|^{\sigma,t} + \|r_d\|^{\sigma,t} + \|u_d\|^{\sigma,t}$$

$$+ \|v_d\|^{\sigma,t} + \|e_d\|^{\sigma,t} + \|y_d\|^{\sigma,t} + \|w_d\|^{\sigma,t})$$
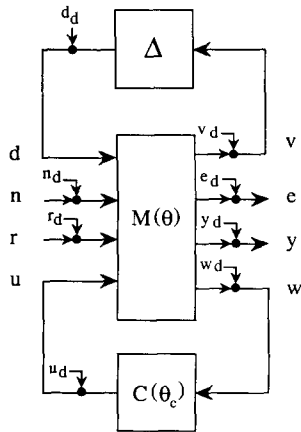
FIG. 3. Signal insertion points for non-degeneracy definition.

for some constant $c < \infty$ independent of θ, and (ii) given $\varepsilon > 0$ there exists a $\delta > 0$ such that if $|\tilde{\theta}(t)| < \delta$ for all t and $\theta_c(t) = f(\theta - \tilde{\theta}(t))$, then

$$\|e\|^{\sigma,t} - (\alpha_2(\theta) + \varepsilon) \|r\|^{\sigma,t} - (\alpha_1(\theta) + \varepsilon)$$

$$\leq (c + \varepsilon)(\|d_d\|^{\sigma,t} + \|n_d\|^{\sigma,t} + \|r_d\|^{\sigma,t} + \|u_d\|^{\sigma,t}$$

$$+ \|v_d\|^{\sigma,t} + \|e_d\|^{\sigma,t} + |y_d\|^{\sigma,t} + \|w_d\|^{\sigma,t}),$$

where c is valid in (i) above, and the choice of δ given ε is independent of θ.

Remark. In practice (and generically, we believe), a particular level of robust performance is invariably more difficult to achieve than mere robust bounded-gain stability. Furthermore, the robust performance of the system is not infinitely sensitive to small parameter variations. The term "non-degeneracy" above is appropriate.

Assumption 15. We assume that for each $\theta \in \Theta_P$, the associated gain scheduled system is non-degenerate.

B. Consistent parameter set

This section provides concepts which are necessary for an understanding of the "efficiency" of an identification process, which in turn is critical to the achievement of robust asymptotic performance.

Definition 7. Consistent parameter set $\Theta(t)$. The consistent parameter set $\Theta(t)$ is defined to be the set of all parameter vectors θ that are consistent with the M and Δ interconnection structure of Fig. 1 and all known inputs on the interval $[0, t)$ and all constructed signals on the interval $[0, t)$ and the uncertainty assumptions $\Delta \in \Delta$, $n \in N$, and $\theta \in \Theta_P$.

Note. With u and w both known, the dependency of u on w as shown in Fig. 1 provides no information regarding θ, and therefore this

aspect of Fig. 1 is omitted from the definition of $\Theta(t)$.

Definition 8. Asymptotic consistent parameter set Θ_∞

$$\Theta_\infty := \lim_{t \to \infty} \left\{ \bigcap_{\tau \geq t} \Theta(\tau) \right\} = \bigcap_{\tau \geq 0} \Theta(\tau).$$

Note. The last inequality and the existence of the limit follows the obvious fact that $\Theta(t_2) \subseteq \Theta(t_1)$ if $t_1 \leq t_2$, and $\Theta(t)$ is compact since Θ_P is compact.

Lemma 1. Θ_∞ is nonempty.

Proof. The system assumptions assert that some value of θ exists such that all residual uncertainty can be captured by $n \in N$ and $\Delta \in \Delta$. This is an element of $\Theta(t)$ for all t, and therefore it is an element of Θ_∞. □

The definition of Θ_∞ involves the entire information content of the measurements and assumptions. It may be difficult to construct identification laws which make such a thorough use of information. Motivated by properties of the robust identifier of Krause *et al.* (1989), we now weaken the notion of a consistent parameter set to make it more attainable to an identification process.

Definition 9. Asymptotic consistent parameter set modulo transients, Θ_∞^+. Recall that e, y, and w are the constructed signals. A parameter vector θ is an element of the set Θ_∞^+ if and only if $\theta \in \Theta_P$ and there exist signals $\hat{e}$, $\hat{y}$, and $\hat{w}$ such that: (i) $\|\hat{e} - e\|^{\sigma,t} + \|\hat{y} - y\|^{\sigma,t} + \|\hat{w} - w\|^{\sigma,t} \leq \varepsilon(t)$ for some $\varepsilon(t)$ which is bounded and has a limit of zero as $t \to \infty$, and (ii) if the constructed signals e, y, and w were conceptually replaced by $\hat{e}$, $\hat{y}$, and $\hat{w}$, then θ would be consistent with the M and Δ interconnection structure of Fig. 1 and all known inputs on the interval $[0, \infty)$ and $\hat{e}$, $\hat{y}$, and $\hat{w}$ as the fictitious constructed signals and the postulated uncertainty assumptions $\Delta \in \Delta$ and $n \in N$.

In words, Θ_∞^+ are parameters which need not be completely consistent with the constructed signals. However, the inconsistency can be no more than transient.

C. Adaptive system robust performance

In stochastic estimation problems, an estimate is said to be efficient if its variance equals the Cramer–Rao bound, and asymptotically efficient if the ratio of the variance over the Cramer–Rao bound approaches one as time approaches infinity (Van Trees, 1968). In this paper, we have a deterministic problem formulation; the

uncertainty satisfies a deterministic bound rather than a stochastic distribution. Consequently, we redefine efficiency for the deterministic formulation, retaining the underlying abstract notion that efficiency is a "thorough use of information".

With the added assumption that the identification mechanism is efficient, or even only asymptotically efficient, we show that the system automatically achieves robust asymptotic performance levels equal to those of the gain scheduled system. No excitation conditions are required.

Definition 10. Identification process. In this paper, an identification process is a strictly causal mapping of known inputs, constructed signals, and prior information to an estimate $\hat{\theta}(t)$ of θ, with $\hat{\theta}(t)$ finite on any finite interval $[0, T]$.

Remark. Finite $\hat{\theta}(t)$ is a mild restriction on the mapping since Θ_P is compact.

Definition 11. Efficiency. An identification process is efficient if and only if

$$\hat{\theta}(t) \in \Theta(t) \quad \forall t.$$

Definition 12. Virtual efficiency. An identification process is virtually efficient if and only if

$$\lim_{t \to \infty} \hat{\theta}(t) =: \hat{\theta}_\infty \quad \text{exists and} \quad \hat{\theta}_\infty \in \Theta_\infty.$$

Definition 13. Asymptotic efficiency. An identification process is asymptotically efficient if and only if

$$\lim_{t \to \infty} \hat{\theta}(t) =: \hat{\theta}_\infty \quad \text{exists and} \quad \hat{\theta}_\infty \in \Theta_\infty^+.$$

Remark. Clearly, the three forms of efficiency defined above are progressively weaker.

Theorem 1. Given the system assumptions 1–13, and given that the adjustable controller parameters satisfy $\theta_c(t) = f(\hat{\theta}(t))$ for a gain schedule f satisfying assumptions 14 and 15, if:

Assumption 16. The gain schedule graph has a robust performance level of $J_1(t, r, \theta)$, and

Assumption 17. $\hat{\theta}(t)$ is generated by an asymptotically efficient identification process,

then all signals shown in Fig. 1 for the adaptive system are in $L^{2,\sigma}$ (this is $L^{2,\sigma}$-BIBO stability), and the adaptive system achieves an asymptotic performance level of $J_1(t, r, \hat{\theta}_\infty)$.

Proof. See Appendix.   □

Corollary. Under the conditions of the theorem, the prior worst case performance guarantee (in the J_1 sense) for the adaptive system is equal to the performance guarantee $\bar{J}_1(t, r)$ for the gain schedule graph.

Paraphrasing Theorem 1, under the given conditions, the adaptive control system has a prior performance guarantee which is as good as that of the gain schedule graph. The performance can be better however; the adaptive system asymptotically performs at least as well as the gain scheduled system $S(\hat{\theta}_\infty, f(\hat{\theta}_\infty))$, even though the asymptotic system is actually $S(\theta, f(\hat{\theta}_\infty))$ where θ need not equal $\hat{\theta}_\infty$.

Interestingly, no assumptions are required regarding the performance or even stability of "untuned" systems $S(\theta, f(\theta_2))$ with $\theta_2 \neq \theta$. For a particular value of θ, there may be values of θ_2 within Θ_P such that $S(\theta, f(\theta_2))$ is unstable. We can conclude from the theorem that when the gain schedule graph has robust performances, the efficiency property of the estimator steers the estimate away from values which make S demonstrate instability or poor performance for the particular applied input. Note, however, that parameter estimate convergence to the true value θ does not necessarily follow.

This is a key point: there is no claim that the asymptotic parameter estimate is correct. We know only that the system will behave (in a certain sense) as if it is correct. [Qualitatively similar observations have been made for different adaptive control problem formulations. See Becker *et al.* (1985) for an analysis which also distinguishes between convergence to a true parameter and convergence to a desired performance.]

4. PREFERRED PARAMETERS

In general, when unmodeled dynamics are included in the problem formulation, the "true" plant parameters are not uniquely defined. Multiple choices of θ may produce the identical input–output map for all possible inputs, for some θ-dependent choice of Δ which satisfies all assumptions. Even in the rare case in which the parameters are uniquely defined using all possible inputs, the particular observed known inputs and constructed signals are generally insufficient to uniquely determine the parameters. That is, $\Theta(t)$ and Θ_∞ are sets with many members.

This nonuniqueness affords freedom to the identification process. Even among efficient identification processes, there is freedom. Theorem 1 provides a clue as to how to exploit this freedom to improve or optimize the

asymptotic performance, as shown in this section.

First we describe the preferred parameters strategy in general terms. Next, we show an interesting special case involving the on-line identification of the magnitude of the non-parametric uncertainty and noise simultaneously with plant parameter identification. Finally, we state a stronger asymptotic performance result which is achieved using a preferential identification mechanism.

A. Preferential identification

Let a preference take the form of a mapping

$$\mathbf{P}: \{\text{subsets of } \mathbb{R}^N\} \rightarrow \{\text{subsets of } \mathbb{R}^N\},$$

which is weakly contractive in the sense that $\mathbf{P}(\Theta) \subset \{\text{closure of } \Theta\}$ for any Θ.

Definition 14. Preferential identification. An asymptotically efficient identification process is also said to be preferential (with respect to the preference $\mathbf{P}$) if and only if

$$\hat{\theta}_\infty \in \mathbf{P}(\hat{\Theta}) \quad \text{for some } \hat{\Theta} \text{ containing } \Theta_\infty.$$

Remark. A simpler definition would result from using Θ_∞ alone rather than a set which contains it. However, it is difficult to build an identifier which makes such thorough use of information as to nonconservatively characterize $\Theta(t)$. It is easier to cover $\Theta(t)$ and, asymptotically, to cover Θ_∞. Consequently, the preferential identification property defined above is easier to achieve in practice.

B. Noise and unmodeled dynamics estimation

In this section we illustrate the potential utility of preferential identification with an example. The example is interesting in its own right: on-line modeling of the size of non-parametric uncertainty. This can be done in many ways with the general parameterization $M(\theta)$; we give only a simple example here.

Consider the special case of Fig. 1 given by Fig. 4. In the figure, θ_1, the first parameter in the parameter vector θ, multiplies the noise and the unmodeled dynamics outputs.

The constant parameter θ_1 scales the size of the non-parametric uncertainty set. Recall that $\mathbf{N}$ and $\mathbf{\Delta}$ are convex sets containing zero. Let $\theta_1\mathbf{N}$ and $\theta_1\mathbf{\Delta}$ denote the scaled uncertainty sets:

$$\hat{n} \in \theta_1\mathbf{N} \quad \text{if and only if} \quad \frac{\hat{n}}{\theta_1} \in \mathbf{N}, \qquad (8a)$$

$$\hat{\Delta} \in \theta_1\mathbf{\Delta} \quad \text{if and only if} \quad \frac{\hat{\Delta}}{\theta_1} \in \mathbf{\Delta}. \qquad (8b)$$
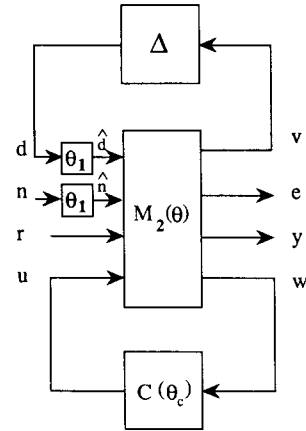

FIG. 4. Uncertainty scaling parameter.

The system M_2 sees a noise $\hat{n} \in \theta_1\mathbf{N}$ and $\hat{d} = \theta_1 \Delta v = \hat{\Delta} v$ where $\hat{\Delta} \in \theta_1\mathbf{\Delta}$.

The gain schedule f can take advantage of the uncertainty scaling θ_1, but is not required to do so. Regardless, the worst case performance of the gain scheduled system is likely to improve as the uncertainty set gets smaller.

A preferential identification process can take advantage of the performance improvement due to smaller-than-worst-case actual uncertainty sets in the following manner. Given a consistent parameter set $\Theta(t)$, the identification process could choose an element which minimizes the uncertainty scaling θ_1. Since $\Theta(t)$ is decreasing with time, $\hat{\theta}_1(t)$ would be monotonically increasing. Since $\hat{\theta}_1$ would also have an upper bound equal to the smallest value which covers the actual uncertainty, it follows that $\hat{\theta}_1(t)$ approaches a limit no larger than the actual uncertainty magnitude.

As we will state formally in the next subsection, the asymptotic performance of the adaptive system will be at least as good as one would achieve if one knew the uncertainty scaling. In effect, if one's prior bounds on the uncertainty were conservative, the adaptation can reduce the conservatism and achieve improved performance.

C. Asymptotic performance preferential identification

Let $\bar{r}$ be a fixed chosen constant. Let $\mathbf{P}_2$ give preference to parameter estimates for which the gain scheduled system gives the best performance. That is, let $\mathbf{P}_2$ be defined by

$$\mathbf{P}_2(\Theta) = \arg \inf_{\theta \in \Theta} J_2(\bar{r}, \theta), \qquad (9)$$

where the argument of the infimum need not be unique.

Theorem 2. Given assumptions 1-17 as in Theorem 1, and in addition

Assumption 18. The identification process is preferential with respect to $\mathbf{P}_2$, and

Assumption 19. $\|r\|^{t,\sigma} \leq \bar{r}$ $\forall t$ then it follows that the adaptive system also achieves an asymptotic performance level of $J_2(\bar{r}, \theta)$.

Proof. See Appendix.

Paraphrasing the theorem, the adaptive system performs at least as well (in the J_2 sense) as the gain scheduled system corresponding to the true plant parameter vector θ. That is, it performs as well as $S(\theta, f(\theta))$ even though the asymptotic system is $S(\theta, f(\hat{\theta}_\infty))$.

Note that there are two fundamental differences between Theorems 1 and 2. First, two different performance measures (J_1 and J_2) are used. Second, different gain scheduled systems are used in the comparison. Theorem 1 used the gain scheduled system corresponding to $\hat{\theta}_\infty$, and Theorem 2 used the gain scheduled system corresponding to the true plant parameter vector θ.

Corollary. Under the conditions of the theorem, the prior worst case performance guarantee (in the J_2 sense) for the adaptive system is equal to the performance guarantee $\bar{J}_2(\bar{r})$ for the gain schedule graph.

5. SIMPLE EXAMPLE OF AN EFFICIENT ESTIMATOR

The purpose of this paper is to provide a decomposition of robust adaptive control into robust control and robust estimation subproblems. It is not our intention to solve these formidable subproblems in this paper. For selected subsets of the robust control subproblem, solutions can be found in the robust control literature. On the estimation side, the literature lacks demonstrations of efficiency of deterministic estimators, since the notion of efficiency given in this paper is new.

Since it is our intention to inspire research into efficient estimation, it is worth demonstrating at the present time that efficient estimators do exist for a nonempty set of special cases. We demonstrate existence here by construction of a simple example for a simple special case.

Consider the system

$$y(t) = \theta\phi(t) + d(t), \tag{10}$$

$$d(t) = \Delta v(t) + n(t), \tag{11}$$

where θ is an uncertain real parameter, Δ is a neglected uncertainty (not to be identified), n is an unknown noise term, ϕ and v are known inputs, and y is a known output. Let the chosen prior membership sets for Δ and n be:

$$\Delta = \{\Delta : \Delta \text{ is real and } |\Delta| \leq 1\}, \tag{12}$$

$$N = \{n : |n(t)| \leq 1 \; \forall t\}. \tag{13}$$

To avoid complicating mathematical nuances later, assume that the membership sets above were chosen to be slightly conservative. That is, assume that it is actually known that for some $\varepsilon_0 > 0$,

$$\Delta \in \Delta_2 = \{\Delta : \Delta \text{ is real and } |\Delta| \leq 1 - \varepsilon_0\}, \tag{14}$$

$$n \in N_2 = \{n : |n(t)| \leq 1 - \varepsilon_0 \; \forall t\}. \tag{15}$$

Assume that the entire adaptive control design and analysis as described in this paper is carried out using the sets Δ and N. We will use the knowledge of the ε-conservatism in the residual uncertainty bounds only in establishing the efficiency of the estimator (below). (*Remark.* We do not know if the ε-conservative bounding technique is necessary, or if it is even useful for more general problems.)

The set of all $d(t)$ which are consistent with the known inputs and uncertainty assumptions (12) and (13) is

$$d \in \{d(t) : |d(t)| \leq |v(t)| + 1 =: \bar{d}(t)\}. \tag{16}$$

From (16) and (10) it is easily seen that the consistent parameter set is

$$\Theta(t) = \bigcap_{\tau \in [0,t), \; \phi(\tau) \neq 0} \left[\frac{y(\tau) - \bar{d}(\tau)}{\phi(\tau)}, \frac{y(\tau) + \bar{d}(\tau)}{\phi(\tau)} \right]. \tag{17}$$

Then Θ_∞^+ must be the set of all θ such that there exists a function $\hat{y}$ satisfying both

$$\theta \in \bigcap_{\tau \in [0,t), \; \phi(\tau) \neq 0} \left[\frac{\hat{y}(\tau) - \bar{d}(\tau)}{\phi(\tau)}, \frac{\hat{y}(\tau) + \bar{d}(\tau)}{\phi(\tau)} \right], \tag{18a}$$

and

$$\|\hat{y} - y\|^{\sigma,t} \to 0 \quad \text{as} \quad t \to \infty. \tag{18b}$$

Let

$$e(t) := y(t) - \hat{\theta}(t)\phi(t) = (\theta - \hat{\theta}(t))\phi(t) + d(t)$$
$$= \tilde{\theta}(t)\phi(t) + d(t), \tag{19}$$

where the last equality is by definition of $\tilde{\theta}(t) := \theta - \hat{\theta}(t)$.

Let T_1 be the set of all t such that $|e(t)| > \bar{d}(t)$. Let the estimation law be

$$\frac{d}{dt}\hat{\theta}(t) = \begin{cases} 0 & \text{if } t \notin T_1 \\ \phi(t)e(t) & \text{if } t \in T_1 \end{cases}. \tag{20}$$

Then,

$$\frac{d}{dt}(\bar{\theta}^2(t)) = -2(\theta - \hat{\theta}(t))\phi(t)e(t)$$

$$= -2(e(t) - d(t))e(t) \quad \forall t \in T_1, \quad (21)$$

where the last equality follows from (19). Now the ε-conservatism in the Δ assumption implies that for some $\varepsilon_2 > 0$,

$$\frac{d}{dt}(\bar{\theta}^2(t)) \le -\varepsilon_2 e^2(t) \quad \forall t \in T_1. \quad (22)$$

Then since $\bar{\theta}^2(t)$ is nonincreasing and bounded below by zero, it has a limit (so $\hat{\theta}(t)$ has a limit) and (22) implies

$$\left(\int_{t \in T_1} e^2(t)dt\right) \text{ is finite.} \quad (23)$$

Now consider the following candidate $\hat{y}$:

$$\hat{y}(t) = \hat{\theta}_\infty \phi(t)$$
$$+ \lim [-\bar{d}(t), \bar{d}(t)]^*(y(t) - \hat{\theta}_\infty \phi(t)), \quad (24)$$

where

$$\lim [a, b]^* c = \begin{cases} a & \text{if } c \le a \\ b & \text{if } c \ge b \\ c & \text{otherwise.} \end{cases}$$

Let T_2 be the set of all t such that $|y(t) - \hat{\theta}_\infty \phi(t)| > \bar{d}(t)$. Clearly,

$$\hat{y}(t) - y(t) = 0 \quad \forall t \notin T_2, \quad (25a)$$

$$|\hat{y}(t) - y(t)| = |y(t) - \hat{\theta}_\infty \phi(t)|$$
$$- \bar{d}(t) \quad \forall t \in T_2. \quad (25b)$$

Since the parameter estimate can only improve as time evolves, one finds that

$$|y(t) - \hat{\theta}_\infty \phi(t)| < |y(t) - \hat{\theta}(t)\phi(t)|$$
$$= |e(t)| \quad \forall t \in T_2. \quad (26)$$

Thus, $T_2 \subset T_1$, and it follows from (23), (25) and (26) that $\hat{y} - y$ is square integrable. Then for any $\sigma > 0$, $\|\hat{y} - y\|^{\sigma,t} \to 0$ as $t \to \infty$.

This completes a proof of asymptotic efficiency of the estimation law.

Remark. In this example, asymptotic efficiency happened to hold for any strictly positive choice of σ in the norm definition. The theorems of this paper only require that the asymptotic efficiency be demonstrated for a particular choice of σ.

6. CONCLUSIONS

On the philosophical side, this paper provides a language for discussion of the performance of robust identification processes and robust adaptive controllers. The language is pragmatic in that the various notions of controller performance arise from a robust control setting. In addition, the identification properties used here have a connection to the familiar notion of efficiency which arises in classical stochastic estimation. This paper makes precise the concept and terminology extensions required to talk about adaptive, rather than fixed, systems, and deterministic rather than stochastic uncertainty.

On the theoretical side, this paper supplies formal proofs of robust asymptotic performance for classes of adaptive systems. In effect, it states a separation principle: robust control paired with efficient estimation provides robust adaptive control. The parameterization and uncertainty characterizations for which this separation principle holds are exceptionally general. In fact, the separate problems of robust control and efficient estimation have not been solved in such a general setting. Still, current robust control can handle a significant subclass of problems, while efficient estimation is a new idea and is not nearly as far along. In effect, this paper provides a goal (namely asymptotic efficiency) for robust identification research, and reveals the adaptive controller performance payoff of achieving the goal.

REFERENCES

Becker, A., P. R. Kumar and C. Z. Wei (1985). Adaptive control with the stochastic approximation algorithm: Geometry and convergence. *IEEE Trans. Aut. Control,* **AC-30,** 330–338.

Bode, H. W. (1945). *Network Analysis and Feedback Amplifier Design.* Van Nostrand, Princeton, NJ.

Doyle, J. C. and G. Stein (1981). Multivariable feedback design: Concepts for a classical/modern synthesis. *IEEE Trans. Aut. Control,* **AC-26,** 4–16.

Doyle, J. C., J. E. Wall and G. Stein (1982). Performance and robustness analysis for structured uncertainty. *Proceedings of the 20th Conference on Decision and Control.*

Doyle, J. C. (1985). Structured uncertainty in control system design. *Proc. 24th Conf. on Decision and Control,* Ft Lauderdale, FL, 260–265.

Enns, D. F. (1987). Multivariable flight control for an attack helicopter. *IEEE Control Syst. Magazine,* **7,** 34–38.

Enns, D. F., D. J. Bugajski and M. J. Klepl (1987). Flight control for the F-8 oblique wing research aircraft. *Proc. American Control Conference,* Minneapolis, MN.

Enns, D. F. and D. J. Bugaski (1987). Multivariable control law analysis for a large space antenna. *Proc. SPIE Conf.,* Los Angeles., CA.

Goodwin, G. C., P. J. Ramadge and P. E. Caines (1980). Discrete time multivariable adaptive control. *IEEE Trans. Aut. Control,* **AC-25,** 449–456.

Goodwin, G. C. and M. E. Saldago (1989). Quantification of uncertainty in estimation using an embedding principle. *Proc. 1989 American Control Conference,* Pittsburgh, PA, 1416–1421.

Helmicki, A. J., C. A. Jacobson and C. Nett (1990a).

Identification in H-infinity: Linear algorithms. *Proc. American Control Conference*, San Diego, CA, pp. 2418–2423.

Helmicki, A. J., C. A. Jacobson and C. Nett (1990b). Identification in H-infinity: A robustly convergent nonlinear algorithm. *Proc. American Control Conference*, San Diego, CA, pp. 386–391.

Horowitz, I. M. (1963). *Synthesis of Feedback Systems.* Academic Press, New York.

Ioannou, P. A. and K. S. Tsakalis (1986). A robust direct adaptive controller. *IEEE Trans. Aut. Control*, AC-31, 1033–1043.

Kosut, R. L., M. Lau and S. Boyd (1990). Identification of systems with parametric and nonparametric uncertainty. *Proc. American Control Conference*, San Diego, CA, pp. 2412–2417.

Kosut, R. L. (1987). Adaptive uncertainty modeling: On-line robust control design. *Proc. American Control Conference*, Minneapolis, pp. 245–250.

Kosut, R. L. (1988). Adaptive robust control via transfer function uncertainty estimation. *Proc. American Control Conference*, Atlanta, pp. 349–354.

Krause, J. M., P. P. Khargonekar and G. Stein (1989). Robust adaptive control: Stability and asymptotic performance. *Proc. 28th Conf. on Decision and Control, (IEEE Trans. Aut. Control*, to appear).

Krisselmeier, G. and B. D. O. Anderson (1986). Robust model reference adaptive control. *IEEE Trans. Aut. Control*, AC-31, 127–133.

LaMaire, R., L. Valavani, M. Athans and G. Stein (1987). A frequency-domain estimator for use in adaptive control systems. *Proc. American Control Conference*, Minneapolis, MN, pp. 238–244.

Lehtomaki, N. (1981). Practical robustness measures in multivariable control system analysis. Ph.D Dissertation, Massachusetts Institute of Technology, MA.

Morse, A. S. (1980). Global stability of parameter adaptive control systems. *IEEE Trans. Aut. Control*, AC-25, 433–439.

Narendra, K. S., Y. H. Lin and L. S. Valavani (1980). Stable adaptive controller design, part II: Proof of stability. *IEEE Trans. Aut. Control*, AC-25, 440–448.

Narendra, K. S. and A. M. Annaswamy (1987). A new adaptive law for robust adaptation without persistent excitation. *IEEE Trans. Aut. Control*, AC-32, 134–145.

Ortega, R., L. Praly and I. D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Control*, AC-30, 1179–1187.

Poolla, K. and J. Shamma (1990). Asymptotic performance through adaptive robust control. *29th IEEE Conf. on Decision and Control*, Honolulu, 3176–3181.

Smith, R. S. and J. C. Doyle (1989). Model invalidation: A connection between robust control and identification. *Proc. American Control Conference*, Pittsburgh, 1435–1440.

Smith, R. S. (1990). Model validation for uncertain systems. Ph.D Dissertation, California Institute of Technology, CA.

Van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory.* Wiley, Chichester.

Willems, J. C. (1971). *The Analysis of Feedback Systems.* MIT Press, Cambridge, MA.

Younce, R. C. (1989). Identification with non-parametric uncertainty. Ph.D Dissertation, University of Notre Dame, IN.

APPENDIX: PROOFS

Proof of Theorem 1. Since $\theta_\infty \in \Theta_\infty^+$, there exist e_d, y_d, and w_d satisfying

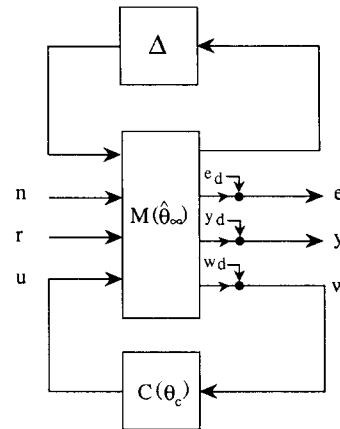$$\|e_d\|^{\sigma,t} + \|y_d\|^{\sigma,t} + \|w_d\|^{\sigma,t} \to 0 \quad \text{as} \quad t \to \infty$$

Fig. 5. Model for analysis.

such that the system of Fig. 5 with $n \in \mathbf{N}$ and $\Delta \in \Delta$ is a valid mathematical model for the constructed signals which result from the known inputs, for all $t \ge 0$.

By assumption 14, since $\hat{\theta}(t)$ has a limit in Θ_P as $t \to \infty$, $\theta_c(t)$ also has a limit as $t \to \infty$, namely $\theta_{c\infty} = f(\hat{\theta}_\infty)$.

Now take any $\varepsilon > 0$ and let $\delta > 0$ be chosen to satisfy the nondegeneracy expression (see definition 6). Since $\theta_c \to \theta_{c\infty}$, there exists a time $T < \infty$ such that $|\theta_c(t) - \theta_{c\infty}| < \delta$ for all $t \ge T$.

Prior to time T, $\hat{\theta}(t)$ is bounded (definition 10). Then since f is locally Lipschitzian, $\theta_c(t)$ is also bounded on $t \in [0, T]$. Then, by the well-posedness assumption on the interconnection of Δ and $M(\theta)$, all of the signals shown in Fig. 1 as well as the internal states of M exist (are bounded) on $[0, T]$. Now recall that the effect of nonzero initial conditions on M and past excitation of Δ can be reflected with an added transient noise term added to n. For the purpose of asymptotic analysis, shift the "start time" to time T (time T becomes the new time 0) and model the previous excitation (during $[0, T]$) by adding the appropriate n_d to n. Recall from Section II.C. that the added n_d satisfies $\|n_d\|^{\sigma,t} \to 0$ as $t \to \infty$.

For $t \ge T$, non-degeneracy of the gain scheduled system corresponding to $\hat{\theta}_\infty \in \Theta_P$ implies

$$\|e\|^{\sigma,t} \le (\alpha_2(\hat{\theta}_\infty) + \varepsilon)\|r\|^{\sigma,t} + (\alpha_1(\hat{\theta}_\infty) + \varepsilon)$$
$$+ (c + \varepsilon)(\|n_d\|^{\sigma,t} + \|e_d\|^{\sigma,t}$$
$$+ \|y_d\|^{\sigma,t} + \|w_d\|^{\sigma,t}),$$

which approaches $(\alpha_2(\hat{\theta}_\infty) + \varepsilon)\|r\|^{\sigma,t} + (\alpha_1(\hat{\theta}_\infty) + \varepsilon)$ as $t \to \infty$. Since ε is arbitrary, the theorem is proved.   □

Remark. The proof appears to make the commonplace argument that an asymptotically autonomous robust system acts in some respects like its limit. However, the proposition which is proved is subtly different. The asymptotic performance is that of a gain scheduled system with the plant parameters equal to $\hat{\theta}_\infty$, yet the actual plant parameters θ need not equal $\hat{\theta}_\infty$. The system merely behaves as if they do for the particular excitation and within the tolerances implied by the inclusion of n and Δ.

Proof of Theorem 2. By Theorem 1, the system achieves an asymptotic performance level of $J_1(t, r, \hat{\theta}_\infty)$ which equals $\alpha_1(\hat{\theta}_\infty) + \alpha_2(\hat{\theta}_\infty)\|r\|^{\sigma,t}$. Applying the bound on $\|r\|^{\sigma,t}$, the performance meets or exceeds $\alpha_1(\hat{\theta}_\infty) + \alpha_2(\hat{\theta}_\infty)\bar{r} = J_2(\bar{r}, \hat{\theta}_\infty)$.

Of course, the true parameters θ are in Θ_∞. The preferential identification produces $\hat{\theta}_\infty \in \mathbf{P}_2(\hat{\Theta})$ where $\Theta_\infty \subset \hat{\Theta}$. Thus $J_2(\bar{r}, \hat{\theta}_\infty) \le J_2(\bar{r}, \theta)$, by the definition of $\mathbf{P}_2$.   □