Systems & Control Letters 3 (1983) 311–319
North-Holland

December 1983

On parameter convergence in adaptive control *

Stephen BOYD and Shankar SASTRY

Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720, USA

Received 27 July 1983
Revised 8 September 1983

It is well known that the parameter error as well as the model–plant mismatch error in a model reference adaptive scheme tends exponentially to zero iff a certain *sufficient richness* condition holds for signals inside the time-varying plant control loop. In this paper we give conditions on the reference signal (the exogenous input to the adaptive loop) – namely, that it have as many spectral lines as there are unknown parameters, in order to guarantee parameter convergence.

Keywords: Model reference adaptive systems, Parameter convergence, Sufficient richness, Persistent excitation.

1. Problem statement

In recent work [1,2,8] on continuous time model reference adaptive systems, it has been shown that under a suitable choice of adaptive control law the output of the controlled plant y_P asymptotically tracks the output y_M of a stable reference model, despite the fact that the parameter error vector may not convergence to zero (indeed, it may not converge at all). Consider, for example, the case when the reference input is a step. In this case it may be shown that the parameter error vector converges, not necessarily to zero but to a value such that the (asymptotic) closed loop plant transfer function matches the model transfer function at D.C. (0 rad/sec). This observation suggests the following intuitive argument: assuming that the parameter vector *does* converge, the plant loop is 'asymptotically time invariant'. If the input r has spectral lines at frequencies $\nu_1,\ldots,\nu_N$, we expect y_P will also; since $y_P \to y_M$, we 'conclude' that the asymptotic closed loop plant transfer function matches the model transfer function at $s = j\nu_1,\ldots,j\nu_N$. If N is large enough, this implies that the asymptotic closed loop transfer function *is precisely* the model transfer function so that the parameter error converges to *zero*. It is the purpose of this paper to make this intuitive argument formal.

Results that have appeared in the literature on parameter error convergence (notably [3,4,5,13]) have established the uniform asymptotic and (equivalently) the exponential stability of the adaptive schemes under a certain *sufficient richness condition*. As is widely recognized, e.g. [14], the principal drawback to this condition is that it applies to a certain vector of signals $w(t)$ appearing *inside* the *time varying* feedback loop around the unknown plant. As a result, it is presently impossible to determine a priori whether a given reference input will result in a sufficiently rich $w(t)$ and subsequent parameter error convergence to zero. In this paper, we remedy this deficiency. Specifically, we show that *when the reference input* (which is the exogenous input to the adaptive system) *has as many spectral lines* as there are unknown parameters, then the output error $y_P - y_M$ and parameter error converge to zero exponentially. We also sketch how prior parameter and plant–model state error bounds can be used along with the methods of [4] to give an estimate of the rate of exponential convergence.

We agree with the authors of [12] that the issue of parameter convergence is important, not just for its own sake, but as a first step in tackling important questions like robustness to unmodelled dynamics, slowly time-varying plants, etc. that have recently been raised (e.g. [9,10]).

The organization of the paper is as follows: Section 2 briefly describes the model reference adaptive system; in Section 3, we state and prove our main result for the relative degree 1 case; in Section 4, we discuss the extension to the higher relative degree cases. Section 5 contains concluding remarks.

* Research supported in part by the Air Force Office of Scientific Research (AFSC) United States Air Force Contract F49620-79-C-0178.

S. Boyd gratefully acknowledges the support of the Fannie and John Hertz Foundation.

2. The model reference adaptive system

To fix notation, we briefly review the model reference adaptive system of Narendra, Valavani, et al. [1,2]. The single-input single-output plant is assumed to be represented by a transfer function

$$\hat{W}_P(s) = k_P \frac{\hat{n}_P(s)}{\hat{d}_P(s)} \tag{2.1}$$

where $\hat{n}_P(s)$, $\hat{d}_P(s)$ are relatively prime monic polynomials of degree m, n respectively and k_P is a scalar. The following are assumed known about the plant transfer function:

(A1) The degree of the polynomial $\hat{d}_P$, i.e. n, is known.

(A2) The relative degree of $\hat{W}_P$, i.e. $(n-m)$, is known.

(A3) The sign of k_P is known (say, + without loss of generality).

(A4) The transfer function $\hat{W}_P$ is assumed to be minimum phase, i.e. $\hat{n}_P$ is Hurwitz.

Remark. (A1) may be replaced by the weaker assumption that an upper bound on the degree of $\hat{d}_P$ is known. We use (A1) here for simplicity.

The objective of adaptive control is to build a dynamic compensator so that the plant output asymptotically matches that of a stable reference model $\hat{W}_M(s)$ with input $r(t)$, output $y_M(t)$ and transfer function

$$\hat{W}_M(s) = k_M \frac{\hat{n}_M(s)}{\hat{d}_M(s)} \tag{2.2}$$

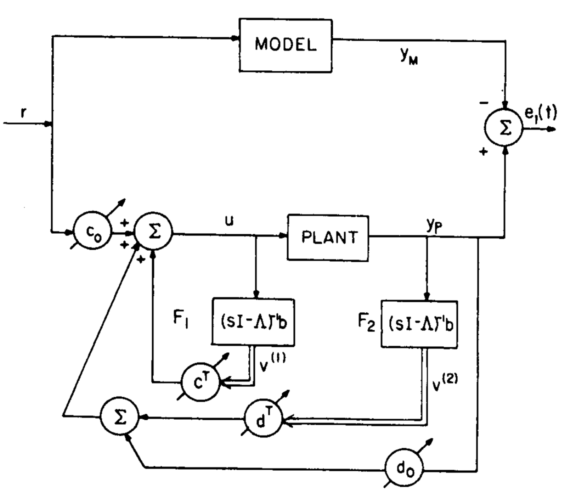

Fig. 1. The adaptive system for the relative degree 1 case.

where $\hat{n}_M$, $\hat{d}_M$ are monic polynomials of degree m^*, n^* respectively, $k_M > 0$. Since our interest in this paper is in parameter convergence we will assume $n^* = n$, $m^* = m$. We do not, however, need $\hat{n}_M$ and $\hat{d}_M$ to be relatively prime. If we denote the input and output of the plant $u(t)$ and $y_P(t)$ respectively, the objective may be stated as: choose $u(t)$ such that $y_P(t) - y_M(t) \to 0$ as $t \to \infty$.

2.1. *Relative degree* 1 *case*

By suitable prefiltering, if necessary, we may assume that the model $\hat{W}_M(s)$ is strictly positive real. The adaptive scheme in this case is as shown in Figure 1.

The dynamic compensation blocks F_1, F_2 are identical one input, $(n-1)$ output systems, each with transfer function

$$(sI-\Lambda)^{-1}b; \quad \Lambda \in \mathbb{R}^{(n-1)\times(n-1)}, b \in \mathbb{R}^{(n-1)},$$

where Λ is chosen so that the eigenvalues of Λ are the zeros of $\hat{n}_M$. We assume that the pair (Λ, b) is in controllable canonical form so that

$$(sI-\Lambda)^{-1}b = \frac{1}{\hat{n}_M(s)} \begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-2} \end{bmatrix}. \tag{2.3}$$

The adaptive gains $c \in \mathbb{R}^{n-1}$ are in the pre-compensator block for the purpose of cancelling the plant zeros and replacing them by the model zeros, $d \in \mathbb{R}^{n-1}$, $d_0 \in \mathbb{R}$ in the feedback compensator for the purpose of assigning the plant poles. The adaptive gain c_0 adjusts the overall plant gain. Thus, the vector of $2n$ adjustable parameters denoted θ is

$$\theta^T = [c_0, c^T, d_0, d^T].$$

If the signal vector $w \in \mathbb{R}^{2n}$ is defined by

$$w^T = [r, v^{(1)T}, y_P, v^{(2)T}], \tag{2.4}$$

we see that the input to the plant u is given by

$$u = \theta^T w. \tag{2.5}$$

It may be verified that there exists a unique constant $\theta^* \in \mathbb{R}^{2n}$ such that when $\theta = \theta^*$, the transfer function of the plant plus controller equals $\hat{W}_M(s)$. Further, it has been shown that under the update law

$$\dot{\theta} = -e_1 w \tag{2.6}$$

then $\lim_{t\to\infty} e_1(t) = 0$ provided $r(t)$ is bounded. Further, all signals in the loop, viz. $u(t)$, $v^{(1)}(t)$, $v^{(2)}(t)$, $y_P(t)$, $y_M(t)$ are bounded. Define the parameter error $\phi = \theta - \theta^*$. Then we have from [1] that

$$\phi \in L^2 \cap L^\infty, \quad \dot{\phi} \in L^\infty \quad \text{and } \dot{\phi} \to 0 \text{ as } t \to \infty.$$

However, we cannot say anything as yet about the convergence of $\phi(t)$ and hence of $\theta(t)$.

2.2. Relative degree 2 case

In this case $\hat{W}_M$ cannot be chosen positive real; however, we may assume (using suitable prefiltering, if necessary) that there is $L(s) = (s + \delta)$, with $\delta > 0$, such that $\hat{W}_M \hat{L}$ is positive real. The scheme of Figure 1 is modified (see [1]) [1] by replacing each of the gains θ_i, viz. c_0, d_0, c, d by the gain $\hat{L}\theta_i\hat{L}^{-1}$ which in turn are given by

$$\hat{L}\theta_i\hat{L}^{-1} = \theta_i + \dot{\theta}_i\hat{L}^{-1}, \quad i = 1, \ldots, 2n.$$

We now define the signal vector

$$\zeta^T(t) = \left[\hat{L}^{-1}r, \hat{L}^{-1}v^{(1)}, \hat{L}^{-1}y_P, \hat{L}^{-1}v^{(2)}\right]. \tag{2.7}$$

Then

$$\dot{\theta} = -e_1\zeta \tag{2.8}$$

yields that $e_1(t) \to 0$ as $t \to \infty$ provided $r(t)$ is bounded.

2.3. The case of relative degree $\geqslant 3$

As in Section 2.2, pick a stable Hurwitz polynomial $\hat{L}$ so that $\hat{L}\hat{W}_M$ is positive real. The trick

[1] Λ is now chosen to be exponentially stable, with the zeros of $\hat{n}_M$ a subset of the eigenvalues of Λ.

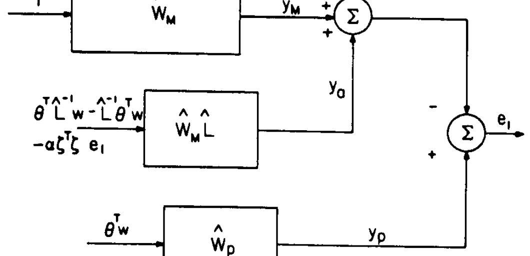

Fig. 2. Schematic of the adaptive system when the relative degree $\geqslant 3$.

used in Section 2.2, namely, to replace each θ_i by $\hat{L}\theta_i\hat{L}^{-1}$, is no longer possible since $\hat{L}\theta_i\hat{L}^{-1}$ depends on second and (possibly higher) derivatives of θ_i. To obtain a positive real error equation we retain the original configuration of Figure 1, and augment the model output by

$$\hat{W}_M\hat{L}[\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta^T]w$$

as shown in Figure 2. In addition to obtain $\dot{\phi} \in L^2$ and thereby prove stability of the adaptive scheme, we add an additional quadratic term to y_a to get the total augmented model output y_a

$$y_a = \hat{W}_M\hat{L}\left\{[\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta^T]w + \alpha\zeta^T\zeta e_1\right\} \tag{2.9}$$

where $\alpha > 0$ and ζ is defined in (2.7). The update law

$$\dot{\theta} = -e_1\zeta \tag{2.8}$$

yields that as $t \to \infty$, $e_1(t) \to 0$, $y_a(t) \to 0$ so that $y_M(t) \to y_P(t)$. As before, the parameter error ϕ satisfies

$$\phi \in L^2 \cap L^\infty, \quad \dot{\phi} \in L^\infty \quad \text{and} \quad \dot{\phi} \to 0 \text{ as } t \to \infty.$$

Again, nothing can be said about the convergence of $\phi(t)$.

3. Spectral lines and sufficient richness in the relative degree 1 case

Consider the adaptive system of Section 2.1 for the case of relative degree 1. We noted that the control law of (2.5) with the adaptive law of (2.6) yield that

$$\lim_{t\to\infty} e_1(t) = 0$$

provided $r(t)$ is bounded. Without additional conditions, however, we cannot guarantee

$$\lim_{t\to\infty} \theta(t) = \theta^*$$

(or in fact that θ converges at all). It has been shown by Morgan and Narendra [3], Anderson [4], Kreisselmeier [5] that $e_1(t) \to 0$, $\theta(t) \to \theta^*$ *exponentially* iff the signal vector $w(t)$ is *sufficiently rich*, in the following sense: There is $\delta > 0$, $\alpha > 0$ such that for all $s \in \mathbb{R}_+$

$$\int_s^{s+\delta} w(t)w^T(t)\,dt \geqslant \alpha I. \tag{3.1}$$

Recall from the definition of $w(t)$ in (2.5) that it contains signals $v^{(1)}(t)$, $v^{(2)}(t)$, $y_P(t)$ generated *inside* the *time varying* feedback loop around the *unknown* plant. Conditions on the reference input $r(t)$ required for (3.1) to hold are, to our knowledge, so far unknown. In the remainder of this section we will show that if $r(t)$ *has* $2n$ *spectral lines* (in a sense that will be made precise), *then we have exponential convergence of* $e_1(t)$ *to* 0 *and* $\theta(t)$ *to* θ^*. The proof is in two steps.

Step 1 consists of transcribing the condition (3.1) into an analogous condition for the model, which is a linear *time-invariant* system.

Step 2 consists of showing that the condition analogous to (3.1) for the model is obtained when the reference signal $r(t)$ has $2n$ spectral lines. We now discuss these steps in detail:

For Step 1, redraw Figure 1 as shown in Figure 3 with the model represented (in non-minimal form) as the plant with dynamic compensator and $\theta = \theta^*$. The signal vector $w_M \in \mathbb{R}^{2n}$ in the model-loop is given by

$$w_M^T = \left[r, v_M^{(1)}, y_M, v_M^{(2)}\right].$$

We have that $w_M \to w$ as $t \to \infty$. Hence, it seems reasonable to expect that if w_M is sufficiently rich then so is w. The foregoing is indeed true if $\dot{w}$ and $\dot{w}_M$ are bounded. However, we will use no supplementary assumptions on w, w_M but rather the conclusion from Narendra and Valavani [1] that $w(\cdot) - w_M(\cdot) \in L^2$. Further, it follows from their proof (specifically, Equations 16, 17, 18 of [1]) that

$$\begin{aligned}\|w(\cdot) - w_M(\cdot)\|_2 &\leqslant K_0\big(\|\theta(0) - \theta^*\| \\ &\quad + \|x_M(0) - x_P(0)\| + \|v^{(1)}(0) - v_M^{(1)}(0)\| \\ &\quad + \|v^{(2)}(0) - v_M^{(2)}(0)\|\big) \qquad (3.2)\end{aligned}$$

where x_M, x_P are the state variables in minimal representations for the plant in the model loop, plant loop respectively. Hence, from *prior bounds on the parameter error, and initial state errors* a bound on the L_2 norm of $w(\cdot) - w_M(\cdot)$ is obtained. Further, from [1], it follows that there is a K_2 such that

$$\|w(t)\|, \|w_M(t)\| \leqslant K_2 \quad \forall t. \qquad (3.3)$$

The bound K_2 depends as before on

$$\|\theta(0) - \theta^*\|, \quad \|x_M(0) - x_P(0)\|,$$

$$\|v^{(1)}(0) - v_M^{(1)}(0)\|, \quad \|v^{(2)}(0) - v_M^{(2)}(0)\|.$$

We now have:

Theorem 3.1. *Suppose*

$$\|w(t)\|, \|w_M(t)\| \leqslant K_2$$

and

$$\|w(\cdot) - w_M(\cdot)\|_2 = K_1 < \infty.$$

Then, $w(t)$ *is sufficiently rich* $\Leftrightarrow$ $w_M(t)$ *is sufficiently rich.*

Proof. The argument is symmetric between w and w_M. Hence, we only show ($\Rightarrow$). w sufficiently rich implies that $\exists\ \alpha, \delta > 0$ such that $\forall\ s \in \mathbb{R}_+$, $z \in \mathbb{R}^{2n}$

$$z^T\left[\int_s^{s+\delta} ww^T\, dt\right] z \geqslant \alpha z^T z. \qquad (3.4)$$

Iterating on (3.4) p times we get that $\forall\ p \in \mathbb{Z}_+$

$$z^T\left[\int_s^{s+p\delta} ww^T\, dt\right] z = \int_s^{s+p\delta} (z^T w)^2\, dt \geqslant \alpha p z^T z. \qquad (3.5)$$

Now, note that

$$(z^T w)^2 - (z^T w_M)^2 = z^T(w - w_M) z^T(w + w_M) \leqslant z^T z 2K_2 \|w - w_M\|.$$

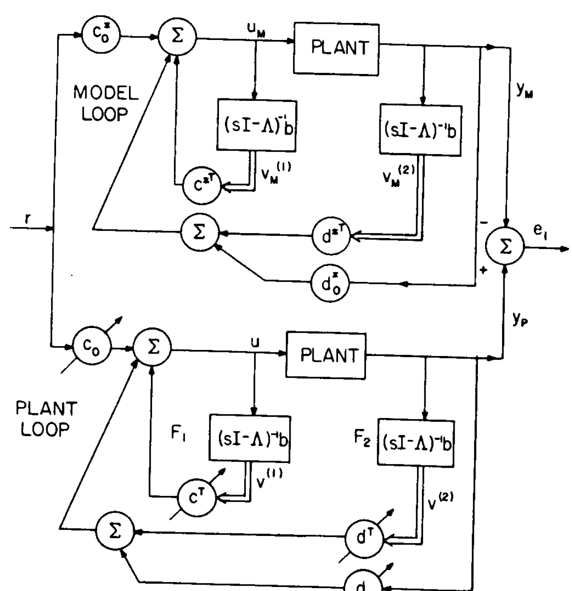

Fig. 3. The adaptive system of Figure 1 with a new representation for the model.

Hence

$$\int_s^{s+p\delta}(z^Tw)^2-(z^Tw_M)^2\,dt \leqslant z^Tz2K_2\int_s^{s+p\delta}\|w-w_M\|\,dt. \tag{3.6}$$

But, by Cauchy–Schwarz

$$\int_s^{s+p\delta}\|w-w_M\|dt \leqslant (p\delta)^{1/2}\int_s^{s+p\delta}\|w-w_M\|^2\,dt \leqslant K_1(p\delta)^{1/2}. \tag{3.7}$$

Using (3.7) in (3.6), and (3.4), we obtain that $\forall\, p\in\mathbb{Z}_+$

$$z^T\left[\int_s^{s+p\delta}w_Mw_M^T\,dt\right]z \geqslant z^Tz\left(\alpha p-2K_2K_1(p\delta)^{1/2}\right).$$

Choose p_0 sufficiently large so that

$$\bar\alpha := \alpha p_0 - 2K_2K_1(p_0\delta)^{1/2} > 0$$

and define $\bar\delta = p_0\delta$. Then we have that $\forall\, s\in\mathbb{R}_+$

$$\left[\int_s^{s+\bar\delta}w_Mw_M^T\,dt\right] \geqslant \bar\alpha I. \tag{3.8}$$

Thus w_M is sufficiently rich. □

Remark. We have shown that we have exponential convergence of parameter error and $e_1(t)$ provided that w_M is sufficiently rich (i.e. (3.8) holds). This completes Step 1.

Step 2. We now give conditions on $r(t)$ so that $w_M(t)$ is sufficiently rich, using the classical concept of a spectral line (see Wiener [6]).

Definition 3.2. A function $u(t):\mathbb{R}_+\to\mathbb{R}^n$ is said to have a *spectral line at frequency* ν *of amplitude* $\hat u(\nu)\in\mathbb{C}^n$ iff

$$\frac{1}{T}\int_s^{s+T}u(t)\,e^{-j\nu t}\,dt \tag{3.9}$$

converges to $\hat u(\nu)$ as $T\to\infty$, uniformly in s. When $\hat u(\nu)\neq 0$ we will say that *u has a spectral line at* ν.

Remark. u does *not* have to be *almost periodic* to have a spectral line at frequency ν_0; for example (3.9) need not converge for $\nu\neq\nu_0$.

The following lemma is immediate:

Lemma 3.3. *Let $u(t)$, $y(t)$ be the input and output, respectively, of a stable linear time-invariant system with transfer function $\hat L(s)$ (and arbitrary initial condition). If u has a spectral line at frequency ν then so does y, with amplitude*

$$\hat y(\nu) = \hat L(j\nu)\hat u(\nu). \tag{3.10}$$

Remark. Since the initial condition contributes a decaying exponential to $y(t)$ it does not appear in (3.10). $\hat y(\nu)$ in (3.10) may be zero if $\hat L(s)$ has a zero on the imaginary axis.

The second lemma is key to our main result:

Lemma 3.4. *Let $x(t)\in\mathbb{R}^N$ have spectral lines at frequencies $\nu_1, \nu_2,\ldots,\nu_N$. Further, let*

$$\{\hat x(\nu_1), \hat x(\nu_2),\ldots,\hat x(\nu_N)\}$$

be linearly independent in $\mathbb{C}^N$. Then, $x(t)$ is sufficiently rich, i.e. $\exists\ \alpha, \delta>0$ such that $\forall\, s\in\mathbb{R}_+$

$$\int_s^{s+\delta}xx^T\,dt \geqslant \alpha I. \tag{3.11}$$

Proof. Define the $N\times N$ matrix $X(s,\delta)$ by

$$X(s,\delta) := \frac{1}{\delta}\int_s^{s+\delta}\begin{bmatrix} e^{-j\nu_1 t}\\ \vdots \\ e^{-j\nu_N t}\end{bmatrix}x^T(t)\,dt$$

and the $N\times N$ matrix X_0 which is the (uniform in s) limit of $X(s,\delta)$ as $\delta\to\infty$,

$$X_0 := \begin{bmatrix}\hat x^T(\nu_1)\\ \vdots \\ \hat x^T(\nu_N)\end{bmatrix}.$$

By hypothesis X_0 is non-singular. Hence for δ sufficiently large $X(s,\delta)$ is invertible and

$$\|X(s,\delta)^{-1}\| \leqslant 2\|X_0^{-1}\|$$

for $\delta\geqslant\delta^*$ and all s. Now for $z\in\mathbb{R}^N$ with $\|z\|=1$, and any $\nu\in\mathbb{R}$ we have

$$\frac{1}{\delta}\int^{s+\delta}(x^Tz)^2\,dt = \frac{1}{\delta}\int_s^{s+\delta}|x^Tz\,e^{-j\nu t}|^2\,dt \geqslant \left|\frac{1}{\delta}\int_s^{s+\delta}x^Tz\,e^{-j\nu t}\,dt\right|^2 \tag{3.12}$$

(by Jensen's inequality). Using (3.12) for $\nu=\nu_1$,

$\nu_2,\ldots,\nu_N$ we have

$$\frac{1}{\delta}\int_s^{s+\delta}(x^{\mathrm{T}}z)^2\,\mathrm{d}t \geqslant \frac{1}{N}\sum_{k=1}^{N}\left|\frac{1}{\delta}\int_s^{s+\delta}x^{\mathrm{T}}z\,\mathrm{e}^{-\mathrm{j}\nu_k t}\,\mathrm{d}t\right|^2$$
$$= \frac{1}{N}\|X(s,\delta)z\|^2$$
$$\geqslant \frac{1}{N}\|X(s,\delta)^{-1}\|^{-2} \quad \text{for } \delta \geqslant \delta_*$$
$$\geqslant \frac{1}{4N}\|X_0^{-1}\|^{-2}.$$

Equation (3.11) now holds with $\delta = \delta^*$ and

$$\alpha = \frac{1}{4N}\|X_0^{-1}\|^{-2} > 0. \qquad \square$$

We now apply Lemmas (3.3), (3.4) to prove the main result of this section.

Theorem 3.5. *Suppose $r(t)$ has spectral lines at ν_1, $\nu_2,\ldots,\nu_{2n}$. Then $w_{\mathrm{M}}(t)$ is sufficiently rich.*

Remark. Once we have shown $w_{\mathrm{M}}(t)$ is sufficiently rich, Theorem 3.1 guarantees that $w(t)$ is also sufficiently rich which in turn guarantees exponential convergence of $e_1(t)$ to 0 and $\theta(t)$ to θ^*.

Proof. Recall that

$$w_{\mathrm{M}}^{\mathrm{T}}(t) = \left[r,\, v_{\mathrm{M}}^{(1)\mathrm{T}},\, y_{\mathrm{M}},\, v_{\mathrm{M}}^{(2)\mathrm{T}}\right].$$

We derive the transfer function from $r(t)$ to $w_{\mathrm{M}}^{\mathrm{T}}(t)$; using (2.3)

$$\hat{Q}^{\mathrm{T}}(s) = \left[1,\, \frac{\hat{W}_{\mathrm{M}}}{\hat{W}_{\mathrm{P}}}\frac{1}{\hat{n}_{\mathrm{M}}},\, \frac{\hat{W}_{\mathrm{M}}}{\hat{W}_{\mathrm{P}}}\frac{s}{\hat{n}_{\mathrm{M}}},\ldots,\frac{\hat{W}_{\mathrm{M}}}{\hat{W}_{\mathrm{P}}}\frac{s^{n-2}}{\hat{n}_{\mathrm{M}}},\right.$$
$$\left.\hat{W}_{\mathrm{M}},\, \frac{\hat{W}_{\mathrm{M}}s}{\hat{n}_{\mathrm{M}}},\ldots,\frac{\hat{W}_{\mathrm{M}}s^{n-2}}{\hat{n}_{\mathrm{M}}}\right]$$
$$= \frac{k_{\mathrm{M}}}{k_{\mathrm{P}}\hat{n}_{\mathrm{P}}\hat{d}_{\mathrm{M}}}\left[\frac{k_{\mathrm{P}}\hat{n}_{\mathrm{P}}\hat{d}_{\mathrm{M}}}{k_{\mathrm{M}}},\, \hat{d}_{\mathrm{P}},\ldots,\hat{d}_{\mathrm{P}}s^{n-2},\right.$$
$$\left.k_{\mathrm{P}}\hat{n}_{\mathrm{P}}\hat{n}_{\mathrm{M}},\, k_{\mathrm{P}}\hat{n}_{\mathrm{P}},\ldots,k_{\mathrm{P}}\hat{n}_{\mathrm{P}}s^{n-2}\right]. \quad (3.13)$$

Since the plant is minimum phase and the model is stable the transfer function $\hat{Q}(s)$ in (3.31) is *stable*. Neglecting the initial conditions (which do not, anyhow, contribute to the spectral lines of $w_{\mathrm{M}}(t)$) we have

$$w_{\mathrm{M}}^{\mathrm{T}} = \hat{Q}^{\mathrm{T}}r(t).$$

Now, the $(n+1)$th entry of $\hat{Q}$ has numerator polynomial $\hat{n}_{\mathrm{P}}\hat{n}_{\mathrm{M}}$ with $\hat{n}_{\mathrm{M}}$ of degree $(n-1)$. Further the first entry of $\hat{Q}$ has numerator polynomial $\hat{n}_{\mathrm{P}}\hat{d}_{\mathrm{M}}$ with $\hat{d}_{\mathrm{M}}$ of degree n. Compare these terms with the last $(n-1)$ entries of $\hat{Q}$, viz. $\hat{n}_{\mathrm{P}},\ldots,\hat{n}_{\mathrm{P}}s^{n-1}$. Using *constant* row operations then we can write

$$w_{\mathrm{M}} = T\overline{w} = T\frac{1}{\hat{n}_{\mathrm{P}}\hat{d}_{\mathrm{M}}}\begin{bmatrix}\hat{d}_{\mathrm{P}}\\ \vdots\\ \hat{d}_{\mathrm{P}}s^{n-2}\\ \hat{n}_{\mathrm{P}}\\ \vdots\\ \hat{n}_{\mathrm{P}}s^{n-2}\\ \hat{n}_{\mathrm{P}}s^{n-1}\\ \hat{n}_{\mathrm{P}}s^{n}\end{bmatrix}r(t) \qquad (3.14)$$

for some $T \in \mathbb{R}^{2n\times 2n}$, a non-singular matrix. It follows that w_{M} is sufficiently rich iff $\overline{w}$ is sufficiently rich. Now by Lemma 3.3 $\overline{w}$ has spectral lines at $\nu_1,\ldots,\nu_{2n}$ of amplitude

$$\frac{1}{\hat{n}_{\mathrm{P}}(\mathrm{j}\nu_i)\hat{d}_{\mathrm{M}}(\mathrm{j}\nu_i)}\begin{bmatrix}\hat{d}_{\mathrm{P}}(\mathrm{j}\nu_i)\\ \vdots\\ \hat{d}_{\mathrm{P}}(\mathrm{j}\nu_i)(\mathrm{j}\nu_i)^{n-2}\\ \hat{n}_{\mathrm{P}}(\mathrm{j}\nu_i)\\ \vdots\\ \hat{n}_{\mathrm{P}}(\mathrm{j}\nu_i)(\mathrm{j}\nu_i)^{n}\end{bmatrix}, \quad i = 1,\ldots,2n.$$

By Lemma 3.4 we need only show that these vectors are linearly independent. If not, there is a row vector $[\beta:\gamma]$ with $\beta^{\mathrm{T}} \in \mathbb{R}^{n-1}$, $\gamma^{\mathrm{T}} \in \mathbb{R}^{n+1}$ such that

$$[\beta:\gamma]\begin{bmatrix}\hat{d}_{\mathrm{P}}(\mathrm{j}\nu_1) & \cdots & \hat{d}_{\mathrm{P}}(\mathrm{j}\nu_{2m})\\ \vdots & & \vdots\\ \hat{d}_{\mathrm{P}}(\mathrm{j}\nu_1)(\mathrm{j}\nu_1)^{n-2} & & \hat{d}_{\mathrm{P}}(\mathrm{j}\nu_{2n})(\mathrm{j}\nu_{2n})^{n-2}\\ \hat{n}_{\mathrm{P}}(\mathrm{j}\nu_1) & & \hat{n}_{\mathrm{P}}(\mathrm{j}\nu_{2n})\\ \vdots & & \vdots\\ \hat{n}_{\mathrm{P}}(\mathrm{j}\nu_1)(\mathrm{j}\nu_1)^{n} & \cdots & \hat{n}_{\mathrm{P}}(\mathrm{j}\nu_{2n})(\mathrm{j}\nu_{2n})^{n}\end{bmatrix} = 0. \qquad (3.15)$$

Defining

$$\hat{\beta}(s) = \beta_1 + \beta_2 s + \cdots + \beta_{n-1}s^{n-2}$$

and

$$\hat{\gamma}(s) = \gamma_1 + \gamma_2 s + \cdots + \gamma_{n+1} s^n,$$

we may write (3.15) as

$$\hat{\beta}(s)\hat{d}_P(s) + \hat{\gamma}(s)\hat{n}_P(s) = 0 \quad \text{at } s = j\nu_1, \ldots, j\nu_{2n}. \tag{3.16}$$

The polynomial in (3.16) has degree $(2n-1)$ so we conclude that it is identically 0 and

$$\hat{\beta}\hat{d}_P \equiv -\hat{\gamma}\hat{n}_P.$$

But, since $\hat{n}_P$ and $\hat{d}_P$ are coprime (by assumption) the zeros of $\hat{\beta}$ must include those of $\hat{n}_P$. But this is impossible since $\hat{\beta}$ has degree $n-2$ and $\hat{n}_P$ has degree $(n-1)$. This establishes the contradiction. Thus $\bar{w}$ and hence w_M are sufficiently rich. □

Comments. (1) We say that $r(t)$ is *persistently exciting at frequencies* $\nu_1, \ldots, \nu_{2n}$ if it has spectral lines at these frequencies. We have shown that when the reference input is persistently exciting at as many frequencies as there are unknown parameters, then $w(t)$ is sufficiently rich resulting in exponential parameter and error convergence.

(2) $r(t)$ does not have to be almost periodic [7] to satisfy the conditions of Theorem 3.5. It need only have spectral lines at $2n$ frequencies. Further the *strength* of the spectral lines figures only in an estimate of the *rate* of exponential convergence (which may be derived using the techniques of [4]). In particular a low intensity persistently exciting signal (i.e. having $2n$ spectral lines) may be added to the $r(t)$ that needs to be tracked in the model to guarantee parameter convergence – see also Remark 6 below.

(3) It is not widely appreciated in the literature that parameter convergence may not occur (even to an incorrect value), unless the signal $w(t)$ is sufficiently rich. If it were known that $\lim_{t\to\infty}\theta(t)$ exists, a more elementary proof could be given – though the convergence proven need not be either exponential or uniform.

(4) The hypothesis of the theorem can be weakened. For instance, we do not need $r(t)$ to have spectral lines at $\nu_1, \ldots, \nu_{2n}$; it is adequate that

$$\liminf \left| \frac{1}{T} \int_s^{s+T} r(t)\, e^{-j\nu_k t}\, dt \right| > 0 \quad \text{uniformly in } s$$

for $k = 1, \ldots, 2n$.

(5) Most periodic functions (specifically, those having at least $2n$ non-zero Fourier coefficients) for $r(t)$ yield exponential parameter convergence.

(6) Our estimate for the rate of convergence of the parameter error given the magnitude of the spectral line would (in principle) proceed as follows: use the estimates of Lemma 3.4 to obtain the α, δ in the definition of sufficient richness for w_M. Then, use the prior bounds on parameter and initial error to bound the L^2 difference between w and w_M, and obtain using Theorem 3.1 the α, δ in the definition of sufficient richness for w. From here, the techniques of [4] may be used to obtain a (conservative!) rate of convergence estimate.

4. Parameter convergence when the relative degree $\geqslant 2$

Consider first the relative degree 2 case of Section 2.2. In this case, the sufficient richness condition for exponential parameter and error convergence is on the signal vector $\zeta(t)$ of (2.7), i.e. $\exists\ \alpha$, $\delta > 0$, $\forall\ s \in \mathbb{R}_+$

$$\int_s^{s+\delta} \zeta\zeta^T\, dt \geqslant \alpha I. \tag{4.1}$$

Even though the adaptive scheme has changed, redraw the *model* exactly as in Figure 3. Now define from the w_M of the model the signal vector

$$\zeta_M^T = \left[\hat{L}^{-1}r,\ \hat{L}^{-1}v_M^{(1)T},\ \hat{L}^{-1}y_M,\ \hat{L}_M^{-1}v_M^{(2)T}\right], \tag{4.2}$$

i.e. ζ_M is obtained by filtering each component of w_M through the stable system with transfer function $\hat{L}^{-1}$. Now, if $r(t)$ has $2n$ spectral lines we have by Theorem 3.5 that $\hat{w}_M(\nu_1)$, $\hat{w}_M(\nu_2), \ldots, \hat{w}_M(\nu_{2n})$ are linearly independent. From the definition of ζ_M in (4.2) and the fact that $\hat{L}^{-1}(s)$ is stable, it follows that

$$\hat{\zeta}_M(\nu_i) = \frac{1}{\hat{L}(j\nu_i)}\hat{w}_M(\nu_i), \quad i = 1, \ldots, 2n,$$

are linearly independent. Hence ζ_M is sufficiently rich.

Further, the stability proof [1] yields that $\zeta(\cdot) - \zeta_M(\cdot) \in L^2$, so that ζ is sufficiently rich thereby guaranteeing exponential parameter convergence.

Now consider the scheme of Figure 2 for the relative degree $\geqslant 3$ case. Redraw the model as in Figure 3 and define ζ_M as in (4.2) above. The same argument, as above, yields that when r has $2n$