PERSISTENCY OF EXCITATION IN POSSIBLY UNSTABLE CONTINUOUS TIME SYSTEMS AND PARAMETER CONVERGENCE IN ADAPTIVE IDENTIFICATION

by

N. Nordström and S. S. Sastry

Memorandum No. UCB/ERL M86/30

11 April 1986

COVER PAGE

1

PERSISTENCE OF EXCITATION IN POSSIBLY UNSTABLE CONTINUOUS TIME SYSTEMS AND PARAMETER CONVERGENCE IN ADAPTIVE IDENTIFICATION

by

N. Nordström and S. S. Sastry

Memorandum No. UCB/ERL M86/30

11 April 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

PERSISTENCE OF EXCITATION IN POSSIBLY UNSTABLE CONTINUOUS TIME SYSTEMS AND PARAMETER CONVERGENCE IN ADAPTIVE IDENTIFICATION

by

N. Nordström and S. S. Sastry

Memorandum No. UCB/ERL M86/30

11 April 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Persistency of Excitation in Possibly Unstable Continuous Time Systems and Parameter Convergence in Adaptive Identification

N. Nordström and S.S. Sastry

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

ABSTRACT

We find conditions on the input signal of an output reachable *possibly unstable* linear system, under which the output is persistently exciting. The conditions are given in both frequency and in time domain versions. Interpreting these results in the context of controllable or observable state space realizations we obtain some interesting facts relating persistency of excitation of the input, state and output signals.

To illustrate the importance of our results we propose an adaptive identification scheme with "least squares" update law for multivariable plants with proper transfer function. We prove that parameter convergence is guaranteed for any stationary piecewise uniformly continuous input with nonzero minimum interdiscontinuity distance and at least $2n + 1$ points of strong support of its spectral measure, where n is the McMillan degree of the plant. With covariance resetting the convergence rate is shown to be exponential. Without covariance resetting we prove, that the convergence rate is as $1/t$ for sufficiently fast identifiers, and in any case at least as fast as $1/\sqrt{t}$.

Keywords: Persistent Excitation, Adaptive Identification, Parameter Convergence.

February 15, 1986

Research supported by: JSEP-AFOSR-F49620-79-C0178, Paulson Scholarship and Alfred F. Moore and Chella D. Moore Scholarship.

1. Introduction

It has been a well known fact for some time, that the state trajectory of a stable controllable linear time invariant system realization driven by a stationary input is persistently exciting if the input spectrum has at least n points of support, where n is the dimension of the realization. This has been proven for continuous time systems [1,2] as well as for discrete time systems [3]. For discrete time systems and stationary input signals the input spectrum condition has been shown to be equivalent to the, in discrete time more commonly used, time domain concept of persistency of excitation of order n [3,4]. In this form the result has been extended to include the output function of any output reachable system with the state trajectory of a controllable realization as a special case [5].

A somewhat unexpected difference between the continuous and the discrete time proofs, is that in continuous time the system has to be stable, an unnecessary requirement in the discrete time case. This observation has not surprisingly inspired attempts to prove, that the result is valid independently of system stability in continuous time as well as in discrete time [6]. At least one such "proof" has appeared in the literature, but we have found it to be incorrect, and as far as we know, a correct version has yet not been published. In this paper we therefore give another proof of this fact. More specifically, we determine input conditions, under which the output of an output reachable *possibly unstable* continuous time system with proper rational transfer function is persistently exciting. In particular these input conditions ensure that the state trajectory of a controllable realization is persistently exciting. The input conditions are given in both frequency domain and time domain versions. Our results are stated for general multivariable systems, and take advantage of the possibility of having a controllability index lower than the McMillan degree.

The difficulty with the proof for unstable systems is due to complications arising from the zero-input response. In discrete time this response can be cancelled by means of a trick involving the Cayley-Hamilton theorem. The usual way of translating results from discrete time to continuous time by replacing the shift operator by differentiation does not work in this case, because the differentiation operator, unlike the shift operator, is unbounded. We have therefore utilized a generalized continuous time version of the "Cayley-Hamilton trick", and thereby found a large class of interesting signals producing persistently exciting outputs.

To illustrate the importance of our results, we apply them to show parameter convergence of an adaptive identification scheme with least squares update law. The scheme we propose is a multi-input-single-output (MISO) version of the Narendra-scheme for single-input-single-output (SISO) systems [7]. It applies to multivariable proper possibly unstable plants of known McMillan degree. We prove that under appropriate but reasonable input conditions, the estimated parameters converge to the true parameters, and we find the convergence rate for schemes with and without covariance resetting to be exponential and as $1/t$ respectively.

During the course of proving these results we define a few new concepts such as "piecewise uniform continuity", "minimum interdiscontinuity distance", "strong support of a positive semidefinite matrix valued measure", etc. and develop a collection of useful results relating them to various other properties of functions.

We point out, that our results are not only of general interest, in that they remove part of the difference between what is known about continuous and discrete time systems. Our convergence proof for the adaptive identification scheme shows, that they are also of practical importance. Here the fact that persistency of excitation can be guaranteed regardless of stability, means that we can ensure parameter convergence, without any prior knowledge about the pole locations of the plant. It also means, that the convergence does not rely on the zero-input response to fade out, hence showing that the convergence rate is not directly related to the distance of the poles of the plant to the right half plane. This indicates some robustness of the scheme even when applied only to the class of stable plants.

The paper is organized as follows: In the following section we introduce some notation. In section 3 we define a few key concepts, and develop a set of useful related propositions. In section 4 we prove that the output of an output reachable system is persistently exciting under certain input conditions. In section 5 we then establish a few facts governing persistency of excitation of the input, the state and the output of a state space realization. Finally in section 6 we apply the results developed in section 4, to adaptive identification as discussed above. In the interest of brevity some of the more procedural proofs are omitted. These proofs are available by the author.

2. Notation

The following notation is used throughout this paper.

$\lfloor r \rfloor$ largest integer $\leqslant r$

$\lceil r \rceil$ smallest integer $\geqslant r$

e_i column vector with i th element = 1, and all other elements = 0

$\underline{x}_l :=$ $[x^{(0)T} \cdots x^{(l-1)T}]^T$ for any $l-1$ times differentiable function $x : \mathbf{R} \to \mathbf{C}^n$.

Maximum and minimum are denoted by $\vee$ and $\wedge$ respectively. A $\hat{}$ sign indicates a Fourier or Laplace transformed function. A * is used for complex conjugate as well as convolution. Finally all vector- and matrix norms are 2-norms.

3. Preliminary Definitions and Propositions

To be precise about what we mean by a few concepts, which will arise in some of the results later on, we begin this paper with some preliminary definitions. (This is necessary since some of these concepts are not standard in the literature.)

Definition 1: Let $X \subseteq \mathbf{R}$, and let (M, ρ) be a metric space. We say that the function $f : X \to M$ is *piecewise uniformly continuous* if

(i) f is piecewise continuous, i.e. $\exists$ a countable set $D \subseteq X$ such that f is continuous on $X \setminus D$ and $B \cap D$ is finite for every bounded set $B \subseteq \mathbf{R}$

(ii) $\forall \epsilon > 0$, $\exists \delta(\epsilon) > 0$ such that $\rho(f(x_2), f(x_1)) < \epsilon$
$\forall x_2 \in (\sup \{x_1 - \delta(\epsilon)\} \cup D \cap (-\infty, x_1] ; \inf \{x_1 + \delta(\epsilon)\} \cup D \cap [x_1, \infty))$, $\forall x_1 \in X$

We see that piecewise uniform continuity is nothing but piecewise continuity with a uniform modulus of continuity.

Definition 2: Let $f : M \to Y$ be a function from a metric space (M, ρ) to a topological space Y. Let $D \subseteq M$ be the set of discontinuity points of f. We define the ***minimum interdiscontinuity distance*** $\kappa(f)$ of f by

$$\kappa(f) := \inf_{d_1, d_2 \in D} \rho(d_1, d_2) \tag{3.1}$$

The following two propositions show that most of the input signals of our interest will more or less automatically be bounded with bounded derivatives.

Proposition 1: If a piecewise uniformly continuous signal $u : \mathbf{R} \to \mathbf{C}^m$ with minimum interdiscontinuity distance $\kappa > 0$ is stationary, i.e. has well defined autocorrelation function

$$R_u(\tau) := \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} u(t) u^H(t+\tau) dt \qquad \text{uniformly in } t_0 \tag{3.2}$$

then it is bounded.

Proof: Suppose u is stationary. Then $\exists T \in (0, \infty)$ such that

$$\left\| \frac{1}{T} \int_{t_0}^{t_0+T} u(t) u^H(t) dt - R_u(0) \right\| < 1 \qquad \forall t_0 \in \mathbf{R} \tag{3.3}$$

Let

$$\beta := \|R_u(0)\| + 1 \tag{3.4}$$

and

$$\bar{\kappa} := \kappa \wedge T \tag{3.5}$$

Then

$$\inf_{t \in [t_0, t_0+\bar{\kappa}]} \|u(t)\|^2 \leqslant \frac{1}{\bar{\kappa}} \int_{t_0}^{t_0+\bar{\kappa}} \|u(t)\|^2 dt \leqslant \frac{1}{\bar{\kappa}} \operatorname{tr} \int_{t_0}^{t_0+T} u(t) u^H(t) dt \tag{3.6}$$

$$\leqslant \frac{m}{\bar{\kappa}} \| \int_{t_0}^{t_0+T} u(t) u^H(t) dt \| < \frac{mT\beta}{\bar{\kappa}} \qquad \forall t_0 \in \mathbf{R}$$

Now u has minimum interdiscontinuity distance $\kappa \geqslant \bar{\kappa} > 0$, so $\forall t \in \mathbf{R} \; \exists$ a half open interval $I(t)$ containing t, with endpoints $t_0(t)$ and $t_0(t) + \bar{\kappa}$ on which u is continuous. Since moreover u is piecewise uniformly continuous, $\exists \delta > 0$ independent of t such that $\|u(t_2) - u(t_1)\| < 1 \;\; \forall t_1, t_2 \in I(t)$ with $|t_2 - t_1| < \delta$. Hence

$$\|u(t)\| \leqslant \inf_{t \in I(t)} \|u(t)\| + \left\lceil \frac{\bar{\kappa}}{\delta} \right\rceil \leqslant \left(\frac{mT\beta}{\bar{\kappa}}\right)^{\frac{1}{2}} + \frac{\bar{\kappa}}{\delta} + 1 < \infty \quad \blacksquare \tag{3.7}$$

Proposition 2: If $u: \mathbf{R} \rightarrow \mathbf{C}^m$ is bounded and $u^{(l)}$ is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$, then $u^{(0)}, \ldots, u^{(l)}$ are bounded and piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$.

Proof: Consider first a scalar signal $u: \mathbf{R} \rightarrow \mathbf{C}$. Since $u^{(l)}$ is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$, $\forall t_0 \in \mathbf{R} \; \exists$ an interval $I(t_0)$ containing t_0 and of length $\bar{\kappa} := \kappa \wedge 1$ such that $u^{(l)}$ is continuous on $I(t_0)$. Hence by Taylor's formula

$$u(t) = \sum_{i=0}^{l} a_i(t_0) \tau^i + \bar{a}_l(t_0, \tau) \tau^l \qquad \forall t \in I(t_0) \tag{3.8}$$

where

$$\tau = t - t_0 \tag{3.9}$$

$$a_i(t_0) = \frac{u^{(i)}(t_0)}{i!} \qquad i = 0, \ldots, l \tag{3.10}$$

$$\bar{a}_l(t_0, \tau) \leqslant M_{\bar{a}} < \infty \qquad \forall t_0 \in \mathbf{R} \tag{3.11}$$

By a tedious exercise in calculus one can then show that

$$|u^{(j)}(t_0)| \leqslant (l+1)!\left(\frac{2}{\bar{\kappa}}\right)^l (8l^2)^{l^2} (M_u + M_{\bar{a}}\bar{\kappa}^l) < \infty \tag{3.12}$$

$$\forall t_c \in \mathbf{R}, \ j = 0, \ldots, l$$

where

$$M_u := \sup_{t \in \mathbf{R}} |u(t)| < \infty \tag{3.13}$$

If u is multidimensional, the proof holds for each of its components, and hence for u itself. Finally since $u^{(1)}, \ldots, u^{(l)}$ are bounded, it follows that $u^{(0)}, \ldots, u^{(l-1)}$ are uniformly continuous. ∎

Definition 3: We say that $x: \mathbf{R} \to \mathbf{C}^n$ is *persistently exciting (p.e)* if $\exists$ constants $\Delta < \infty$ and $\alpha > 0$ such that

$$\int_{t_0}^{t_0+\Delta} x(t)x^H(t)dt \geqslant \alpha I \qquad \forall t_0 \in \mathbf{R} \tag{3.14}$$

The next proposition shows that a non-persistently exciting signal can in general not be made persistently exciting by filtering with a finite impulse response filter. This means that persistency of excitation can be inferred from a filtered version of a signal, and for some interesting filters linear combinations of derivatives of such a version.

Proposition 3: Consider two functions $q: \mathbf{R} \to \mathbf{C}$ and $y: \mathbf{R} \to \mathbf{C}^p$. Assume that $q \in L^2(\mathbf{R})$ has compact support, and that $q * y$ is persistently exciting. Then y is persistently exciting as well.

Proof: Let $\|q\|$ denote the L^2-norm of q, and let $[t_1, t_2]$ be a compact interval containing its support. Then using the Cauchy-Schwarz inequality, $\forall v \in \mathbf{C}^p$, $\forall t_0 \in \mathbf{R}$, $\forall \Delta \geqslant 0$, we have

$$v^H \int_{t_0}^{t_0+\Delta} (q * y)(t)(q * y)^H(t)dt \ v = \int_{t_0}^{t_0+\Delta} \left| \int_{-\infty}^{\infty} q(\tau)v^H y(t-\tau)d\tau \right|^2 dt \tag{3.15}$$

$$\leqslant \Delta \sup_{t \in [t_0, t_0+\Delta]} \Big| \int_{-\infty}^{\infty} q(\tau) v^H y(t-\tau) dt \Big|^2 \leqslant \Delta \sup_{t \in [t_0, t_0+\Delta]} \|q\|^2 \int_{t_1}^{t_2} |v^H y(t-\tau)|^2 d\tau$$

$$\leqslant \Delta \|q\|^2 \int_{t_0-t_2}^{t_0-t_1+\Delta} |v^H y(\tau)|^2 d\tau = \Delta \|q\|^2 v^H \int_{t_0-t_2}^{(t_0-t_2)+(t_2-t_1+\Delta)} y(\tau) y^H(\tau) d\tau \, v$$

Since $\|q\| < \infty$ and $t_2 - t_1 < \infty$, we see that the persistency of excitation of $q * y$ implies that of y. ∎

Definition 4: We say that $x : \mathbf{R} \to \mathbf{C}^n$ is *persistently exciting of order* l if x is $l-1$ times differentiable, and the vector valued function $\underline{x}_l := [x^{(0)T} \cdots x^{(l-1)T}]^T$ is persistently exciting.

The following definition generalizes the idea of support of a measure to the class of positive semidefinite matrix valued measures.

Definition 5: Let $S : B^1 \to \mathbf{C}^{m \times m}$ be a positive semi definite matrix valued measure on the Borel sets B^1 of $\mathbf{R}$.

We say that $\omega_0 \in \mathbf{R}$ is a *point of support* of S, and write $\omega \in \mathrm{supp}(S)$, if $S(O)$ is positive definite $\forall$ neighborhoods O of ω_0.

We say that ω_0 is a *point of strong support* of S, and write $\omega \in \mathrm{ssupp}(S)$, if

$$\inf_{\|v(\omega)\| \geqslant 1 \, \forall \, \omega \in \mathbf{R}} \int_O v^H(\omega) dS(\omega) v(\omega) > 0 \qquad (3.16)$$

$\forall$ neighborhoods O of ω_0.

Remarks:

1) The definitions above make perfect sense with $\mathbf{R}$ and the Borel sets on $\mathbf{R}$ replaced by any measurable space. This however, will not be used in the following discussion.

2) If S is one dimensional then $\mathrm{ssupp}(S) = \mathrm{supp}(S)$.

3) $\mathrm{ssupp}(S) \subseteq \mathrm{supp}(S)$, because

$$\inf_{\|v\| \geq 1} v^H S(O)v = \inf_{\|v\| \geq 1} \int_O v^H dS(\omega)v \geq \inf_{\|v(\omega)\| \geq 1 \,\forall\, \omega \in \mathbf{R}} \int_O v^H(\omega)dS(\omega)v(\omega) \quad (3.17)$$

4) A point of support may not be a point of strong support. Consider, for example, the case when

$$S(\Omega) = \int_\Omega \hat{r}(\omega)\hat{r}^H(\omega)d\,\omega \tag{3.18}$$

where $\hat{r}(\omega) \in \mathbf{C}^m$, $m \geq 2$ and the functions $\hat{r}_1(\omega), \ldots, \hat{r}_m(\omega)$ are linearly independent over $\mathbf{C}$ on every non degenerate interval. Then $S(O)$ is positive definite, but

$$\begin{aligned} &\inf_{\|v(\omega)\| \geq 1 \,\forall\, \omega \in \mathbf{R}} \int_O v^H(\omega)dS(\omega)v(\omega) \\ &= \inf_{\|v(\omega)\| \geq 1 \,\forall\, \omega \in \mathbf{R}} \int_O |v^H(\omega)\hat{r}(\omega)|^2 d\,\omega = 0 \end{aligned} \tag{3.19}$$

$\forall$ neighborhoods O of ω_0, $\forall\, \omega_0 \in \mathbf{R}$.

For the class of spectral measures the notion of support is related to persistency of excitation according to the following proposition.

Proposition 4: Let $x : \mathbf{R} \to \mathbf{C}^n : t \to x(t)$ be a stationary function with autocorrelation $R_x : \mathbf{R} \to \mathbf{C}^{n \times n}$, and spectral measure S_x. Then

$$\mathrm{supp}(S_x) \neq \emptyset \Rightarrow S_x(\mathbf{R}) > 0 \Leftrightarrow x \text{ is p.e.} \tag{3.20}$$

Proof: The proof, whose non-trivial part was given in [2], follows from the fact that

$$\frac{1}{T}\int_{t_0}^{t_0+T} x(t)x^H(t)dt \underset{T\to\infty}{\longrightarrow} R_x(0) = S_x(\mathbf{R}) \tag{3.21}$$

uniformly in t_0 on $\mathbf{R}$. We leave the details to the reader. ∎

To be able to make use of *proposition 3* in the sense described above, we need the following lemma, which proves the existence of FIR filters preserving the excitation properties of a signal.

Lemma 1: Consider a bounded signal $u : \mathbb{R} \to \mathbb{C}^m$, persistently exciting of order l and such that $u^{(l-1)}$ is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$. Let $\{d_a : \mathbb{R} \to \mathbb{C} \mid a \in (0,\infty)\}$ be a collection of integrable functions such that

$$\operatorname{supp}(d_a) \subseteq R(a) := [-r(a), r(a)] \tag{3.22}$$

where

$$r(a) \underset{a \to \infty}{\to} 0 \tag{3.23}$$

and

$$\left| \int_{-\infty}^{\infty} d_a(\tau) d\tau \right| \geq \beta > 0 \quad \forall a \in (0,\infty) \tag{3.24}$$

Then $d_a * u$ is also persistently exciting of order l for a large enough.

Proof: Since u is bounded and $u^{(l-1)}$ is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$, it follows from *proposition* 2 that $u^{(0)}, \ldots, u^{(l-1)}$ are all bounded and piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$. From this one can readily show that:

$$s(a) := \sup_{\substack{t_0 \in \mathbb{R} \\ \tau, \sigma \in R(a)}} \int_{t_0}^{t_0+\Delta} \| \underline{u}_l(t-\tau)\underline{u}_l^H(t-\sigma) - \underline{u}_l(t)\underline{u}_l^H(t) \| dt \underset{a \to \infty}{\to} 0 \tag{3.25}$$

Since u is persistently exciting of order l, $\exists \Delta < \infty$ and $\alpha > 0$ such that

$$\int_{t_0}^{t_0+\Delta} \underline{u}_l(t)\underline{u}_l^H(t) dt \geq \alpha I \tag{3.26}$$

Interchanging order of integration we have

$$\int_{t_0}^{t_0+\Delta} \underline{v}_l(t)\underline{v}_l^H(t) dt = \int_{t_0}^{t_0+\Delta} (d_a * \underline{u}_l)(t)(d_a * \underline{u}_l)^H(t) dt \tag{3.27}$$

$$= \left| \int_{-\infty}^{\infty} d_a(\tau) d\tau \right|^2 \int_{t_0}^{t_0+\Delta} \underline{u}_l(t)\underline{u}_l^H(t) dt$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d_a(\tau) \int_{t_0}^{t_0+\Delta} [\underline{u}_l(t-\tau)\underline{u}_l^H(t-\sigma) - \underline{u}_l(t)\underline{u}_l^H(t)] dt \; d_a^*(\sigma) d\tau \, d\sigma$$

$$\geq \beta\alpha I - \left(\int_{-\infty}^{\infty} |d_a(\tau)| d\tau \right)^2 s(a) I \underset{a \to \infty}{\to} \beta\alpha I > 0 \quad \blacksquare$$

Next we define what we mean by a spectral line of a vector valued stationary signal. Although impossible to generate in practice, spectral lines yet serve as a good idealization of what one would accomplish in trying to generate a signal whose spectral measure has isolated points of strong support.

Definition 6: A stationary function $x : \mathbf{R} \rightarrow \mathbf{C}^n$ with spectral measure S_x is said to have a *spectral line* at ω_0 if $S_x(\{\omega_0\})$ is positive definite.

Remark:

If x has a spectral line at $\omega_0 \in \mathbf{R}$, then ω_0 is a point of strong support of S_x, because for every neighborhood O of ω_0

$$\inf_{\|v(\omega)\| \geq 1 \,\forall\, \omega \in \mathbf{R}} \int_O v^H(\omega) dS(\omega) v(\omega) \tag{3.28}$$
$$\geq \inf_{\|v(\omega)\| \geq 1 \,\forall\, \omega \in \mathbf{R}} v^H(\omega_0) S(\{\omega_0\}) v(\omega_0) > 0$$

A natural way to generate a signal $u : \mathbf{R} \rightarrow \mathbf{C}^m$, which approximately has a spectral line, at ω_0, say, would be to let $u(t) = \phi(t) e^{j\omega_0 t}$, where the components $\phi_1, \ldots, \phi_m$ of ϕ are square waves whose periods are distinct odd multiples of some strictly positive number.

Finally we establish the relation between the order of excitation of a stationary signal and the number of points of strong support of its spectral measure. As a first step we check the conditions under which the correlation function of the derivative of a signal can be expressed in terms of the derivative of the correlation function of the signal itself.

Proposition 5: Consider two signals $u : \mathbf{R} \rightarrow \mathbf{C}^m$ and $v : \mathbf{R} \rightarrow \mathbf{C}^k$. Assume that

(i) u and v are jointly stationary, i.e. $[u^T \; v^T]^T$ is stationary.

(ii) u and $\dot{v}$ are jointly stationary.

(iii) u is bounded.

(iv) v is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$.

Then

$$R_{u\dot{v}}(\tau) = \frac{d}{d\tau} R_{uv}(\tau) \tag{3.29a}$$

and

$$R_{\dot{v}u}(\tau) = -\frac{d}{d\tau} R_{vu}(\tau) \tag{3.29b}$$

Proof: Let D be the set of discontinuity points of $\dot{v}$. Since $\dot{v}$ is piecewise uniformly continuous, $\exists$ functions $\chi_1, \ldots, \chi_m : \mathbf{R} \to (-|h|, |h|)$ such that for $t \in C(h) := \{t \in \mathbf{R} : t - d > |h| \; \forall d \in D\}$

$$\frac{v_i(t+h) - v_i(t)}{h} = \dot{v}_i(t + \chi_i(t)) \tag{3.30}$$

and

$$s(h) := \sup_{t \in C(h)} \left\| \frac{v(t+h) - v(t)}{h} - \dot{v}(t) \right\| \underset{h \to 0}{\to} 0 \tag{3.31}$$

Since $\dot{v}$ has minimum interdiscontinuity distance $\kappa > 0$, the set $\mathbf{R} \setminus C(h)$ has Lebesgue measure not greater than $2hT / \kappa$, and from *proposition 1* we know that $\dot{v}$ is bounded, say by $M_{\dot{v}} < \infty$. Thus

$$\left\| \frac{v(t+h) - v(t)}{h} - \dot{v}(t) \right\| = \left\| \frac{1}{h} \int_t^{t+h} [\dot{v}(\tau) - \dot{v}(t)] d\tau \right\| \leq 2M_{\dot{v}} \tag{3.32}$$

$\forall t \in \mathbf{R} \setminus C(h)$

Since u is bounded, say by $M_u < \infty$, from (3.31) and (3.32) it then follows that

$$\left\| \frac{R_{uv}(\tau+h) - R_{uv}(\tau)}{h} - R_{u\dot{v}}(\tau) \right\| \tag{3.33}$$

$$= \left\| \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} u(t) \left[\frac{v^H(t+\tau+h) - v^H(t+\tau)}{h} - \dot{v}^H(t+\tau) \right] dt \right\|$$

$$\leq M_u s(h) + \frac{4hM_{\dot{v}}}{\kappa} \underset{h \to 0}{\to} 0$$

which proves (3.29a). Equation (3.29b) then follows, because

$$R_{\dot{v}u}(\tau) = R_{u\dot{v}}^H(-\tau) = \frac{d}{d(-\tau)} R_{uv}^H(-\tau) = -\frac{d}{d\tau} R_{vu}(\tau) \quad \blacksquare \tag{3.34}$$

Proposition 6: Consider a signal u with spectral measure S_u. Assume $u^{(0)}, \ldots, u^{(l-1)}$ are jointly stationary and that $u^{(l-1)}$ is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$. Then

$$\text{cardinality ssupp}(S_u) \geqslant l \implies u \text{ is p.e. of order } l \tag{3.35}$$

Proof: Note that the fact that $u^{(l-1)}$ is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$, implies by *proposition 1* that $u^{(l-1)}$ is bounded. Hence $u^{(l-2)}$ is also (piecewise) uniformly continuous (with minimum interdiscontinuity distance $= \infty$), and thus by induction $u^{(0)}, \ldots, u^{(l-1)}$ are all bounded piecewise uniformly continuous with minimum interdiscontinuity distance $\geqslant \kappa > 0$. From *proposition 5* it then follows by induction that

$$R_{u^{(i)}u^{(j)}}(\tau) = (-1)^i R_u^{(i+j)}(\tau) \qquad i,j \in \{0, \ldots, l-1\} \tag{3.36}$$

Since u and hence R_u are bounded, this has Fourier transform

$$\hat{R}_{u^{(i)}u^{(j)}}(\omega) = (-j\omega)^i (j\omega)^j \hat{R}_u(\omega) \tag{3.37}$$

where the tempered distribution $\hat{R}_u$ is the Fourier transform of R_u. Hence

$$R_{\underline{u}_l}(0) = \begin{bmatrix} R_{u^{(0)}u^{(0)}}(0) & \cdots & R_{u^{(0)}u^{(l-1)}}(0) \\ \vdots & & \vdots \\ R_{u^{(l-1)}u^{(0)}}(0) & \cdots & R_{u^{(l-1)}u^{(l-1)}}(0) \end{bmatrix} \tag{3.38}$$

$$= \int_{\mathbb{R}} \begin{bmatrix} (-j\omega)^0 I_m \\ \vdots \\ (-j\omega)^{l-1} I_m \end{bmatrix} dS_u(\omega) \left[(j\omega)^0 I_m \cdots (j\omega)^{l-1} I_m \right]$$

For an arbitrary vector $v \in \mathbb{C}^{lm} \setminus \{0\}$ let

$$\hat{V}(s) := [s^0 I_m \cdots s^{l-1} I_m] v \in \mathbb{C}^m[s] \tag{3.39}$$

Since the polynomial vector $\hat{V}$ has at most $l-1$ zeros, there is at least one point ω_v of strong support of S_u, such that $\hat{V}(j\omega_v) \neq 0$. Now $\hat{V}$ is continuous, so $\exists$ a neighborhood O_v of ω_v such that

$$\|\hat{V}(j\omega)\| > \frac{\|\hat{V}(j\omega_v)\|}{2} > 0 \qquad \forall\, \omega \in O_v \tag{3.40}$$

and since $\omega_v \in \mathrm{ssupp}(S_u)$,

$$v^H R_{\underline{u}_i}(0)v = \int_{\mathbf{R}} \hat{V}^H(j\omega)dS_u(\omega)\hat{V}(j\omega) \tag{3.41}$$

$$\geq \int_{O_v} \hat{V}^H(j\omega)dS_u(\omega)\hat{V}(j\omega) \geq \frac{\|\hat{V}(j\omega_v)\|^2}{4} \inf_{\|v(\omega)\| \geq 1 \; \forall \; \omega \in \mathbf{R}} \int_{O_v} v^H(\omega)dS_u(\omega)v(\omega) > 0$$

Thus $R_{\underline{u}_i}(0)$ is positive definite. The proposition then follows from *proposition 4*. ∎

4. Output Reachability and Persistency of Excitation

In this section we consider a continuous time *possibly unstable* linear system with a $p \times m$ complex proper rational transfer function $\hat{H}(s) \in \mathbf{C}_p^{p \times m}(s)$. We will with abuse of language but without ambiguity use the symbol $\hat{H}$ to refer to the system itself as well as its transfer function. For the following discussion we assume that

$$\hat{H}(s) = \hat{N}(s)\hat{D}^{-1}(s) \quad \hat{D}(s) \in \mathbf{C}^{m \times m}[s], \; \hat{N}(s) \in \mathbf{C}^{p \times m}[s] \tag{4.1}$$

is a polynomial, column reduced, right coprime matrix fraction description (MFD) of $\hat{H}$, and that

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{4.2a}$$

$$y(t) = Cx(t) + Du(t) \tag{4.2b}$$

with $u(t) \in \mathbf{C}^m$, $x(t) \in \mathbf{C}^n$, $y(t) \in \mathbf{C}^p$ is a minimal state space realization of $\hat{H}$. Thus $\hat{H}$ has McMillan degree $=n$, and Markov parameters

$$(M_0, M_1, \ldots, M_i, \ldots) = (D, CB, \ldots, CA^{i-1}B, \ldots) \tag{4.3}$$

The (minimum) relative degree of $\hat{H}$ and the maximum column degree of $\hat{D}$, we denote by r and μ respectively, i.e.

$$r := \bigwedge_{i,j} \text{rel deg} \hat{H}_{ij} \tag{4.4a}$$

$$\mu := \bigvee_{i,j} \deg \hat{D}_{ij} \tag{4.4b}$$

where $\hat{H}_{ij}$ and $\hat{D}_{ij}$ are the (i,j) entries of the matrices $\hat{H}$ and $\hat{D}$ respectively. Recall that μ is independent of choice of column reduced right coprime MFD $\hat{N}(s)\hat{D}^{-1}(s)$, and that in the usual case of a non-constant $\hat{H}$ (i.e. for $n \neq 0$), μ is the controllability index of any minimal state space realization of $\hat{H}$. More precisely

$$\mu = \min\{\nu \in \mathbf{N} \mid \mathrm{rk}[B\ AB\ \cdots\ A^{\nu-1}B] = n\} \qquad n \in \mathbf{N} \tag{4.5}$$

For any function f taking values in a vector space, we write

$$Z(f) := f^{-1}(\{0\}) \tag{4.6}$$

We then recall that the characteristic polynomial $\hat{\chi}$ of $\hat{H}$ is given by

$$\hat{\chi}(s) = \sum_{i=0}^{n} \chi_i s^i = \det \hat{D}(s) = \det(sI - A) \tag{4.7}$$

so

$$Z(\hat{\chi}) = Z(\det \hat{D}) = \sigma(A) \tag{4.8}$$

and

$$\deg \hat{H} := \deg \hat{\chi} = \deg \det \hat{D} = \dim A = n \tag{4.9}$$

Finally we recall that $\hat{H}$ is said to be *output reachable* (or *output controllable*) if $\forall\, t_0 \in \mathbf{R}$, $\forall\, y \in \mathbf{C}^r$, $\exists t_1 > t_0$ and an input $u : \mathbf{R} \to \mathbf{C}$ such that $x(t_0) = 0 \Longrightarrow y(t_1) = y$.

With this notation we are now ready to state the key result of this paper. It essentially gives frequency domain input conditions under which the output of an output reachable system is persistently exciting.

Theorem 1: Consider the system $\hat{H}$ above with input u and output y. Assume that:

(C1) $\hat{H}$ is output reachable.

(C2) u is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$.

(C3) u is stationary, and the strong support Ω of its spectral measure S_u satisfies at least one of the following:

(i) cardinality $\Omega > n - r$

(ii) cardinality $[\Omega \backslash j\sigma(A)] > \mu - r$

Then y is persistently exciting.

We prove *theorem 1* by building up a few intermediate results, from which the theorem follows as an easy consequence. Until then we assume that the conditions (C1) - (C3) are satisfied. The first of these results gives the link between persistency of excitation of the system input and its autocorrelation and spectral content. It is stated in the following lemma.

Lemma 2: For any compact set K

$$\frac{1}{\Delta}\int_{t_0}^{t_0+\Delta} u(t)u^H(t+\tau)dt \underset{\Delta\to\infty}{\to} R(\tau) \tag{4.10}$$

uniformly in (t_0,τ) on $\mathbf{R}\times K$.

Proof: The lemma follows from the conditions (C2) - (C3) of *theorem 1*. The proof is technically involved and therefore omitted. It is available on request. ∎

We now address the construction of an *initial condition killer*. By this we mean a scalar function, which when convolved with the output y of the system $\hat{H}$, "kills off" the zero input response. This technique of getting around the "unpredictable" interaction between the zero input response and the zero state response of the system, is the trick that makes the proof of *theorem 1* valid regardless of whether $\hat{H}$ is exponentially stable or not. If $\hat{H}(s)$ is exponentially stable this technique is unnecessary.

According to condition (C3) of *theorem 1*, we can pick l distinct frequencies $\omega_1,\ldots,\omega_l \in \mathbf{R}$ such that

$$\omega_1,\ldots,\omega_l \in \Omega \quad l > n-r \tag{4.11a}$$

or

$$\omega_1,\ldots,\omega_l \in \Omega\setminus j\sigma(A) \quad l > \mu-r \tag{4.11b}$$

In either case choose a function $d:\mathbf{R}\to\mathbf{C} \in C^n(\mathbf{R})$ with compact support $\mathrm{supp}(d) = K$ and Fourier transform $\hat{d}:\mathbf{R}\to\mathbf{C}$, such that

$$\hat{d}(-\omega_i) \neq 0 \quad i = 1,\ldots,l \tag{4.12}$$

A simple choice would be:

$$d(t) := \mathop{\ast}_{i=1}^{n+2} a\,\mathrm{rect}(at) \qquad a > \bigvee_{i=1}^{l} \frac{|\omega_i|}{2\pi} \tag{4.13}$$

With $d : \mathbf{R} \to \mathbf{C}$ so chosen, we define a new function $q : \mathbf{R} \to \mathbf{C}$ by

$$q(t) := \hat{\chi}(\frac{d}{dt})d(t) = \sum_{i=0}^{n} \chi_i d^{(i)}(t) \tag{4.14}$$

We then have the following:

Fact 1:

(i) $\hat{d}$ is continuous.

(ii) q is continuous.

(iii) q has compact support $\mathrm{supp}(q) = K$

(iv) $q * e^{A\cdot} = 0$

Proof: Property (i) follows because d is continuous with compact support. Indeed, since the map $K \times \mathbf{R} \to \mathbf{C}$: $(t, \omega) \to e^{-j\omega t}$ is continuous and K is compact, the collection $\{\mathbf{R} \to \mathbf{C}: \omega \to e^{-j\omega t} \mid t \in K\}$ is equicontinuous. Thus

$$|\hat{d}(\omega) - \hat{d}(\omega_0)| \leq \int_K |d(t)|\,|e^{-j\omega t} - e^{-j\omega_0 t}|\,dt \tag{4.15}$$

$$\leq (\sup K - \inf K) \sup_{t \in K} |d(t)| \sup_{t \in K} |e^{-j\omega t} - e^{-j\omega_0 t}| \underset{\omega \to \omega_0}{\to} 0 \qquad \forall\, \omega_0 \in \mathbf{R}$$

Since $d \in C^n(\mathbf{R})$ has compact support, (ii) and (iii) hold. Moreover using integration by parts and the Cayley-Hamilton theorem we have

$$(q * e^{A\cdot})(t) = \sum_{i=0}^{n} \chi_i \int_{-\infty}^{\infty} d^{(i)}(\tau) e^{A(t-\tau)} d\tau \tag{4.16}$$

$$= \sum_{i=0}^{n} \chi_i \int_{-\infty}^{\infty} d(\tau) A^i e^{A(t-\tau)} d\tau = \hat{\chi}(A)(d * e^{A\cdot})(t) = 0$$

which proves (iv). ∎

We call the property (iv) the *killing property* of q. Any continuous function with compact support having this property is, for reasons given above, referred to as an *initial condition killer* of $\hat{H}$ on $\mathrm{supp}(q)$.

Now define a map $X : \mathbf{R} \to \mathbf{C}^{p \times m}$ by

$$X(\sigma) := \int_{-\infty}^{\sigma} q(\tau) C e^{A(\sigma-\tau)} B d\tau + q(\sigma) D \qquad \sigma < 0 \tag{4.17a}$$

$$:= -\int_{\sigma}^{\infty} q(\tau) C e^{A(\sigma-\tau)} B d\tau + q(\sigma) D \qquad \sigma \geqslant 0 \tag{4.17b}$$

and let $\hat{X} : \mathbf{R} \to \mathbf{C}^{p \times m}$ denote its Fourier transform. Some of the properties of these functions are summarized below:

Fact 2:

(i) X has compact support.

(ii) X is continuous.

(iii) $\hat{X}$ is continuous.

(iv) The rows of $\hat{X}$ are linearly independent over $\mathbf{C}$ on $-\Omega$

Proof: From (4.17) we see that $\text{supp}(X)$ is contained in the compact set

$$K_X := [\inf K, 0] \cup [0, \sup K] \tag{4.18}$$

Hence (i) is true. Since q is continuous with compact support, and the map $(\sigma, \tau) \to e^{A(\sigma-\tau)}$ is continuously differentiable, by the bounded convergence theorem $X(\sigma) - q(\sigma)D$ is differentiable and hence continuous when restricted to either $\sigma < 0$ or $\sigma \geqslant 0$. Moreover by the killing property of q

$$\lim_{\sigma \uparrow 0} [X(\sigma) - q(\sigma)D] - [X(0) - q(0)D] \tag{4.19}$$

$$= \lim_{\sigma \uparrow 0} \int_{-\infty}^{\sigma} q(\tau) C e^{A(\sigma-\tau)} B d\tau + \int_{0}^{\infty} q(\tau) C e^{-A\tau} B d\tau$$

$$= C(q * e^{A\cdot})(0)B = 0$$

Thus $X(\sigma) - q(\sigma)D$ is continuous on $\mathbf{R}$, and (ii) follows from the continuity of q. Condition (iii) then follows from (i) and (ii). Finally by changing order of integration (permissible because q is continuous and has compact support), and again using the killing property of q, for $\omega \in \mathbf{R} \setminus -j\sigma(A)$ we obtain

$$\hat{X}(\omega) = \hat{H}(j\omega)\hat{q}(\omega) = \hat{H}(j\omega)\hat{\chi}(j\omega)\hat{d}(\omega) = \hat{N}(j\omega)\,\text{adj}\,\hat{D}(j\omega)\hat{d}(\omega) \tag{4.20}$$

Since $\hat{X}(\omega)$, $\hat{N}(j\omega)$, $\mathrm{adj}\,\hat{D}(j\omega)$ and $\hat{d}(\omega)$ are all continuous functions of ω on $\mathbf{R}$, we conclude that

$$\hat{X}(j\omega) = \hat{N}(j\omega)\,\mathrm{adj}\,\hat{D}(j\omega)\hat{d}(\omega) \qquad \forall\, \omega \in \mathbf{R} \tag{4.21}$$

(even though $\hat{H}(j\omega)$ is not well defined for $\omega \in -j\sigma(A)$). From (4.12) we then have that

$$\begin{aligned} &\mathrm{rk}[\hat{X}(-\omega_1) \cdots \hat{X}(-\omega_l)] \\ &= \mathrm{rk}[\hat{N}(-j\omega_1)\,\mathrm{adj}\,\hat{D}(-j\omega_1) \cdots \hat{N}(-j\omega_l)\,\mathrm{adj}\,\hat{D}(-j\omega_l)] \end{aligned} \tag{4.22}$$

Assume for a moment that (4.11a) is satisfied. Since the system $\hat{H}$ is output reachable, the rows of $\hat{H}$ and hence those of $\hat{N}\ \mathrm{adj}\,\hat{D} = \hat{H}\hat{\chi}$ are linearly independent over $\mathbf{C}$. Thus $\forall\, v \in \mathbf{C}^p\backslash\{0\}$, $\exists\, j \in \{1, \ldots, m\}$ such that $v^H\hat{N}(s)\,\mathrm{adj}\,\hat{D}(s)e_j$ is a *nonzero* polynomial. Since

$$\hat{N}(s)\,\mathrm{adj}\,\hat{D}(s) = \hat{H}(s)\hat{\chi}(s) \qquad \forall\, s \in \mathbf{C}\backslash\sigma(A) \tag{4.23}$$

we have that

$$\deg v^H\hat{N}\ \mathrm{adj}\,\hat{D}e_j \leqslant n - r < l \tag{4.24}$$

Therefore $\forall\, v \in \mathbf{C}^p\backslash\{0\}$, $\exists\, \omega_v \in \{\omega_1, \ldots, \omega_l\}$ such that

$$v^H\hat{N}(-j\omega_v)\,\mathrm{adj}\,\hat{D}(-j\omega_v) \neq 0 \tag{4.25}$$

This shows that

$$[\hat{N}(-j\omega_1)\,\mathrm{adj}\,\hat{D}(-j\omega_1) \cdots \hat{N}(-j\omega_l)\,\mathrm{adj}\,\hat{D}(-j\omega_l)] \tag{4.26}$$

has full rank. From (4.22) it then follows that the rows of $\hat{X}$ are linearly independent over $\mathbf{C}$ on $\{-\omega_1, \ldots, -\omega_l\} \subseteq -\Omega$

If (4.11a) is not satisfied, (4.11b) must be, in which case

$$\det \mathrm{adj}\,\hat{D}(-j\omega_i) \neq 0 \qquad i = 1, \ldots, l \tag{4.27}$$

Hence (4.22) reduces to

$$\mathrm{rk}[\hat{X}(-\omega_1) \cdots \hat{X}(-\omega_k)] = \mathrm{rk}[\hat{N}(-j\omega_1) \cdots \hat{N}(-j\omega_k)] \tag{4.28}$$

Since the rows of $\hat{H} = \hat{N}\hat{D}^{-1}$ are linearly independent over $\mathbf{C}$, so are those of $\hat{N}$. Thus $\forall\, v \in \mathbf{C}^p\backslash\{0\}$, $\exists\, j \in \{1, \ldots, m\}$ such that $v^H\hat{N}(s)e_j$ is a *nonzero* polynomial. Since $\hat{N}(s) = \hat{H}(s)\hat{D}(s)$ we have that

$$\deg v^H \hat{N} e_j \leqslant \mu - r < l \tag{4.29}$$

By analogy with the previous case, from (4.28) it therefore follows that the rows of $\hat{X}$ are linearly independent over $\mathbb{C}$ on $-\Omega$ This completes the proof of (iv). ∎

Remarks:

1) Only (iv), which follows from (C3), depends on the conditions of *theorem 1.*

2) The jump from $n-r+1$ to $\mu-r+1$ in the required number of points of strong support of the input spectrum taking place if these points are restricted to be outside the set $j\sigma(A)$ may seem a bit strange. There however exist output reachable multi input systems with $\mu < n$ and finite sets $\Omega = \{\omega_1, \ldots, \omega_n\}$ such that $\hat{d}(-\omega_1) = \cdots = \hat{d}(-\omega_n) = 1$, but $[\hat{X}(-\omega_1) \cdots \hat{X}(-\omega_n)]$ does not have full rank. Consider for example the minimal realization

$$A = \operatorname{diag}(-j\omega_1, \ldots, -j\omega_{n-1}, 1 - j\omega_n) \tag{4.30a}$$

$$B = C = D = I_n \tag{4.30b}$$

Then

$$e_n^H [\hat{X}(-\omega_1) \cdots \hat{X}(-\omega_n)] = 0 \quad \blacksquare \tag{4.31}$$

Next let η denote the function $\eta := q * y : \mathbf{R} \to \mathbb{C}^p$. We then have the following.

Fact 3: $\exists$constants $\Delta < \infty$ and $\alpha > 0$ such that

$$\int_{t_0}^{t_0+\Delta} \eta(t)\eta^H(t)dt \geqslant \alpha\Delta I \qquad \forall t_0 \in \mathbf{R} \tag{4.32}$$

i.e. η is persistently exciting.

Proof: Using the killing property of q, and changing order of integration, (permissible since q is continuous and has compact support,) it is straight forward to show that

$$\eta = X * u \tag{4.33}$$

Then by another straight forward calculation

$$\frac{1}{\Delta}\int_{t_0}^{t_0+\Delta}\eta(t)\eta^H(t)dt = \int_{K_X}\int_{K_X} X(\sigma)\frac{1}{\Delta}\int_{t_0}^{t_0+\Delta}u(t-\sigma)u^H(t-\tau)dt\ X^H(\tau)d\sigma d\tau \quad (4.34)$$

$$\geq \int_{K_X}\int_{K_X} X(\sigma)R_u(\sigma-\tau)X^H(\tau)d\sigma d\tau$$

$$-\int_{K_X}\int_{K_X} \| R_u(\sigma-\tau) - \frac{1}{\Delta}\int_{t_0}^{t_0+\Delta}u(t-\sigma)u^H(t-\tau)dt \| X(\sigma)X^H(\tau)d\sigma d\tau$$

Since $X: \mathbf{R} \to \mathbf{C}^{p\times m}$ is bounded, *lemma 2* implies that the last integral above tends to 0 uniformly in t_0 on $\mathbf{R}$ as $\Delta \to \infty$. Therefore using the convolution theorem we have

$$\frac{1}{\Delta}\int_{t_0}^{t_0+\Delta}\eta(t)\eta^H(t)dt \underset{\Delta\to\infty}{\to} \int_{K_X}\int_{K_X} X(\sigma)R_u(\sigma-\tau)X^H(\tau)d\sigma d\tau \quad (4.35)$$

$$= \int_{-\infty}^{\infty}\hat{X}(-\omega)dS_u(\omega)\hat{X}^H(-\omega)$$

uniformly in t_0 on $\mathbf{R}$. Let $v \in \mathbf{C}^p\setminus\{0\}$. Since the rows of $\hat{X}$ are linearly independent over $\mathbf{C}$ on $-\Omega$, $\exists\omega_v \in \Omega$ such that $v^H\hat{X}(-\omega_v) \neq 0$. Since moreover $\hat{X}$ is continuous, $\exists$ a neighborhood O_v of ω_v such that

$$\|\hat{X}^H(-\omega)v\| \geq \frac{\|\hat{X}^H(-\omega_v)v\|}{2} > 0 \quad \forall\,\omega \in O_v \quad (4.36)$$

Since ω_v is a point of strong support of S_u, it follows, just as in *proposition 6* that

$$v^H\int_{-\infty}^{\infty}\hat{X}(-\omega)dS_u(\omega)\hat{X}^H(-\omega)\,v > 0 \quad (4.37)$$

This shows that the right hand side of (4.35) is positive definite. The fact then follows from the uniform convergence of the left hand side in the same equation. ∎

From *proposition 3* and *fact 3* it now follows that y is persistently exciting. This completes the proof of *theorem 1*.

Remarks:

1) For single input systems $\hat{D}$ is scalar. Hence

$$\mu = \deg\hat{D} = \deg\det\hat{D} = n \quad (4.38)$$

which means that the two subconditions (i) and (ii) of (C3) in *theorem 1* are equivalent.

2) For multi input systems neither of these two conditions implies the other. Consider for example a system with minimal state space realization

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & j\omega_0 & 0 \\ 0 & 0 & -j\omega_0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \quad C = I_3 \quad D = 0 \quad \omega_0 \neq 0 \tag{4.39}$$

Then

$\Omega = \{0, -\omega_0, \omega_0\} \Longrightarrow$ (i) is satisfied while (ii) is not satisfied.

whereas for $\omega_1 \notin \{0, -\omega_0, \omega_0\}$

$\Omega = \{-\omega_1, \omega_1\} \Longrightarrow$ (ii) is satisfied while (i) is not satisfied.

3) The output reachability condition (C1) is necessary, to guarantee that y is persistently exciting for all initial conditions. This is obvious since the zero state response takes values only in the space $R\,[D\ CB\ CAB\ \cdots CA^{n-1}B]$, which is equal to $\mathbb{C}^p$ iff $\hat{H}$ is output reachable.

4) The input conditions (C2) - (C3) are not necessary, as can easily be verified by simple first order SISO examples such as

$$A = B = C = 0, \quad D = 1, \quad u(t) = t \tag{4.40a}$$

$$A = B = C = 0, \quad D = 1, \quad u(t) = \text{random telegraph signal} \tag{4.40b}$$

$$A = B = C = 1, \quad D = 0, \quad u(t) = \mathrm{rect}(t - \frac{1}{2}) \tag{4.40c}$$

5) Note that it is not true in general, that the required number of points of strong support in condition (C3) can be reduced, if the system possesses unstable modes. For example in the case of a strictly proper first order system, only $n - r + 1 = 1$ point of strong support is required. However, for every input u with compact support T, $\exists$ an initial state

$$x(0) = -\int_T Ce^{-A\tau}Bu(\tau)d\,\tau \tag{4.41}$$

such that the output y is not persistently exciting, even if the single mode of the system is unstable. Thus an attempt to compensate for a point of strong support of the input spectral measure, by exciting an unstable mode with an input of finite duration might fail.

6) Note that because $\hat{X}^H(\omega)v$ varies over O_v, the *strong* support assumption is necessary in the proof above.

The rest of this section is devoted to another version of *theorem 1* for which the input conditions are expressed in the time domain rather than the frequency domain. It does neither follow from nor imply, *theorem 1*, and the proof is considerably different. The two theorems are however closely related. We begin the proof of the theorem with the following lemma, which may be a useful for other purposes as well.

Lemma 3: Consider the system $\hat{H}$ above with input u and output y. Assume that:

(C1) $\hat{H}$ is output reachable.

(C2) u is locally integrable, e.g. piecewise uniformly continuous.

(C3) $\exists$ a function $d : \mathbf{R} \to \mathbf{C} \in C^n(\mathbf{R})$ with compact support, such that $d * u$ is persistently exciting of order $n - r + 1$.

Then y is persistently exciting.

Proof: Let

$$q(t) := \hat{\chi}(\frac{d}{dt})d(t) = \sum_{i=0}^{n} \chi_i d^{(i)}(t) \tag{4.42}$$

As was shown in the proof of *fact 1*, the conditions on d ensure that q is an initial killer of $\hat{H}$. Let furthermore

$$v := d * u \tag{4.43}$$

$$\eta := q * y \tag{4.44}$$

Then

$$\eta(t) = \int_{-\infty}^{\infty} q(\tau) C e^{-A\tau} x(t) d\tau \tag{4.45}$$

$$+ \int_{-\infty}^{\infty} q(\tau) \int_{t}^{t-\tau} C e^{A(t-\tau-\sigma)} B u(\sigma) d\sigma d\tau + \int_{-\infty}^{\infty} q(\tau) D u(t-\tau) d\tau$$

Due to the killing property of q, the first term on the RHS of (4.45) vanishes $\forall\, t \in \mathbf{R}$. Changing order of integration (permissible since the integrand is continuous and has compact support), the second term on the RHS of (4.45) can be written

$$\int_0^{\infty} C \sum_{i=0}^{n} \chi_i \int_{-\infty}^{-\sigma} d^{(i)}(\tau) e^{A(-\tau-\sigma)} d\tau\, Bu(t+\sigma) d\sigma \tag{4.46}$$
$$- \int_{-\infty}^{0} C \sum_{i=0}^{n} \chi_i \int_{-\sigma}^{\infty} d^{(i)}(\tau) e^{A(-\tau-\sigma)} d\tau\, Bu(t+\sigma) d\sigma$$

Since d is smooth with compact support, using integration by parts (4.46) can be expressed as

$$\int_0^{\infty} C \sum_{i=0}^{n} \chi_i \,[\sum_{j=0}^{i-1} A^j d^{(i-1-j)}(-\sigma) + A^i \int_{-\infty}^{-\sigma} d(\tau) e^{A(-\tau-\sigma)} d\tau\,] Bu(t+\sigma) d\sigma \tag{4.47}$$
$$- \int_{-\infty}^{0} C \sum_{i=0}^{n} \chi_i \,[- \sum_{j=0}^{i-1} A^j d^{(i-1-j)}(-\sigma) + A^i \int_{-\sigma}^{\infty} d(\tau) e^{A(-\tau-\sigma)} d\tau\,] Bu(t+\sigma) d\sigma$$

By the Cayley-Hamilton theorem and (4.3) this reduces to:

$$\int_{-\infty}^{\infty} \sum_{i=0}^{n} \chi_i \sum_{j=0}^{i-1} M_{j+1} d^{(i-1-j)}(-\sigma) u(t+\sigma) d\sigma \tag{4.48}$$

Observing that the last term of the RHS of (4.45) can be written as

$$\int_{-\infty}^{\infty} \sum_{i=0}^{n} \chi_i M_0 d^{(i)}(\tau) u(t-\tau) d\tau \tag{4.49}$$

we therefore have

$$\eta(t) = \int_{-\infty}^{\infty} \sum_{i=0}^{n} \chi_i \sum_{j=0}^{i} M_j\, d^{(i-j)}(\tau) u(t-\tau) d\tau \tag{4.50}$$

Since d is smooth with compact support and u is locally integrable, the differentiation can be moved outside the integral sign in (4.50). Changing summation index and recalling that $M_0 = \cdots = M_{r-1} = 0$, we thus obtain

$$\eta(t) = \sum_{j=0}^{n} \sum_{i=j}^{n} \chi_i M_{i-j} (d * u)^{(j)}(t) \tag{4.51}$$
$$= [M_r \cdots M_n] \begin{bmatrix} \chi_r I_m & \cdots & \chi_n I_m \\ \vdots & ⋰ & \\ \chi_n I_m & & \bigcirc \end{bmatrix} \underline{v}_{n-r+1}(t)$$

Since the system is output reachable and the leading characteristic polynomial coefficient

$\chi_n = 1$, the matrix

$$M := [M_r \cdots M_n] \begin{bmatrix} \chi_r I_m & \cdots & \chi_n I_m \\ \vdots & \cdot & \\ \vdots & \cdot & \bigcirc \\ \chi_n I_m & & \end{bmatrix} \tag{4.52}$$

has full rank. Since moreover v is persistently exciting of order $n-r+1$, $\exists \Delta < \infty$ and $\alpha > 0$ such that

$$\int_{t_0}^{t_0+\Delta} \eta(t)\eta^H(t)dt \geqslant \lambda_{\min}(MM^H)\alpha I > 0 \quad \forall t_0 \in \mathbf{R} \tag{4.53}$$

Since q is continuous with compact support, the rest of the lemma follows from *proposition 3.* ∎

Theorem 2: Consider the system $\hat{H}$ above with input u and output y. Assume that:

(C1) $\hat{H}$ is output reachable.

(C2) u is bounded and $u^{(n-r)}$ is piecewise uniformly continuous with minimum inter-discontinuity distance $\kappa > 0$.

(C3) u is persistently exciting of order $n-r+1$.

Then y is persistently exciting.

Proof: Let

$$d_a(t) := \mathop{\ast}_{i=1}^{n+2} a \operatorname{rect}(at) \qquad a > 0 \tag{4.54}$$

Then

$$\operatorname{supp}(d_a) = [-\frac{n+2}{2a}, \frac{n+2}{2a}] \tag{4.55}$$

and

$$\int_{-\infty}^{\infty} d_a(\tau)d\tau = 1 > 0 \qquad \forall a > 0 \tag{4.56}$$

so by *lemma 1* $\exists \bar{a} < \infty$ such that $d_{\bar{a}} * u$ is persistently exciting of order $n-r+1$. Moreover $d_{\bar{a}} \in C^n(\mathbf{R})$, and by *proposition 2* u is piecewise uniformly continuous and hence locally integrable. The theorem therefore follows from *lemma 3*. ∎

Remark:

Most of *theorem 1* could have been deduced by choosing a smooth function $d : \mathbf{R} \to \mathbf{C}$ with compact support, and whose Fourier transform is nonzero at at least $n-r+1$ of the points of strong support of the spectral measure of the input signal, for example

$$d(t) := \mathop{\ast}_{i=1}^{n+2} a \operatorname{rect}(at) \qquad a > 0 \tag{4.57}$$

for a large enough. Let $v := d * u$. Using *lemma 2* it can be shown that $v^{(0)}, \ldots, v^{(n-r)}$ are jointly stationary, so by *proposition 6* v is persistently exciting of order $n-r+1$. It then follows by *lemma 3* that y is persistently exciting. One reason for choosing the other approach is that it brings out the role played by the controllability index μ.

5. State Space Realizations and Persistency of Excitation

The theorems in the previous section relates persistency of excitation of the output of a linear system to the spectral content and the time domain behavior respectively of its input. For a given (not necessarily minimal) state space realization (A,B,C,D) there are of course similar relations between the input and the state and between the state and the output. In the discussion of these relations below we will use the same notation as in the previous section, but we relax the over all assumption, that A, B, C, D refer to a *minimal* realization. The following input/state-relation is a simple consequence of *theorem 1*

Corollary 1: If (A, B) is controllable and the input conditions (C2) - (C3) of *theorem 1* or *theorem 2* are satisfied, then the state trajectory x in (4.2) is persistently exciting.

Proof: The assumption that (A, B) is controllable is equivalent to the assumption that the system with input u, output x and transfer function $(sI - A)^{-1}B$ is output reachable. It therefore follows from one of the two theorems, that x is persistently exciting. ∎

Some state/output results are summarized below. Let

$$\underline{M}_{l,j} := [0 \cdots 0\ M_0 \cdots M_{l-1-j}] \in \mathbb{C}^{lp \times m} \tag{5.1}$$

where the j leading zeros are $p \times m$ matrices, and $\{M_j\}_{j=0}^{\infty}$ are the Markov parameters of the system under consideration. We then have the following test for persistency of excitation of the state trajectory in terms of excitation properties of the input and output functions.

Theorem 3: Assume that (C, A) is observable with observability index ν. If the input u is $\nu-1$ times differentiable and the vector valued function

$$z(t) := \underline{y}_\nu(t) - \sum_{j=0}^{\nu-1} \underline{M}_{\nu,j} u^{(j)}(t) \tag{5.2}$$

is persistently exciting, then the state trajectory $x : \mathbb{R} \to \mathbb{C}^n$ is also persistently exciting.

Proof: Note that for $l < \nu$ we have

$$y^{(l)}(t) = CA^l x(t) + \sum_{j=0}^{l} M_{l-j} u^{(j)}(t) \tag{5.3}$$

Therefore by direct calculation

$$z(t) = \begin{bmatrix} y^{(0)}(t) - \sum_{j=0}^{0} M_{0-j} u^{(j)}(t) \\ \cdot \\ \cdot \\ y^{(\nu-1)}(t) - \sum_{j=0}^{\nu-1} M_{\nu-1-j} u^{(j)}(t) \end{bmatrix} = \begin{bmatrix} CA^0 x(t) \\ \cdot \\ \cdot \\ CA^{\nu-1} x(t) \end{bmatrix} = O_\nu x(t) \tag{5.4}$$

where the $\nu p \times n$ matrix

$$O_\nu := \begin{bmatrix} CA^0 \\ \cdot \\ \cdot \\ CA^{\nu-1} \end{bmatrix} \tag{5.5}$$

has full column rank by the definition of ν. Hence

$$\int_{t_0}^{t_0+\Delta} x(t)x^H(t)dt \geqslant \lambda_{\min}\left[\int_{t_0}^{t_0+\Delta} z(t)z^H(t)dt\right]\lambda_{\max}(O_\nu^H O_\nu) \qquad (5.6)$$

$$\forall\, t_0 \in \mathbf{R}, \quad \forall\, \Delta \in (0,\infty)$$

from which the theorem follows. ∎

Remarks:

1) The Markov parameters can, in principle, be derived from any input-output description of the system or from simple experiments. Thus, in principle, the test does not require knowledge of the order of the system, let alone a state space parametrization. The computation of $z(t)$ does however require reliable differentiation of the input, output and system impulse response signals. To make practical use of the test one would have to rely on *proposition 3* and *lemma 1*.

2) If we replace all derivatives in (5.2) by forward shifts, we obtain the corresponding result for discrete time systems. In this case the Markov parameters are just the impulse response of the system. Hence the i th component of $z(t)$ is given by

$$e_i^H z(t) = y(t+i-1) - \sum_{j=0}^{i-1} M_{i-1-j}\, u(t+j) = y(t+i-1) - \rho[i-1, u(\cdot+t)] \qquad (5.7)$$

where $\rho(t, u)$ denotes the zero state response at time t to the input u. We see that z can readily be computed from the input- and output signals (without any unreliable operations such as differentiation). If, furthermore, the system is stable, $\rho[i-1, u(\cdot+t)]$ $\quad i = 1, \ldots, \nu$ can be obtained from the system itself without knowledge of the Markov parameters. This method can obviously be used in practice. In the special case when $u = 0$, this modified test (5.2) reduces to a previously known result [4].

For unforced systems *theorem 3* implies the following.

Corollary 2: Assume that (C, A) is observable with observability index ν. If the input $u(t) \equiv 0$ and the output y is persistently exciting of order ν, then the state trajectory $x : \mathbf{R} \rightarrow \mathbf{C}^n$ is persistently exciting.

This fact can also be expressed in terms of the spectral measure of the input:

Corollary 3: Assume that (C, A) is observable with observability index ν. If the input $u(t) \equiv 0$, the output has jointly stationary derivatives of all orders less than ν, and the output spectral measure S_y has at least ν points of support, then the state trajectory $x : \mathbf{R} \rightarrow \mathbf{C}^n$ is persistently exciting.

Proof: From *proposition 6* we know that $\underline{y}_\nu$ is persistently exciting. Since $u(t) \equiv 0$ we have that

$$\underline{y}_\nu(t) = O_\nu x(t) \tag{5.8}$$

where O_ν defined as above has full column rank $= n$. Hence the transfer function $\hat{J}$ from $\underline{y}_\nu$ to x is given by

$$\hat{J}(s) \equiv (O_\nu^H O_\nu)^{-1} O_\nu^H \tag{5.9}$$

Since this $n \times \nu p$ matrix has full row rank, $\hat{J}$ represents an output reachable system. The corollary then follows from *theorem 1*. ∎

Finally we note some facts which are true for strictly proper systems only:

Proposition 7: For a strictly proper system ($D = 0$) the following is true.

(i) $\operatorname{rk} C < p \Rightarrow y$ is *not* p.e. $\Rightarrow \operatorname{supp}(S_y) = \varnothing$

(ii) If $\operatorname{rk} C = p$, then

$\operatorname{supp}(S_x) \neq \varnothing \Rightarrow x$ is p.e. $\Rightarrow y$ is p.e.

(iii) If $p = n$ and C is invertible, i.e. the observability index $\nu = 1$, then

$\mathrm{supp}(S_y) \neq \emptyset \Rightarrow y$ is p.e. $\Rightarrow x$ is p.e.

Proof: Immediate from *proposition 4*, *theorem 1* and the fact that

$$\int_{t_0}^{t_0+\Delta} y(t) y^H(t) dt = C \int_{t_0}^{t_0+\Delta} x(t) x^H(t) dt \; C^H \qquad \forall t_0 \in \mathbf{R} \quad \blacksquare \tag{5.10}$$

Remark:

1) Note that the conclusion in (iii) is true under the weaker condition that $\mathrm{rk}\, C = n$. If $\mathrm{rk}\, C < p$, however, it is impossible for y to be persistently exciting.

6. Application to Adaptive Identification

To illustrate the importance of the results in the previous sections, we will consider an adaptive identification scheme for proper, possibly unstable, plants, and prove that it ensures parameter convergence. We thereby extend the applicability of adaptive identification techniques to the class of unstable plants. This is obviously of interest for the purpose of identifying unstable plants. But more importantly it relaxes the requirements of a priori knowledge about whether the plant to be identified is stable or unstable. After all the parameters of the plant to be identified are unknown (Why else identify?), so that the stability properties of the plant are not necessarily available.

In the first few subsections of this section we give a detailed convergence proof for identification of SISO plants. We then present the natural extension of this result to identification of multi input single output (MISO) plants, along with the parts of the proof that differ from the SISO case. Finally we outline how these results can be used for identification of multi input multi output (MIMO) plants.

6.1. General Assumptions

Consider a plant with input $u(t) \in \mathbf{C}$, output $y(t) \in \mathbf{C}$ and proper rational transfer function

$$\hat{P}(s) = \frac{\hat{n}(s)}{\hat{d}(s)} \tag{6.1}$$

where $\hat{n}$ and $\hat{d}$ are coprime polynomials and $\hat{d}$ is monic of known degree $= n$. We assume that the observable modes of the plant are not both unstable and uncontrollable. The structure of the adaptive identifier is shown in fig. 1.

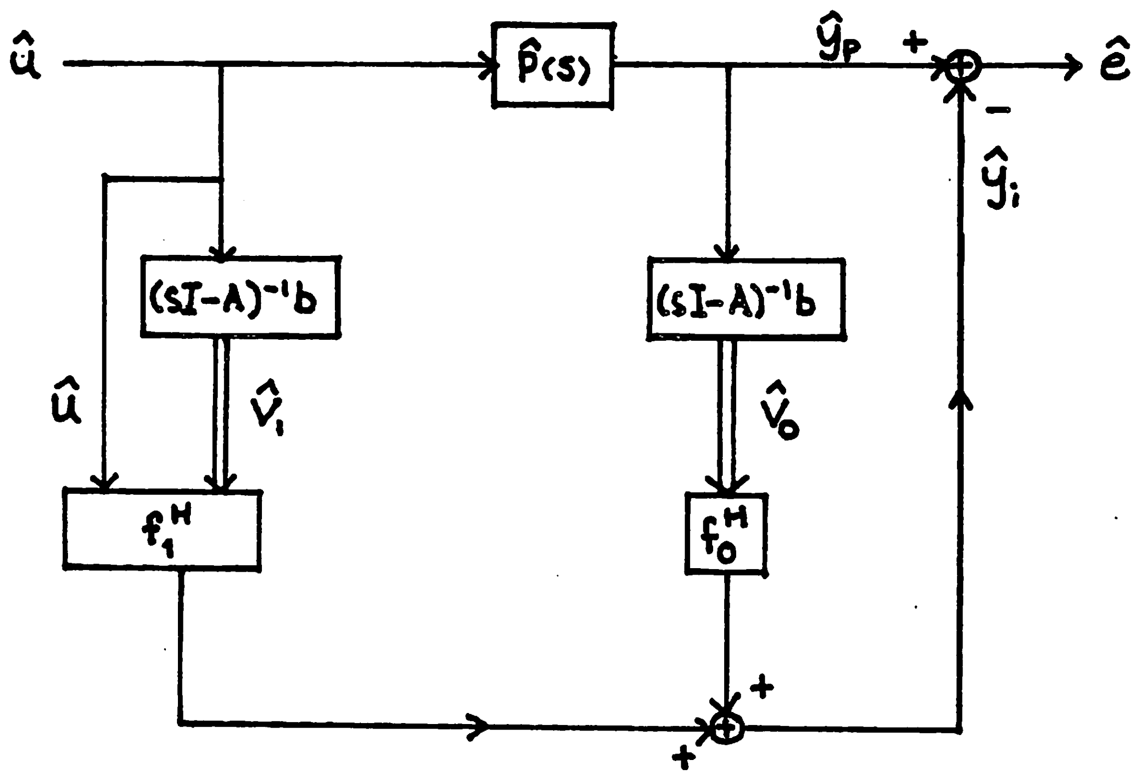

Figure 1

where $A \in \mathbb{C}^{n \times n}$, $b, f_0(t) \in \mathbb{C}^n$, $f_1(t) \in \mathbb{C}^{n+1}$. Here and throughout the rest of this paper it is understood that matrices and vectors of zero dimension represent non-existing signal paths, states etc. With this interpretation the block diagram of the identifier in fig. 1 makes sense even when $n = 0$. In this case the blocks corresponding to A, b and f_0 do not exist, so $y_i(t) = f_1^*(t)u(t)$, $f_1(t) \in \mathbb{C}$. We make the following assumptions about the adaptive identifier design and the input signal.

(A1) (A, b) is controllable.

(A2) $\sigma(A) \in \mathbf{C}_-^c$

(I1) u is piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$.

(I2) u is stationary, and its spectral measure S_u has at least $2n + 1$ points of strong [1] support.

We will w.l.o.g. assume that (A, b) is on canonical controllable form. It can then readily be verified, that under the assumptions (A1) and (A2) there exists a *unique* parametrization $g = [g_0^T \; g_1^T]^T$, $g_0 \in \mathbf{C}^n$, $g_1 \in \mathbf{C}^{n+1}$ of the plant $\hat{P}$ on the form given in fig. 2.

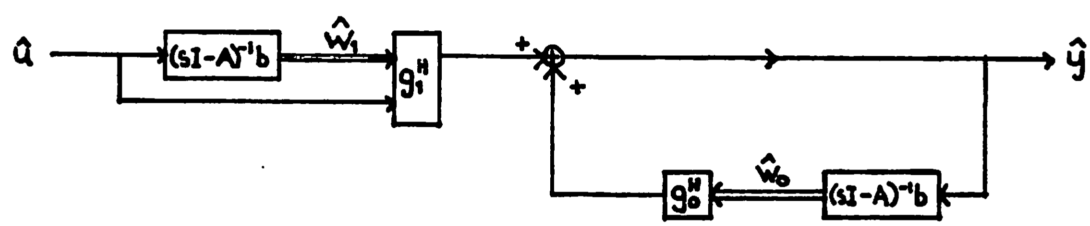

Figure 2

6.2. Notation

We introduce the following notation for the analysis of the adaptive identifier.

Identifier state: $$v(t) := \begin{bmatrix} v_0(t) \\ v_1(t) \\ u(t) \end{bmatrix}$$

Plant state: $$w(t) := \begin{bmatrix} w_0(t) \\ w_1(t) \\ u(t) \end{bmatrix}$$

1) For single input plants the attribute "strong" in condition (I2) has no significance. When referring to (I2) in the multi input case it is important however that "strong" not be omitted.

State error: $\epsilon(t) := \begin{bmatrix} \epsilon_0(t) \\ \epsilon_1(t) \\ 0 \end{bmatrix} := \begin{bmatrix} w_0(t) \\ w_1(t) \\ u(t) \end{bmatrix} - \begin{bmatrix} v_0(t) \\ v_1(t) \\ u(t) \end{bmatrix}$

Estimated parameters: $f(t) := \begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix}$

True parameters: $g := \begin{bmatrix} g_0 \\ g_1 \end{bmatrix}$

Parameter error: $\phi(t) := \begin{bmatrix} \phi_0(t) \\ \phi_1(t) \end{bmatrix} := \begin{bmatrix} g_0 \\ g_1 \end{bmatrix} - \begin{bmatrix} f_0(t) \\ f_1(t) \end{bmatrix}$

Identifier output: $y_i(t)$

Plant output: $y_p(t)$

Output error: $e(t) := y_p(t) - y_i(t) = g^H w(t) - f^H(t) v(t)$

$$= \phi^H(t) v(t) + g^H \epsilon(t)$$

Note that the identifier- and plant "states" are not states strictly speaking in that they include the input as well.

6.3. Update Law

For the convergence analysis we assume that the identifier is updated according to the "least squares with covariance resetting" update law:

$$P(t) = \beta I > 0 \quad \forall\, t \in R_T := \{nT\}_{n=0}^{\infty} \tag{6.2a}$$
$$\dot{P}(t) = -P(t) v(t) v^H(t) P(t) \quad \forall\, t \in R_T^c := [0,\infty) \setminus R_T \tag{6.2b}$$
$$\dot{f}(t) = P(t) v(t) e^H(t) \quad \forall\, t \geq 0 \tag{6.2c}$$

for some $T \in (0,\infty]$. Note that the "ordinary least squares" update law *without* covariance resetting is included as the special case for which $T = \infty$. In any case it follows immediately that

$$P^{-1}(0) > 0 \tag{6.3a}$$
$$\frac{d}{dt} P^{-1}(t) = v(t) v^H(t) \quad \forall\, t \in R_T^c \tag{6.3b}$$
$$\dot{\phi}(t) = -\dot{f}(t) = -P(t) v(t) e^*(t) \quad \forall\, t \geq 0 \tag{6.3c}$$

6.4. Convergence Analysis

From fig. 1 we see, that, for $n \geqslant 0$, the transfer function $\hat{Q}(s)$ from the input $\hat{u}(s)$ to the identifier state $\hat{v}(s)$ is given by

$$\hat{Q}(s) := \begin{bmatrix} (sI - A)^{-1}b\hat{P}(s) \\ (sI - A)^{-1}b \\ 1 \end{bmatrix} = \begin{bmatrix} s^0\hat{P}(s) \\ \cdot \\ \cdot \\ s^{n-1}\hat{P}(s) \\ s^0 \\ \cdot \\ \cdot \\ s^{n-1} \\ \hat{\chi}(s) \end{bmatrix} \frac{1}{\hat{\chi}(s)} = \begin{bmatrix} s^0\hat{n}(s) \\ \cdot \\ \cdot \\ s^{n-1}\hat{n}(s) \\ s^0\hat{d}(s) \\ \cdot \\ \cdot \\ s^{n-1}\hat{d}(s) \\ \hat{\chi}(s)\hat{d}(s) \end{bmatrix} \frac{1}{\hat{\chi}(s)\hat{d}(s)} \tag{6.4}$$

Since $\deg\hat{\chi} + \deg\hat{d} > n-1 + \deg\hat{d} \geqslant n-1 + \deg\hat{n}$, the transfer function $\hat{Q}(s)$ is strictly proper. Moreover if $[\alpha_0^H \ \alpha_1^H]\hat{Q}(s) \equiv 0$, where $\alpha_0 \in \mathbf{C}^n$, $\alpha_1 \in \mathbf{C}^{n+1}$, then

$$\hat{\alpha}_0(s)\hat{n}(s) \equiv \hat{\alpha}_1(s)\hat{d}(s) \tag{6.5}$$

where $\hat{\alpha}_0(s)$ and $\hat{\alpha}_1(s)$ are polynomials and $\deg\hat{\alpha}_0 < n = \deg\hat{d}$. If $\hat{\alpha}_0(s) \not\equiv 0$, the zeros of the two sides in (6.5) must coincide, and then at least one of the zeros of $\hat{d}$ must also be a zero of $\hat{n}$. This contradicts the assumption that $\hat{d}$ and $\hat{n}$ are coprime. Therefore $\hat{\alpha}_0(s) \equiv 0$ and hence $\hat{\alpha}_1(s) \equiv 0$, which shows that the rows of $\hat{Q}$ are linearly independent over $\mathbf{C}$. It follows that any system with transfer function $\hat{Q}(s)$ is output reachable. Under the input assumptions (I1) - (I2) this means that the hypotheses of *theorem 1* are satisfied. Thus the identifier state v is persistently exciting, i.e. $\exists\, \Delta < \infty$ and $\alpha > 0$ such that

$$\int_{t_0}^{t_0+\Delta} v(\tau)v^H(\tau)d\tau \geqslant \alpha I \qquad \forall\, t_0 \in \mathbf{R} \tag{6.6}$$

From (6.3) it then follows that

$$P^{-1}(t) = P^{-1}(0) + \int_0^t v(\tau)v^H(\tau)d\tau \geqslant \left\lfloor \frac{t}{\Delta} \right\rfloor \alpha I \qquad \forall\, t \geqslant 0 \tag{6.7}$$

Next from fig. 1 and fig. 2 we see that

$$\dot{v}_0(t) = Av_0(t) + by_p(t) \tag{6.8a}$$

$$\dot{v}_1(t) = Av_1(t) + bu(t) \tag{6.8b}$$

$$\dot{w}_0(t) = Aw_0(t) + by_p(t) \tag{6.9a}$$

$$\dot{w}_1(t) = Aw_1(t) + bu(t) \tag{6.9b}$$

Hence

$$\dot{\epsilon}_0(t) = \dot{w}_0(t) - \dot{v}_0(t) = A\,\epsilon_0(t) \tag{6.10a}$$

$$\dot{\epsilon}_1(t) = \dot{w}_1(t) - \dot{v}_1(t) = A\,\epsilon_1(t) \tag{6.10b}$$

Since $\sigma(A) \subseteq \mathbb{C}_-^o$, this implies that $\epsilon(t) \underset{t\to\infty}{\to} 0$ exponentially, i.e. $\exists M < \infty$ and $\lambda > 0$ such that

$$\|\epsilon(t)\| \leqslant Me^{-\lambda t} \qquad \forall\, t \geqslant 0 \tag{6.11}$$

We now introduce the function

$$V(t) := \phi^H(t)P^{-1}(t)\phi(t) - \int_0^t |g^H \epsilon(\tau)|^2 d\tau \tag{6.12}$$

This is almost a Lyapunov function. The only difference is the presence of the second term, which contributes a time varying but bounded (and therefore, for our purposes, harmless shift). From (6.2) and the expression for the output error in section 6.2 we see that $V(t)$ is differentiable on R_T^c and that

$$\dot{V}(t) = -\,|\phi^H(t)v(t) + g^H\epsilon(t)|^2 = -\,|e(t)|^2 \leqslant 0 \qquad \forall\, t \in R_T^c \tag{6.13}$$

Thus

$$V(t) \leqslant V(nT) \qquad \forall\, t \in [nT, nT+T) \qquad \forall\, n \in \mathbb{N}_0 := \mathbb{N}\cup\{0\} \tag{6.14}$$

From (6.1), (6.7), (6.11) and (6.12) it then follows that

$$\beta\,\|\phi(nT)\|^2 + \frac{\|g\|^2 M^2}{2\lambda}\, e^{-2\lambda nT} \geqslant \beta\,\|\phi(nT)\|^2 + \int_{nT}^t |g^H\epsilon(\tau)|^2 d\tau \tag{6.15}$$

$$= V(nT) + \int_0^t |g^H\epsilon(\tau)|^2 d\tau \geqslant \phi^H(t)P^{-1}(t)\phi(t)$$

$$= \phi^H(t)\left[P^{-1}(nT) + \int_{nT}^t v(\tau)v^H(\tau)d\tau\right]\phi(t) \geqslant \left(\beta + \left\lfloor\frac{t-nT}{\Delta}\right\rfloor\alpha\right)\|\phi(t)\|^2$$

$$\forall\, t \in [nT, nT+T) \qquad \forall\, n \in \mathbb{N}_0$$

Hence

$$\|\phi(t)\|^2 \leqslant \frac{\beta\,\|\phi(nT)\|^2 + Le^{-2\lambda nT}}{\beta + \left\lfloor\frac{t-nT}{\Delta}\right\rfloor\alpha} \qquad \forall\, t \in [nT, nT+T), \quad \forall\, n \in \mathbb{N}_0 \tag{6.16}$$

where

$$L = \frac{\|g\|^2 M^2}{2\lambda} \tag{6.17}$$

For the update law *without* covariance resetting ($T = \infty$) with $n = 0$ (6.16) yields:

$$\|\phi(t)\|^2 \leqslant \frac{\beta \|\phi(0)\|^2 + L}{\beta + \left\lfloor \frac{t}{\Delta} \right\rfloor \alpha} \underset{t \to \infty}{\longrightarrow} 0 \tag{6.18}$$

i.e. parameter convergence with rate $1 / \sqrt{t}$. For an update law *with* covariance resetting ($T < \infty$) (6.16) yields:

$$\|\phi((n+1)T)\|^2 \leqslant a \|\phi(nT)\|^2 + b e^{-2\lambda n T} \qquad \forall\, n \in \mathbb{N}_0 \tag{6.19}$$

where

$$a = \frac{\beta}{\beta + \left\lfloor \frac{T}{\Delta} \right\rfloor \alpha} \tag{6.20a}$$

$$b = \frac{L}{\beta + \left\lfloor \frac{T}{\Delta} \right\rfloor \alpha} \tag{6.20b}$$

Hence

$$\|\phi(nT)\|^2 \leqslant a^n \|\phi(0)\|^2 + b \, \frac{a^n - e^{-2\lambda T n}}{a - e^{-2\lambda T}} \qquad a \neq e^{-2\lambda T} \tag{6.21a}$$

$$\|\phi(nT)\|^2 \leqslant a^n \|\phi(0)\|^2 + n a^{n-1} b \qquad a = e^{-2\lambda T} \tag{6.21b}$$

In any case for a resetting period $T \geqslant \Delta$, (6.20a) shows that

$$\|\phi(nT)\| \underset{n \to \infty}{\longrightarrow} 0 \tag{6.22}$$

exponentially. This shows that exponential parameter convergence can be obtained by sampling the estimated parameter vector $f(t)$ at resetting times $t \in R_T$ only. To relax this, we see from (6.16) and (6.22) that

$$\sup_{t \in [nT, nT+T)} \|\phi(t)\|^2 \leqslant \|\phi(nT)\|^2 + \frac{L}{\beta} e^{-2\lambda n T} \underset{n \to \infty}{\longrightarrow} 0 \tag{6.23}$$

exponentially.

If the identifier dynamics is fast enough, more precisely if

$$\bigvee_{s \in Z(\dot{d})} \operatorname{Re} s + \bigvee_{s \in \sigma(A)} \operatorname{Re} s < 0 \tag{6.24}$$

the convergence rate indicated by (6.18) for the update law *without* covariance resetting can be improved to $1 / t$, a fact known to be true for discrete time systems [8]. Indeed from (6.3) we have that

$$P^{-1}(t)\phi(t) = -\nu(t)[\nu^H(t)\phi(t) + \epsilon^H(t)g] \qquad (6.25)$$
$$= -\frac{d}{dt}P^{-1}(t)\phi(t) - \nu(t)\epsilon^H(t)g$$

Hence

$$\frac{d}{dt}[P^{-1}(t)\phi(t)] = -\nu(t)\epsilon^H(t)g \qquad (6.26)$$

so

$$P^{-1}(t)\phi(t) = P^{-1}(0)\phi(0) - \int_0^t \nu(\tau)\epsilon^H(\tau)g\,d\tau \qquad (6.27)$$

and thus

$$\|P^{-1}(t)\phi(t)\| \leqslant \gamma := \frac{1}{\beta}\|\phi(0)\| + \int_0^\infty \|\nu(\tau)\|\,\|\epsilon(\tau)\|\,d\tau\,\|g\| \qquad (6.28)$$

If $\gamma < \infty$, then

$$\|\phi(t)\| \leqslant \|P(t)\|\,\|P^{-1}(t)\phi(t)\| \leqslant \frac{\gamma}{\lambda_{\min}P^{-1}(t)} \leqslant \frac{\gamma}{\left\lfloor \frac{t}{\Delta} \right\rfloor \alpha} \underset{t\to\infty}{\to} 0 \qquad (6.29)$$

Now from fig. 2 and (6.8) we have

$$\begin{bmatrix} \dot{w}_0(t) \\ \dot{w}_1(t) \end{bmatrix} = \bar{A}\begin{bmatrix} w_0(t) \\ w_1(t) \end{bmatrix} + \bar{b}u(t) \qquad (6.30)$$

where

$$\bar{A} = \begin{bmatrix} A + bg_0^H & * \\ 0 & A \end{bmatrix} \qquad (6.31)$$

with spectrum

$$\sigma(\bar{A}) = \sigma(A + bg_0^H) \cup \sigma(A) = Z(\hat{d}) \cup \sigma(A) \qquad (6.32)$$

Since u is stationary and piecewise uniformly continuous with minimum interdiscontinuity distance $\kappa > 0$, by *proposition 1* it is also bounded. Together with (6.30) this implies that $\forall\, \lambda_w > \bigvee_{s \in \sigma(\bar{A})} \mathrm{Re}\, s$, $\exists M_{w0}, M_{w1} < \infty$ such that $\|w(t)\| \leqslant M_{w0} + M_{w1}e^{\lambda_w t}$ $\forall\, t \geqslant 0$. Likewise since $\sigma(A) \subseteq \mathbb{C}_-^\circ$, from (6.10) we know that $\forall\, \lambda_\epsilon > \bigvee_{s \in \sigma(A)} \mathrm{Re}\, s$, $\exists$ $M_\epsilon < \infty$ such that $\|\epsilon(t)\| \leqslant M_\epsilon e^{\lambda_\epsilon t}$ $\forall\, t \geqslant 0$. We therefore see that if (6.24) holds, then

$\exists \lambda_\epsilon < 0$ and $\lambda_w < -\lambda_\epsilon$ such that

$$\int_0^\infty \|v(\tau)\| \|\epsilon(\tau)\| d\tau \leqslant \int_0^\infty (\|w(\tau)\| + \|\epsilon(\tau)\|) \|\epsilon(\tau)\| d\tau \tag{6.33}$$

$$\leqslant \int_0^\infty (M_{w0} + M_{w1} e^{\lambda_w t}) M_\epsilon e^{\lambda_\epsilon t} dt + \int_0^\infty M_\epsilon^2 e^{2\lambda_\epsilon t} dt < \infty$$

in which case $\gamma < \infty$ and (6.29) holds.

6.5. Extension to Multi Input Single Output Plants

The adaptive identifier in the previous subsections can readily be extended to identification of MIMO plants. We do this in two steps. In this subsection we discuss the MISO case. The extension to multi-output plants is even simpler and treated in the next subsection.

Consider a plant with input $u(t) \in \mathbb{C}^m$, output $y(t) \in \mathbb{C}$ and proper rational transfer function

$$\hat{P}(s) = [\hat{P}_1(s) \cdots \hat{P}_m(s)] \tag{6.34}$$

and left coprime polynomial MFD

$$\hat{P}(s) = \hat{D}_L^{-1}(s) \hat{N}_L(s) = \frac{1}{\hat{d}(s)} [\hat{n}_1(s) \cdots \hat{n}_m(s)] \tag{6.35}$$

where $\hat{d}$ is a monic polynomial of known degree n. In analogy with the previous subsections we propose the adaptive identifier structure in fig. 3 below. where $A \in \mathbb{C}^{n \times n}$, b, $f_0(t) \in \mathbb{C}^n$, $f_1(t), \ldots, f_m(t) \in \mathbb{C}^{n+1}$ and (A, b) satisfies the same conditions (A1) - (A2) as in the previous subsection. It is again straight forward to check that there exists a unique parametrization $g := [g_0^T \cdots g_m^T]^T$, $g_0 \in \mathbb{C}^n$, $g_1, \ldots, g_m \in \mathbb{C}^{n+1}$ of the plant $\hat{P}$ given in fig. 4.

If we replace the vectors v, w, ϵ, f, g and ϕ in the notation of the previous subsection by

$$v := [v_0^T \quad v_1^T u_1 \quad v_2^T u_2 \quad \cdots \quad v_m^T u_m]^T \tag{6.36}$$

$$w := [w_0^T \quad w_1^T u_1 \quad w_2^T u_2 \quad \cdots \quad w_m^T u_m]^T \tag{6.37}$$

$$\epsilon := [\epsilon_0^T \quad \epsilon_1^T 0 \quad \epsilon_2^T 0 \quad \cdots \quad \epsilon_m^T 0]^T = w - v \tag{6.38}$$

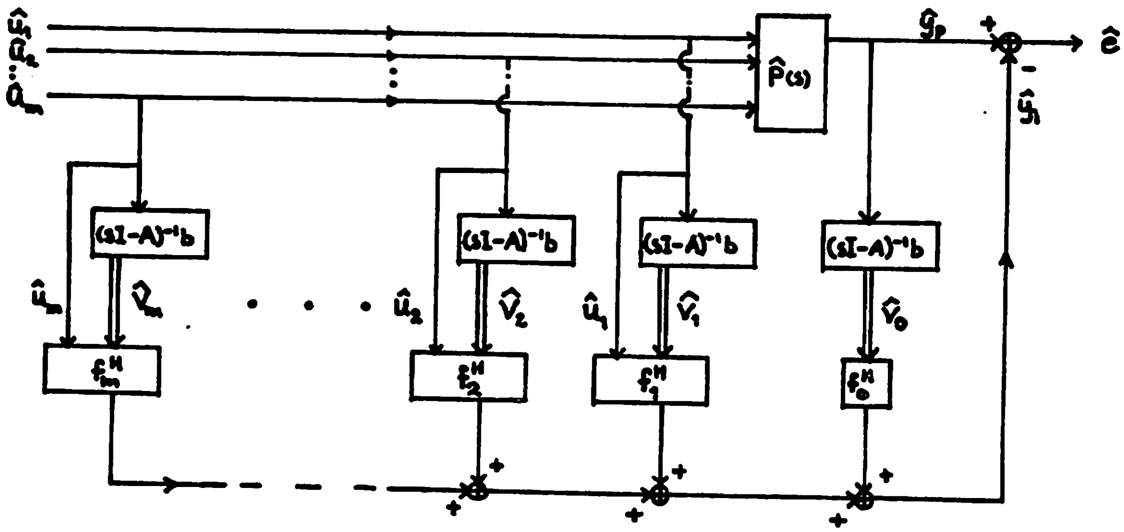

Figure 3

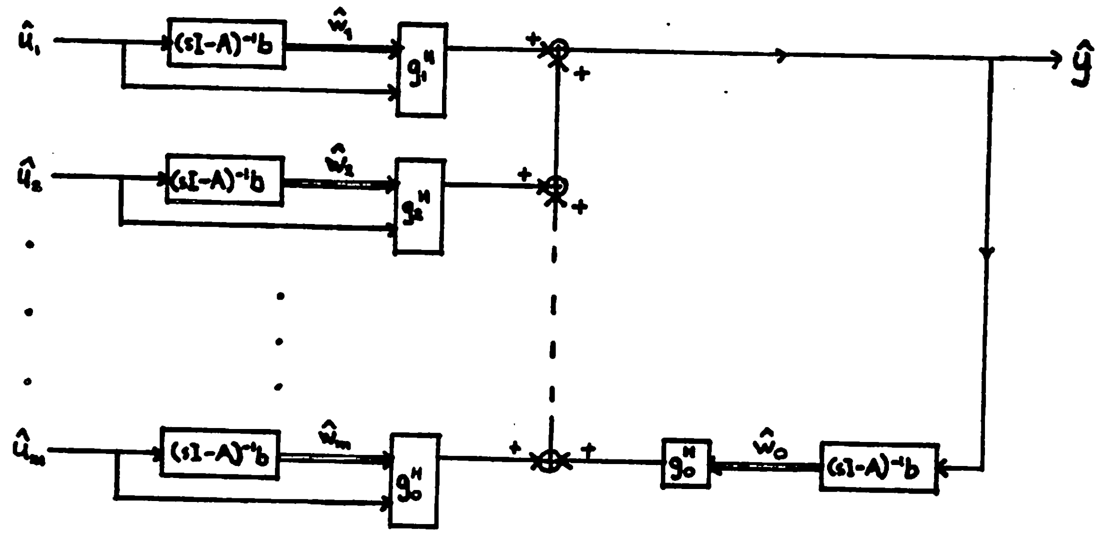

Figure 4

$$f := [f_0^T \cdots f_m^T]^T \tag{6.39}$$

$$g := [g_0^T \cdots g_m^T]^T \tag{6.40}$$

$$\phi := [\phi_0^T \cdots \phi_m^T]^T = g - f \tag{6.41}$$

where all elements except the g_i's are functions of time, and use the same update law (6.2), the convergence analysis that follows is almost identical to that in the foregoing SISO case. The only things we have to check are that the identifier state v is persistently exciting, and that the state error $\epsilon(t) \underset{t\to\infty}{\to} 0$ exponentially fast.

Indeed, from fig. 3 we see that the transfer function $\hat{Q}(s)$ from the input $\hat{u}(s)$ to the identifier state $\hat{v}(s)$ is given by

$$\hat{Q}(s) = \begin{bmatrix} (sI - A)^{-1}b\hat{P}(s) \\ I_m \otimes \begin{bmatrix} (sI - A)^{-1}b \\ 1 \end{bmatrix} \end{bmatrix} \tag{6.42}$$

We note that $\hat{Q}(s)$ is proper. Moreover if the $1 \times m$ matrix $\alpha^H \hat{Q}(s) \equiv 0$ where $\alpha = [\alpha_0^T \cdots \alpha_m^T]^T$, $\alpha_0 \in \mathbb{C}^n$, $\alpha_1, \ldots, \alpha_m \in \mathbb{C}^{n+1}$, then

$$\begin{aligned} 0 \equiv \alpha^H \hat{Q}(s) e_j &= \begin{bmatrix} \alpha_0^H \cdots \alpha_m^H \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}b\hat{P}(s)e_j \\ e_j \otimes \begin{bmatrix} (sI - A)^{-1}b \\ 1 \end{bmatrix} \end{bmatrix} \\ &= \alpha_0^H (sI - A)^{-1} b \hat{P}_j(s) + \alpha_j^H \begin{bmatrix} (sI - A)^{-1}b \\ 1 \end{bmatrix} \\ &= [\hat{\alpha}_0(s)\hat{n}_j(s) + \hat{\alpha}_j(s)\hat{d}(s)] \frac{1}{\hat{\chi}(s)\hat{d}(s)} \qquad j = 1, \ldots, m \end{aligned} \tag{6.43}$$

where $\hat{\alpha}_j$, $j = 0, \ldots, m$ are polynomials and $\deg \hat{\alpha}_0 < n = \deg \hat{d}$. Exactly as in the single input case this implies that

$$\hat{\alpha}_0(s) \equiv \hat{\alpha}_1(s) \equiv \cdots \equiv \hat{\alpha}_m(s) \equiv 0 \tag{6.44}$$

Otherwise

$$Z(\hat{d}) \cap \bigcap_{j=1}^{m} Z(\hat{n}_j) \neq \emptyset \tag{6.45}$$

which would contradict the coprimeness of $\hat{D}_L$ and $\hat{N}_L$. Hence the rows of $\hat{Q}$ are linearly independent over $\mathbb{C}$. Under the input conditions (I1) - (I2) it then follows by *theorem 1*, that the identifier state is persistently exciting.

For the state error dynamics we observe by inspection of fig. 3 and fig. 4 that

$$\dot{v}_0(t) = Av_0(t) + by_p(t) \tag{6.46a}$$

$$\dot{v}_i(t) = Av_0(t) + bu_i(t) \qquad i = 1, \ldots, m \tag{6.46b}$$

$$\dot{w}_0(t) = Aw_0(t) + by_p(t) \tag{6.47a}$$

$$\dot{w}_i(t) = Aw_0(t) + bu_i(t) \quad i = 1, \ldots, m \tag{6.47b}$$

Hence

$$\dot{\epsilon}_i(t) = A\,\epsilon_i(t) \quad i = 0, \ldots, m \tag{6.48}$$

which shows that $\epsilon(t) \underset{t \to \infty}{\to} 0$ exponentially fast. The results for the (SISO) plant therefore extend to the adaptive identifications scheme for (MIMO) plants above.

6.6. Main Result

We summarize the results above in the following theorem:

Theorem 4: Consider the adaptive identifier in fig. 1 or fig. 3 with satisfied conditions (A1) - (A2) along with the update law (6.2). If the input satisfies conditions (I1) - (I2), then the estimated parameters converge to their true values, i.e.

$$f(t) \underset{t \to \infty}{\to} g \tag{6.49}$$

With covariance resetting with long enough period ($T \geqslant \Delta$) the convergence rate is exponential. Without covariance resetting the convergence is at least as fast as $1/\sqrt{t}$. Moreover for any (stable) identifier whose slowest mode is faster than the fastest unstable mode of the plant, i.e. such that

$$\bigvee_{s \in Z(\dot{d})} \operatorname{Re} s + \bigvee_{s \in \sigma(A)} \operatorname{Re} s < 0 \tag{6.50}$$

the convergence is at least as fast as $1/t$.

6.7. Extension to Multi Output Plants

The adaptive identification scheme for single output plants above can be extended to a scheme for multi output plants, by simply connecting one replica of the single output identifier to each component of the plant output. This means that each row of the plant transfer function is identified separately. An obvious practical difficulty arising from this approach is the need to know the degree of each row of the plant transfer function. By the degree of a row, say the ith row, of a transfer function we mean the degree of the transfer function from the input to the ith component of the output vector. This degree is equal to the degree of the smallest common divisor of the elements in the ith row.

7. Conclusion

We have determined input conditions, under which the output of an output reachable *possibly unstable* multivariable continuous time system with proper rational transfer function is persistently exciting. Although not easily stated in one line, these conditions are readily met by an appropriate choice of input signal. These conditions were also found to ensure persistency of excitation of the state trajectory of a controllable state space realization.

For observable state space realizations, tests for persistency of excitation of the state trajectory in terms of the input and the output were developed, and expressed in both time and frequency domain. Some simple relations regarding persistency of excitation of the state and the output of a state space realization of a strictly proper system were given.

Finally we proposed an adaptive identification scheme with least squares update law for proper *possibly unstable* MISO plants. Using the general framework developed in this paper, parameter convergence was proved under certain excitation conditions on the input signal. With covariance resetting, the convergence rate was found to be exponential. Without covariance resetting it was shown to be at least as $1/\sqrt{t}$ for every stable identifier, and as $1/t$ for sufficiently fast identifiers. It was indicated how repeated versions of this MISO scheme can be used to identify MIMO plants as well.

8. Acknowledgements

The author would like to thank E.W. Bai, S. Salcudean, and V. Anantharam for encouragement and helpful discussions.

9. References

[1] S. Boyd and S. Sastry, "On Parameter Convergence in Adaptive Control," *Systems Control Lett.* vol. 3, pp. 311-319, 1983.

[2] S. Boyd and S. Sastry, "Necessary and Sufficient Condition for Parameter Convergence in Adaptive Control," *Proc. American Control Conf.* San Diego, pp. 1584-1588, 1984.

[3] E. Polak, S. Salcudean and D.Q. Mayne "A Sequential Optimal Redesign Procedure for Linear Feedback Systems," *University of California, ERL Memorandum* no. M84/15, Feb., 1984.

[4] E.W. Bai and S.S. Sastry, "Persistency of Excitation, Sufficient Richness and Parameter Convergence in Discrete Time Adaptive Control," *Systems Control Lett.* vol. 6, pp. 153-163, 1985.

[5] M. Green and J.B. Moore, "Persistence of Excitation in Linear Systems," Submitted to *IEEE Trans. Automat. Contr.*

[6] I. Mareels, "Sufficiency of Excitation," *Systems Control Lett.* no. 5, pp. 159-163, 1984.

[7] G. Luder and K.S. Narendra, "An Adaptive Observer and Identifier for a Linear System," *IEEE Trans. Automat. Contr.* pp. 496-499, 1973.

[8] E.W. Bai and S.S. Sastry, "Parameter Identification Using Prior Information," Submitted to *Int. J. Contr.*