

Persistence of excitation in linear systems

Michael GREEN and John B. MOORE

Department of Systems Engineering, Australian National University, G.P.O. Box 4, Canberra 2601, Australia

Received 6 January 1986

Revised 24 March 1986

Abstract: This paper develops output reachability characterizations of linear finite dimensional multivariate systems, so as to translate excitation properties of system inputs to excitation properties of system outputs, states, or associated regression vectors. Such properties are of fundamental concern for convergence of algorithms involving on-line identification, adaptive state estimation, prediction and control. Persistence of excitation guarantees convergence without a priori stability assumptions and ensures robustness properties.

Keywords: Persistence of excitation, Sufficient richness, Adaptive identification, Reachability of regression vectors, Adaptive control.

1. Introduction

Recent work (e.g. [1–5]) has shown the convergence and robustness properties of adaptive estimation and control algorithms to be closely related to persistence of excitation of regression vectors (equivalently that they have a sufficient number of spectral lines [5]). For parameter convergence it is required that signals which are generated inside the time varying feedback loop of the unknown plant (i.e. the regression vectors) be persistently exciting. Since this is not verifiable, it is necessary to study how persistence of excitation transfers from one signal to another through a linear system (for example from the inputs to the regression vectors). This will allow verifiable input persistence of excitation conditions to guarantee the persistence of excitation of the internal regression vector, and hence parameter convergence. In the single input, single output case such input conditions appeared in [6,7,8] where it is required that the regression vector be reachable from the input. The above theory has, however, been confined to single input, single output systems. Multivariable results on translation of persistence of excitation in linear systems appeared in [9] where it is shown that with $[A, B]$ completely reachable, and the input regression vector \bar{u}_i , where $\bar{u}_i' = [u'_{i-1}, u'_{i-2}, \dots, u'_{i-n}]$ persistently exciting, then the state vector x_i will also be persistently exciting. This result holds irrespective of whether the system is stable or unstable. We aim to extend and develop this result so as to guarantee excitation of multivariable regression vectors.

The organisation and contribution of the paper is as follows: In Section 2, the relationship between excitation and reachability is reviewed and generalized, to multivariable and possibly unstable systems, using reachability characterizations in terms of rank conditions. It is shown that reachability is the only requirement on the *system* which is relevant. Section 3, the main contribution of the paper, develops new results on the reachability of regression vectors. It is shown, using our new characterization of reachability, how the reachability of certain regression vectors is related to coprimeness. In so doing we provide compact coordinate free proof of some known results and develop new results. These results are applied to the generation of persistently exciting inputs, and the study of excitation of regression vectors employed in adaptive estimation and control configurations. In particular we develop results on the reachability of pseudo linear regression vectors arising from extended least squares identification.

2. Excitation and reachability

Consider a linear, finite dimensional, discrete time system with state space description

$$x_{k+1} = Ax_k + Bu_k, \quad y_k = Cx_k + Du_k \quad (2.1a)$$

with $x_k \in R^n$, $u_k \in R^m$, $y_k \in R^p$, and associated proper transfer function

$$T(z) = C(zI - A)^{-1}B + D. \quad (2.1b)$$

Definition. The system (2.1) will be called *output reachable* (equivalently y_k is reachable from u_k) if, for any $y \in R^p$ and arbitrary initial state, there exists an input sequence $\{u_i, i = 0, \dots, k < \infty\}$ to the system (2.1) such that its output at time k , y_k , satisfies $y_k = y$.

Let ν be the McMillan degree of the system (1.1) and let $M_0, M_1, \dots, M_\nu, \dots$, be its Markov parameters. Define the matrix M by

$$M = [M_0 \ M_1 \ \dots \ M_\nu] = [D \ CB \ CAB \ \dots \ CA^{\nu-1}B]. \quad (2.2)$$

Recall [10] that M has full row rank p if and only if the system is output reachable. Observe that (2.1) has the following input/output description:

$$N'\tilde{u}_i = [y_i \ y_{i-1} \ \dots \ y_{i-\nu}]d \quad (2.3)$$

where $d_j =$ coefficient of $z^{\nu-j}$ in $d(z)$, a monic minimal polynomial for A , and

$$N_j = \sum_{l=0}^j d_{j-l}M_l, \quad d_0 = 1, \quad (2.4)$$

$$N' = [N_0 \ N_1 \ \dots \ N_\nu], \quad d' = [1 \ d_1 \ \dots \ d_\nu], \quad (2.5a)$$

$$\tilde{u}'_i = [u'_i \ u'_{i-1} \ \dots \ u'_{i-\nu}]. \quad (2.5b)$$

A matrix fraction description of (2.4) is

$$d(z)y(z) = N(z)u(z) \quad (2.6)$$

where

$$d(z) = z^\nu + d_1z^{\nu-1} + \dots + d_\nu, \quad N(z) = N_0z^\nu + N_1z^{\nu-1} + \dots + N_\nu, \quad (2.7a)$$

$$d^{-1}(z)N(z) = T(z). \quad (2.7b)$$

The following lemma is now apparent:

Lemma 2.1. *The following statements are equivalent:*

- The system (2.1) is output reachable, i.e. y_k is reachable from u_k .
- M has full row rank.
- N' has full row rank.
- $N(z)$ has full row rank over R . That is, $\alpha'N(z) = 0$ for all z implies $\alpha = 0$.
- $T(z)$ has full row rank over R .

Excitation and output reachability

Definition. An infinite q -vector sequence $[z_i]$ is said to be *persistently exciting* if

$$\liminf_{l \rightarrow \infty} \frac{1}{l} \left(\sum_{i=1}^l z_i z'_i \right) > 0, \quad (2.8)$$

A signal z_k will be said to be *exciting over an interval* $[k+1, k+l]$ if for some constant $K_1 > 0$,

$$\sum_{i=k+1}^{k+l} z_i z'_i \geq K_1 I. \quad (2.9)$$

Note that if $l < q$, (2.9) cannot hold. A stronger excitation condition, which implies persistence of excitation, is that of *uniform persistence of excitation*. The signal z_k is l -uniformly persistently exciting if there exists $K < \infty$ such that (2.9) holds for all $k \geq K$.

A generalization of a result from [9] is now studied.

Theorem 2.1. *Consider the system (2.1).*

(1) *For the output vector y_i to be exciting over an interval $[k + 1 - \nu, k + l]$, independent of initial conditions, it is necessary and sufficient that the (p dimensional) vector $N' \tilde{u}_i$ (see (2.5)) be exciting over $[k + 1, k + l]$.*

(2) *For the output vector y_i to be persistently exciting, independent of initial conditions, it is sufficient that the (p dimensional) vector $N' \tilde{u}_i$ be persistently exciting.*

Proof. *Sufficiency.* Let $\alpha \in R^p$, $\alpha \neq 0$. Premultiply (2.3) by α' and apply the Cauchy-Schwarz inequality to obtain

$$\|\alpha' N' \tilde{u}_i\|^2 \leq \|\alpha' [y_i \ \cdots \ y_{i-\nu}]\|^2 \|d\|^2.$$

For part (1), sum over the interval $[k + 1, k + l]$ to obtain

$$\alpha' \left[\sum_{i=k+1}^{k+l} N' \tilde{u}_i \tilde{u}_i' N \right] \alpha \leq \|d\|^2 \alpha' \left[\sum_{i=k+1}^{k+l} \sum_{j=i-\nu}^i y_j y_j' \right] \alpha \leq (\nu + 1) \|d\|^2 \alpha' \left[\sum_{i=k+1-\nu}^{k+1} y_i y_i' \right] \alpha. \tag{2.10}$$

The left hand side is strictly positive since $N' \tilde{u}_i$ is exciting over $[k + 1, k + l]$. Also $1 \leq \|d\|^2 < \infty$. Hence y_i is exciting over $[k + 1 - \nu, k + l]$. For part (2), sum from 1 to l , divide by l and take $\lim \inf$ instead of summing over $[k + 1, k + l]$.

Necessity: Suppose $N' \tilde{u}_i$ is not exciting over $[k + 1, k + l]$. Then there exists a non-zero $\alpha \in R^p$ such that $\alpha' N' \tilde{u}_i = 0$, for all $i \in [k + 1, k + l]$. From (2.5a) we have

$$\alpha' y_i + d_1 \alpha' y_{i-1} + \cdots + d_\nu \alpha' y_{i-\nu} = 0, \quad \text{for all } i \in [k + 1, k + l]. \tag{2.11}$$

Now we can choose initial conditions such that $\alpha' y_i = 0$ for $i = k + 1 - \nu, \dots, k + 1$, and the recurrence (2.11) implies $\alpha' y_i = 0$ for all $i \in [k + 1 - \nu, k + l]$. Thus y_i is not exciting over $[k + 1 - \nu, k + l]$. To see that such a choice of initial conditions is possible, consider an observable realization of the transfer function from u to $\alpha' y$. \square

Remarks. Clearly if the system (2.1) is stable, condition (2) of the theorem is also necessary. Without the stability assumption, particular initial conditions, or input excitation over finite time, can lead to persistently exciting outputs without the inputs being persistently exciting.

Lemma 2.2. *A necessary condition for $N' D_i$ to be exciting (over an interval or persistently) is that N' be full row rank, i.e. the system (2.1) is output reachable.*

Remark. It follows that any system which is not output reachable cannot have exciting output.

Corollary 2.1. *A necessary and sufficient condition for the output of every output reachable time invariant linear system (2.1) of McMillan degree ν to be exciting over $[k + 1 - \nu, k + l]$ (resp. persistently exciting) independent of initial conditions is that \tilde{u}_i be exciting over $[k + 1, k + l]$ (resp. persistently exciting).*

Proof. The conditions of Theorem 2.1 must hold for all possible N' . Note that part (2) of Theorem 2.1 is necessary for stable systems. This implies the result. \square

Remarks. 1. It is Corollary 2.1, rather than Theorem 2.1, which is relevant to adaptive estimation. This is because Theorem 2.1, though much less restrictive on the inputs, requires complete and exact knowledge of the system.

2. If $D = 0$, then $N_0 = 0$. Thus we may replace \tilde{u}_i in the theorem with $\bar{u}_i = [u'_{i-1} \ u'_{i-2} \ \dots \ u'_{i-p}]'$. If in addition $C = I$ so that the outputs y_i are the states x_i and the output reachability condition becomes a state reachability condition, then the sufficiency part of Theorem 2.1 specializes to the result of [9].

3. Observe that since the output reachability condition is invariant under constant state variable feedback $u_k = Lx_k + v_k$, then the results of Theorem 2.1 hold for this case, with \tilde{v}_i replacing \tilde{u}_i .

4. Consider the effect of introducing an additive disturbance signal w_k to a plant with output y_k and an external input u_k . For simplicity assume that $d(z)y(z) = N(z)u(z) + N_w(z)w(z)$. Of course, if y_k is reachable from $[u_i \ w_i]$ and $[\tilde{u}'_i \ \tilde{w}'_i]'$ is persistently exciting, then y_i is persistently exciting. This follows from Corollary 2.1. However when u_k, w_k are uncorrelated it is easily shown that the outputs y_i are persistently exciting if either y_i is reachable from u_i and \tilde{u}_i is persistently exciting, or y_i is reachable from w_i and \tilde{w}_i is persistently exciting.

3. Excitation of regression vectors

Corollary 2.1 reduces the question of the transfer of excitation from inputs to outputs to one of reachability of the outputs from the inputs. In adaptive estimation the vectors which one requires to be exciting are regression vectors involving past outputs, inputs and possibly noise estimates (in extended least squares). Thus the focus of this section is on conditions for these regression vectors to be reachable.

Reachability of regression vectors

Here we build on the reachability characterizations of Lemma 2.1 in order to give conditions for the reachability of regression vectors in terms of coprimeness conditions arising from matrix fraction system descriptions.

Let $T(z)$ be a $p \times m$ proper rational matrix with Markov expansion $T(z) = \sum_{i=0}^{\infty} M_i z^{-i}$. Define $T^i(z)$ to be i -th row of $T(z)$, M_j^i the i -th row of M_j and y_k^i the i -th component of y_k . Because of the applications of interest (adaptive identification), we need to consider regression vectors in which each $y_k^i, i = 1, \dots, p$, can have a different number of delays. For example we will not have the same number of delays for inputs and outputs in input/output regression vectors. Thus, for $l_i \geq 1$ arbitrary integers, we define the general regression vector

$$\psi_{l_1 \dots l_p}(k) = \left[y_k^1 \ \dots \ y_{k-l_1+1}^1 \ y_k^2 \ \dots \ y_{k-l_2+1}^2 \ \dots \ y_k^p \ \dots \ y_{k-l_p+1}^p \right]'. \quad (3.1)$$

Now let

$$T_{l_1 \dots l_p}(z) = \left[T^1(z)' \ z^{-1}T^1(z)' \ \dots \ z^{-l_1+1}T^1(z)' \ \dots \ T^p(z)' \ \dots \ z^{-l_p+1}T^p(z)' \right]' \quad (3.2)$$

and

$$M(l_1 \dots l_p) = \begin{bmatrix} M_0^1 & \dots & M_{l_1-1}^1 & M_{l_1}^1 & \dots \\ & & \vdots & \vdots & \\ 0 & & M_0^1 & M_1^1 & \dots \\ M_0^2 & \dots & M_{l_2-1}^2 & M_{l_2}^2 & \dots \\ & & \vdots & \vdots & \\ 0 & & M_0^2 & M_1^2 & \dots \\ \vdots & & \vdots & \vdots & \\ M_0^p & \dots & M_{l_p-1}^p & M_{l_p}^p & \dots \\ & & \vdots & \vdots & \\ 0 & & M_0^p & M_1^p & \dots \end{bmatrix}.$$

Define also

$$P(l_1 \cdots l_p) = \{ \alpha(z) \in P^p, \alpha(z) = [\alpha_1(z) \ \dots \ \alpha_p(z)]', \text{ with } \deg \alpha_i(z) \leq l_i - 1 \} \tag{3.4}$$

and

$$H(z) = \begin{bmatrix} z^{l_1-l_1} & T^1(z) \\ \vdots & \\ z^{l_1-l_p} & T^p(z) \end{bmatrix}, \quad l = \min\{l_i\}. \tag{3.5}$$

Theorem 3.1. *With the definitions (3.1) to (3.5), the following are equivalent:*

- (1) $\psi_{l_1 \dots l_p}(k)$ is reachable from u_k .
- (2) $T_{l_1 \dots l_p}(z)$ is full row rank over R .
- (3) $M(l_1 \cdots l_p)$ is full rank.
- (4) $P(l_1 \cdots l_p) \cap N(H) = 0$, the zero polynomial, and $N(H)$ denotes the left nullspace of $H(z)$.

Proof. Observe $\psi_{l_1 \dots l_p}(z) = T_{l_1 \dots l_p}(z)u(z)$. The matrix of Markov parameters defined in (2.1) for this system is $M(l_1 \cdots l_p)$. Applying Lemma 2.1 gives (1) \Leftrightarrow (2) \Leftrightarrow (3).

We now prove (2) \Leftrightarrow (4). Let $\alpha' = \{\alpha'_1, \dots, \alpha'_p\}$, $\alpha_i \in R^{l_i}$, and

$$\alpha_i(z) = \alpha_{i,1}z^{l_i-1} + \dots + \alpha_{i,l_i}z + \alpha_{i,l_i},$$

so that $\alpha(z) = [\alpha_1(z) \ \dots \ \alpha_p(z)]' \in P(l_1 \cdots l_p)$, arbitrary. Observe now

$$\alpha(z)'H(z) = 0 \quad \Leftrightarrow \quad \alpha'T_{l_1 \dots l_p}(z) = 0.$$

Hence $T_{l_1 \dots l_p}(z)$ full row rank over R is equivalent to $P(l_1 \cdots l_p) \cap N(H) = 0$. \square

Remarks. 1. This new result is the natural extension to output regression vectors of the results of Lemma 2.1 and forms the basis of subsequent derivations where we use condition (4) of the theorem to characterize reachability. Condition (4), although it may appear impractical, combines very nicely with Forney's characterization of minimal polynomial bases for the nullspace of rational matrices [11] to produce regression vector reachability results with comparative ease. In fact the technique quickly becomes routine. The equivalence of conditions (3) and (4) can also produce some interesting results, and in particular provides a method for calculating when condition (4) is satisfied. We now summarize the relevant facts of Forney's theory, which is also discussed in Section 6.5-4 of [12].

Recall the definition of the left minimal index ν_1 of $T(z)$,

$$\nu_1[T(z)] = \min\{\deg \alpha(z) : \alpha(z) \in P^p, \alpha(z)'T(z) = 0\}. \tag{3.6}$$

Let $r = p - \text{rank}(T(z))$ and $\alpha_i(z) \in P^p$, $i = 1, \dots, r$, such that $F(z)T(z) = 0$, where $F(z) = [\alpha_1(z) \ \alpha_2(z) \ \dots \ \alpha_r(z)]'$. The $\alpha_i(z)$, $i = 1, \dots, r$, form a minimal (polynomial) basis for $N(T)$, the left nullspace of $T(z)$, if and only if $F(z)$ is row reduced and irreducible (see Theorem 6.5-10, [12]). The left minimal indices ν_i , $i = 1, \dots, r$, are the degrees of the corresponding $\alpha_k(z)$. This definition of ν_1 agrees with that in (3.1) if we order the ν_i [i.e. shuffle the rows of $F(z)$] so that $\nu_i \leq \nu_{i+1}$, $i = 1, \dots, r-1$. If $T(z)$ has full row rank (over P), then there are no left minimal indices. For convenience, however, we shall say that if $T(z)$ has full row rank, then $\nu_1 = \infty$.

Persistently exciting input signals

Consider input signals u_k derived from white noise w_k (which is persistently exciting) via a linear system. In order for \tilde{u}_k to be persistently exciting, Corollary 2.1 tells us that u_k must be reachable from

w_k . Let us introduce the notations

$$\tilde{u}(z) = \tilde{T}(z)w(z), \tag{3.7a}$$

$$\tilde{T}(z) = [T'(z) \quad z^{-1}T'(z) \quad \dots \quad z^{-l}T'(z)]', \tag{3.7b}$$

where $u(z) = T(z)w(z)$. Relevant reachability results are now derived as a corollary to Theorem 3.1:

Corollary 3.1. *With definitions above,*

$$\tilde{u}_k \text{ is reachable from } w_k \iff \nu_1[T(z)] > l, \tag{3.8}$$

$$\tilde{u}_k \text{ is reachable from } w_k \text{ for any } l \iff T(z) \text{ is right invertible.} \tag{3.9}$$

Proof. By reorganizing rows, observe \tilde{u}_k is reachable if and only if $\psi_{l+1 \dots l+1}(k)$ is reachable (see (3.1)). Hence, applying Theorem 3.1, \tilde{u}_k is reachable if and only if $P(l+1 \dots l+1) \cap N(T) = 0$. Let $\alpha(z)$ be a left minimal (polynomial) vector for $N(T)$ of degree $\nu_1(T)$ (see (3.6)), so $\alpha(z) \in P(\nu_1 + 1 \dots \nu_1 + 1)$ and hence $\nu_1(T) = \text{Max}\{l: P(l \dots l) \cap N(T) = 0\}$. Thus $P(l+1 \dots l+1) \cap N(T) = 0$ if and only if $l < \nu_1(T)$, establishing (3.8). For (3.9), note that $P(l+1 \dots l+1) \cap N(T) = 0$ for arbitrary l if and only if $N(T) = 0$, i.e. $T(z)$ has full row rank or equivalently, is right invertible. \square

Remark. The condition that \tilde{u}_k is reachable for any l is the definition, for the discrete case, of output function reproducibility. Thus we have proved directly the known result [10] that $T(z)$ is output function reproducible if and only if $T(z)$ is right invertible. In fact, Corollary 3.1 can be viewed as stating the extent to which a given system is output function reproducible.

Excitation of input / output regression vectors

Let us consider a regression vector of interest in deterministic adaptive identification, prediction, and control. Thus define

$$\Phi_{\bar{n}\bar{m}}(k) = [y'_{k-1} \quad y'_{k-2} \quad \dots \quad y'_{k-\bar{n}} \quad u'_{k-1} \quad \dots \quad u'_{k-\bar{m}}]'$$
(3.10)

where $y(z) = T(z)u(z)$, $T(z)$ proper, and y_k and u_k are p and m vectors respectively, as in (2.1).

Now consider u and y related by the multivariable ARMA model

$$y_k + A_1 y_{k-1} + \dots + A_{\bar{n}} y_{k-\bar{n}} = B_1 u_{k-1} + \dots + B_{\bar{m}} u_{k-\bar{m}} \tag{3.11}$$

with transfer matrix

$$T(z) = A^{-1}(z)B(z)z^{\bar{n}-\bar{m}} \tag{3.12a}$$

where

$$B(z) = B_1 z^{\bar{m}-1} + \dots + B_{\bar{m}} \quad A(z) = I z^{\bar{n}} + \dots + A_{\bar{n}} \tag{3.12b}$$

and $A_{\bar{n}}, B_{\bar{m}}$ are not zero. Applying Theorem 3.1 to this model, we obtain the following result, which we state as a corollary to Theorem 3.1.

Corollary 3.2. *With the definitions (3.10), (3.11), (3.12) above, $\Phi_{\bar{n}\bar{m}}(k)$ is reachable from u_k if and only if $[A(z) \ B(z)]$ is irreducible (has full row rank for all z) or, equivalently, $A(z), B(z)$ are left coprime.*

Proof. Let $l = \min(\bar{m}, \bar{n})$ and $H(z) = [z^{l-\bar{n}}T(z)' \quad z^{l-\bar{m}}I]'$. By Theorem 3.1 all we need prove is $P(\bar{n} \dots \bar{n} \ \bar{m} \dots \bar{m}) \cap N(H) = 0$ if and only if $A(z), B(z)$ are left coprime.

Observe $[-A(z) \ B(z)]H(z) = 0$. Hence the rows of $[-A(z) \ B(z)] \in N(H)$, and the rows of $[-A(z) \ B(z)] \in P(\bar{n} + 1 \dots \bar{n} + 1 \ \bar{m} \dots \bar{m})$. Thus $A(z), B(z)$ left coprime ($[A(z) \ B(z)]$ irreducible) is a neces-

sary condition for $P(\bar{n} \cdots \bar{n} \bar{m} \cdots \bar{m}) \cap N(H) = 0$. For sufficiency, note that with $A(z)$, $B(z)$ left coprime, the rows of $[A(z) \ B(z)]$ form a basis for $N(H)$ (possibly nonminimal if $\bar{m} > \bar{n}$), and $A(z)$ is row reduced. By the predictable degree property of row reduced matrices (see Theorem 6.3-13 of [12]), $[\alpha(z)' \ \beta(z)'] \in N(H)$, $\alpha(z) \in P^p$, $\beta(z) \in P^m$, implies $\deg \alpha(z) \geq \bar{n}$, so we must have $P(\bar{n} \cdots \bar{n} \bar{m} \cdots \bar{m}) \cap N(H) = 0$. \square

Remarks. 1. This result is known, the multivariable result being in [13], where it is proved using a state space and system equivalence approach, while we have used a purely transfer matrix one resulting in a considerably shorter coordinate basis free proof.

2. In some situations, it would be desirable to have the components of the u and y vectors delayed by varying amounts (e.g. when the transfer matrix (3.12) is diagonal). This situation can clearly be handled by Theorem 3.1, as the generalization from output to input/output regression vectors merely involves augmenting the transfer matrix by an identity matrix.

3. Consider a plant $T(z)$, for example as in (3.12), with the regression vector $\Phi_{\bar{n}\bar{m}}(k)$ reachable from u_k . Now consider the dynamic output feedback law $u(z) = -G(z)y(z) + v(z)$, with $G(z)$ a proper, rational $m \times p$ transfer matrix. Provided only that the transfer matrix from $v(z)$ to $u(z)$, $I_m + G(z)T(z)$, is nonsingular (as a rational matrix), we see that $\Phi_{\bar{n}\bar{m}}(k)$ is reachable from v_k iff $A(z)$, $B(z)$ are left coprime. This follows immediately, since the transfer matrix from $v(z)$ to $[y(z)' \ u(z)']'$ has the same left nullspace as the transfer matrix from $u(z)$ to $[y(z)' \ u(z)']'$.

We conclude that excitation of a regression vector $\Phi_{\bar{n}\bar{m}}(k)$ is invariant under time invariant dynamic feedback controllers $G(z)$.

Excitation of pseudo linear regression vectors

In stochastic adaptive identification, prediction, and control, improved algorithms are obtained by using pseudo linear regression vectors which include noise estimates.

Thus define

$$\Phi_{\bar{n}\bar{m}}(k) = [y'_{k-1} \ \cdots \ y'_{k-\bar{n}} \ \hat{w}'_{k-1} \ \cdots \ \hat{w}'_{k-\bar{l}} \ u'_{k-1} \ \cdots \ u'_{k-\bar{m}}]'$$
 (3.13)

where $y_k \in R^p$, $\hat{w}_k \in R^p$, $u_k \in R^m$. The associated signal models to work with in this case are autoregressive, moving average exogenous input (ARMAX) models with representations

$$y_k + Ay_{k-1} + \cdots + A_{\bar{n}}y_{k-\bar{n}} = B_1u_{k-1} + \cdots + B_{\bar{m}}u_{k-\bar{m}} + w_k + C_1w_{k-1} + \cdots + C_{\bar{l}}w_{k-\bar{l}}.$$
 (3.14)

The noise estimates \hat{w}_k are generated in terms of parameter estimates $\hat{\theta} = \{\hat{A}_i, i = 1, \dots, \bar{n}, \hat{B}_i, i = 1, \dots, \bar{m}, \hat{C}_i, i = 1, \dots, \bar{l}\}$, via

$$\hat{w}_k + \hat{C}_1\hat{w}_{k-1} + \cdots + \hat{C}_{\bar{l}}\hat{w}_{k-\bar{l}} = y_k + \hat{A}_1y_{k-1} + \cdots + \hat{A}_{\bar{n}}y_{k-\bar{n}} - \hat{B}_1u_{k-1} - \cdots - \hat{B}_{\bar{m}}u_{k-\bar{m}}.$$
 (3.15)

In transfer function notation we have

$$\begin{aligned} \begin{bmatrix} y(z) \\ \hat{w}(z) \end{bmatrix} &= \begin{bmatrix} A^{-1}(z)B(z)z^{\bar{n}-\bar{m}} & A^{-1}(z)C(z)z^{\bar{n}-\bar{l}} \\ \hat{C}^{-1}(z)[\hat{A}(z)A^{-1}(z)B(z) - \hat{B}(z)]z^{\bar{l}-\bar{m}} & \hat{C}^{-1}(z)[\hat{A}(z)A^{-1}(z)C(z)] \end{bmatrix} \begin{bmatrix} u(z) \\ w(z) \end{bmatrix} \\ &= \begin{bmatrix} T_{11}(z) & T_{12}(z) \\ T_{21}(z) & T_{22}(z) \end{bmatrix} \begin{bmatrix} u(z) \\ w(z) \end{bmatrix} \end{aligned}$$
 (3.16)

where

$$\begin{aligned} A(z) &= Iz^{\bar{n}} + A_1z^{\bar{n}-1} + \cdots + A_{\bar{n}}, & B(z) &= B_1z^{\bar{m}-1} + \cdots + B_{\bar{m}}, \\ C(z) &= Iz^{\bar{l}} + C_1z^{\bar{l}-1} + \cdots + C_{\bar{l}}, & & \\ \hat{A}(z) &= Iz^{\bar{n}} + \hat{A}_1z^{\bar{n}-1} + \cdots + \hat{A}_{\bar{n}}, & \hat{B}(z) &= \hat{B}_1z^{\bar{m}-1} + \cdots + \hat{B}_{\bar{m}}, \\ \hat{C}(z) &= Iz^{\bar{l}} + \hat{C}_1z^{\bar{l}-1} + \cdots + \hat{C}_{\bar{l}}. & & \end{aligned}$$
 (3.17)

The focus for the remainder of this subsection is to seek conditions on $\hat{\theta}$ (equivalently \hat{A} , \hat{B} , \hat{C}) which ensure the persistence of excitation of $\Phi_{\bar{n}\bar{m}}(k)$ when the inputs are persistently exciting. By Corollary 2.1, we need conditions which ensure $\Phi_{\bar{n}\bar{m}}(k)$ is reachable, which we do by applying Theorem 3.1 as follows.

Corollary 3.3. *With the definitions (3.13), (3.16), (3.17) above:*

- (1) $\Phi_{\bar{n}\bar{m}}(k)$ is reachable from $[u'_k \ w'_k]'$ if and only if $\hat{A}(z)$, $\hat{B}(z)$, $\hat{C}(z)$ are left coprime.
- (2) $\Phi_{\bar{n}\bar{m}}(k)$ is reachable from u_k if and only if both $A(z)$, $B(z)$ and $\hat{C}(z)$, $A(z)\hat{B}(z) - \hat{A}(z)B(z)$ are each left coprime.
- (3) $\Phi_{\bar{n}\bar{m}}(k)$ is not reachable from w_k for any $\hat{\theta}$ selection.
- (4) $\Phi_{\bar{n}\bar{m}}(k)$ of (3.10) is reachable from $[u'_k \ w'_k]'$ for all $\hat{\theta}$ selections.

Proof. (1) Let

$$H(z) = \begin{bmatrix} z^{l-\bar{n}}T_{11}(z) & z^{l-\bar{n}}T_{12}(z) \\ z^{l-\bar{l}}T_{21}(z) & z^{l-\bar{l}}T_{22}(z) \\ z^{l-\bar{m}}I_m & 0 \end{bmatrix}, \quad l = \min(\bar{l}, \bar{m}, \bar{n}).$$

By Theorem 3.1, all we need prove is $P(\bar{n} \cdots \bar{n} \bar{l} \cdots \bar{l} \bar{m} \cdots \bar{m}) \cap N(H) = 0 \Leftrightarrow \hat{A}(z)$, $\hat{B}(z)$, $\hat{C}(z)$ left coprime. Observe that the rows of $[-\hat{A}, \hat{C}, \hat{B}] \in N(H)$ and the rows of $[-\hat{A}, \hat{C}, \hat{B}] \in P(\bar{n} + 1 \cdots \bar{n} + 1 \bar{l} + 1 \cdots \bar{l} + 1 \bar{m} \cdots \bar{m})$. Proof now proceeds as per the proof of Corollary 3.3.

(2) Let

$$H(z) = \begin{bmatrix} z^{l-\bar{n}}T_{11}(z) \\ z^{l-\bar{l}}T_{21}(z) \\ z^{l-\bar{m}}I_m \end{bmatrix}, \quad l = \min(\bar{n}, \bar{l}, \bar{m}).$$

By Theorem 3.1, all we need prove is $P(\bar{n} \cdots \bar{n} \bar{l} \cdots \bar{l} \bar{m} \cdots \bar{m}) \cap N(H) = 0$ if and only if A , B and \hat{C} , $A\hat{B} - \hat{A}B$ are left coprime. Observe that

$$\begin{bmatrix} -\hat{A}(z) & \hat{C}(z) & \hat{B}(z) \\ -A(z) & 0 & B(z) \end{bmatrix} H(z) = 0.$$

As before, it is necessary that $\begin{bmatrix} -\hat{A} & \hat{C} & \hat{B} \\ A & 0 & B \end{bmatrix}$ be irreducible, or, equivalently, \hat{C} , $A\hat{B} - \hat{A}B$ and A , B be left coprime. For sufficiency, with A , B and \hat{C} , $A\hat{B} - \hat{A}B$ left coprime, note that since $\hat{C}(z)$ is row reduced, $[\alpha(z), \beta(z), \gamma(z)] \in N(H)$, $\alpha(z) \in P^p$, $\beta(z) \in P^p$, $\gamma(z) \in P^m \Rightarrow \beta(z) = 0$ or $\deg \beta(z) \geq \bar{l}$. If $\deg \beta(z) \geq \bar{l}$, then $P(\bar{n} \cdots \bar{n} \bar{l} \cdots \bar{l} \bar{m} \cdots \bar{m}) \cap N(H) = 0$. If $\beta(z) = 0$, we must have $[\alpha(z), \gamma(z)] \in \{\text{space spanned by rows of } [-A(z), B(z)]\} \Rightarrow \deg \alpha(z) \geq \bar{n}$ as $A(z)$ is row reduced. Hence $P(\bar{n} \cdots \bar{n} \bar{l} \cdots \bar{l} \bar{m} \cdots \bar{m}) \cap N(H) = 0$.

(3) Let

$$H(z) = \begin{bmatrix} z^{l-\bar{n}}T_{21}(z) \\ z^{l-\bar{l}}T_{22}(z) \\ 0 \end{bmatrix}, \quad l = \min(\bar{n}, \bar{l}, \bar{m}).$$

Observe

$$\begin{bmatrix} -\hat{A} & \hat{C} & F(z) \\ 0 & 0 & E(z) \end{bmatrix} H(z) = 0,$$

where $F(z)$, $E(z)$ are arbitrary. Hence $P(\bar{n} \cdots \bar{n} \bar{l} \cdots \bar{l} \bar{m} \cdots \bar{m}) \cap N(H) \neq 0$ for any \bar{n} , \bar{l} , \bar{m} , and any $\hat{\theta}$.

(4) Let

$$H(z) = \begin{bmatrix} z^{l-\bar{n}}T_{11}(z) & z^{l-\bar{n}}T_{12}(z) \\ z^{l-\bar{m}}I_m & 0 \end{bmatrix}, \quad l = \min(\bar{m}, \bar{n}).$$

Observe that $H(z)$ has full row rank so that $N(H) = 0$, and $\Phi_{\bar{n}\bar{m}}(k)$ is reachable from $[u'_k \ w'_k]'$ for all $\hat{\theta}$.
□

Remarks. 1. Observe that vector $[y'_{k-1}, \dots, y'_{k-\bar{n}}, \hat{w}'_{k-1}, \dots, \hat{w}'_{k-i}]'$ will be reachable from w_k if and only if $\hat{A}(z), \hat{C}(z)$ are left coprime by virtue of Corollary 3.2.

2. By part (3) of Corollary 3.3, persistently exciting noise w_k will not ensure the pseudo linear regression vector $\Phi_{\bar{n}\bar{m}}(k)$ to be persistently exciting. This is because w_k cannot effect u_k . Suppose, however, that u_k is determined from y_k by dynamic output feedback, i.e. $u(z) = E^{-1}(z)F(z)z^{-q}$, where $E(z) = Iz^r + E_1z^{r-1} + \dots + E_r$ and $F(z) = F_1z^{q-1} + \dots + F_q$, E_r, E_q not zero. Necessary and sufficient conditions for $\Phi_{\bar{n}\bar{m}}(k)$ to be reachable from w_k can be established using the techniques developed above, and are as follows:

$\hat{A}(z), \hat{B}(z), \hat{C}(z)$ are left coprime and, $g_i(z) \cap P(\bar{n} \dots \bar{n} \bar{m} \dots \bar{m}) = 0$ for $i = 1, \dots, m$, where $\{g_i(z) \in P^{p+m}, i = 1, \dots, m\}$ is a minimal polynomial basis for the rational vector space spanned by the rows of $[F(z) \ E(z)z^\kappa]$, where $\kappa = \bar{m} - \bar{n} + q - r$.

There are three cases, depending on the controller order (i.e. on q and r), which are profitable to consider, as they simplify the above condition on $E(z), F(z)$.

(i) If $q < \bar{n}$ and $r < \bar{m}$, then $\Phi_{\bar{n}\bar{m}}(k)$ is not reachable from w_k .

(ii) If $q \leq \bar{n}$ and $r \leq \bar{m}$, then the $E(z), F(z)$ condition is equivalent to $E(z), F(z)$ left coprime and $q = \bar{n}$ or $r = \bar{m}$.

(iii) If q and r are unconstrained, no simplification is possible.

Notice that *the order of the controller is crucial in deciding reachability*. In particular, low order controllers (i.e. case (i) above) mean that persistently exciting system noise (e.g. white noise) is not enough to provide persistence of excitation of the pseudo linear regression vector.

3. To ensure reachability of $\Phi_{\bar{n}\bar{m}}(k)$ from system inputs for all $\hat{\theta}$ selections, additional inputs could be used, as suggested in the modified adaptive schemes of [14], where dither signals are added at the output. It is easy to see that in this case the pseudo linear regression vector is reachable for any θ , thus ensuring convergence.

5. Final remarks and conclusions

The theory of this paper exposes the relationship between reachability properties of a linear discrete time system and the translation of excitation properties of system inputs and outputs. The theory is applied to specific signal models and ARMA and ARMAX based adaptive schemes such as are of interest in adaptive identification and control. In particular, a theory for excitation of regression and pseudo regression vectors is developed for polynomial matrix representations. The underlying reachability conditions are expressed as coprimeness conditions on polynomial matrices, confirming some known results with a co-ordinate basis free proof and providing new results.

These reachability results extend to continuous time systems, with delays replaced by differentiation. Unfortunately the basic result relating reachability and excitation, Theorem 2.1, does not easily extend to continuous time, the main problem being to bound the derivatives of $y(t)$ in terms of $y(t)$. However, results analogous to Theorem 2.1 have been obtained for continuous time systems in [7,8].

The results of this paper, in particular Corollary 2.1, can be extended to piece-wise constant and slowly time varying systems [15]. For piece-wise constant systems the persistence of excitation condition is replaced by uniform persistence of excitation. Slowly time varying systems are then handled by approximating them with piece-wise constant systems using an output error criterion, the level of input excitation

required being proportional to the output error bound. For the special case of a constant system with a time varying feedback controller this means that the allowable deviation from a piece-wise constant controller is inversely proportional to the norm of the system state.

The results of this paper are seen to be useful in guiding the design of adaptive schemes so as to ensure persistence of excitation and thereby consistent parameter identification and asymptotic optimality.

References

- [1] B.D.O. Anderson, Exponential stability of linear equations arising in adaptive identification, *IEEE Trans. Automat. Control* **22** (1977) 83–88.
- [2] G. Kreisselmeier, Adaptive observers with exponential rate of convergence, *IEEE Trans. Automat. Control* **22** (1977) 2–9.
- [3] J.S.-C. Yuan and W.M. Wonham, Probing signals for model reference identification, *IEEE Trans. Automat. Control* **22** (1977) 530–538.
- [4] A.P. Morgan and K.S. Narandra, On the uniform asymptotic stability of certain linear non-autonomous differential equations, *SIAM J. Control Optim.* **15** (1977) 5–24.
- [5] S. Boyd and S. Sastry, On parameter convergence in adaptive control, *Systems Control Lett.* **3** (1983) 311–319.
- [6] B.D.O. Anderson and C.R. Johnson Jr., Exponential convergence of adaptive identification and control algorithms, *Automatica* **18** (1982) 1–13.
- [7] S. Dasgupta, B.D.O. Anderson and A.C. Tsoi, Input conditions for continuous-time adaptive systems problems, Tech. Report, Australian National University (1983).
- [8] I. Mareels, Sufficiency of excitation, *Systems Control Lett.* **5** (1985) 159–163.
- [9] J.B. Moore, Persistence of excitation in extended least squares, *IEEE Trans. Automat. Control* **28** (1982) 60–68.
- [10] W.A. Wolovich, *Linear Multivariable Systems* (Springer, Berlin–New York, 1974).
- [11] G.D. Forney, Jr., Minimal bases of rational vector spaces, with application of multivariable linear systems, *SIAM J. Control* **13** (1975) 493–520.
- [12] T. Kailath, *Linear Systems* (Prentice Hall, Englewood Cliffs, NJ, 1980).
- [13] J. Fadavi-Ardekani, S.K. Mitra and B.D.O. Anderson, Extended state-space models of discrete-time dynamical systems, *IEEE Trans. Circuits and Systems* **29** (1982) 547–556.
- [14] J.B. Moore, Dither signals, on-line spectral factorization, and adaptive prewhitening in adaptive control, *Ricerche Automat.*, to appear.
- [15] M. Green and J.B. Moore, Persistence of excitation in linear systems, *Proc. Amer. Control Conf.*, Boston, MA (June 1985) 412–417.