beneficial to apply the proposed algorithm to obtain more accurate results (by the parsimony principle [2]) and more efficient computations than is possible with the usual unconstrained models. In the white noise case, the algorithm becomes a recursive prediction error method, and as such its covariance attains the Cramér–Rao lower bound asymptotically when the true order is used (see, e.g., [3]). In the more general nonwhite noise case, the algorithm's covariance can be evaluated using methods described, e.g., in [2, ch. 7]. This algorithm can be used for filter design and adaptive Nyquist rate estimation. The basic method used here can be applied for deriving system identification algorithms for other constrained transfer functions, such as band-pass and band-stop. Extension to adaptive parameter estimation of constrained ARMA signals with unknown inputs in the presence of noise is presented in [12].

Appendix

The Derivative of Polynomial Coefficients with Respect to Polynomial Zeros

In this Appendix we provide a simple proof of the formula (13). Let C denote the following companion matrix:

$$C=\begin{bmatrix} -a_1 & 1 & & 0\\ \vdots & & \ddots & \\ -a_{n-1} & 0 & & 1\\ -a_n & 0 & \cdots & 0\end{bmatrix}$$

associated with the polynomial $z^nA(z^{-1}) = z^n + a_1z^{n-1} + \cdots + a_n$. It is well known that the zeros $\{\lambda_k\}$ of $z^nA(z^{-1})$ are equal to the eigenvalues of C (see, e.g., [11]). Let

$$u_k=[u_{1,k} \cdots u_{n,k}]^T \neq 0$$

denote a (nonzero) eigenvector corresponding to λ_k. Thus,

$$Cu_k=\lambda_k u_k \tag{A.1}$$

or, in a more detailed form

$$\begin{cases} u_{2,k}=a_1u_{1,k}+\lambda_k u_{1,k}\\ \vdots \qquad \vdots\\ u_{n,k}=a_{n-1}u_{1,k}+\lambda_k u_{n-1,k}\\ 0=a_nu_{1,k}+\lambda_k u_{n,k}. \end{cases} \tag{A.2}$$

It readily follows by contradiction from (A.2) that $u_k \neq 0$ implies $u_{1,k} \neq 0$. Thus, we can set $u_{1,k} = 1$. In the following we assume that the eigenvector u_k has been normalized such that $u_{1,k} = 1$.

Using forward substitution, we find from (A.2) that

$$u_k=Hv_k \tag{A.3}$$

where H is the Hankel matrix

$$H=\begin{bmatrix} & & & 1\\ & & \cdot^{\cdot^{\cdot}} & a_1\\ & \cdot^{\cdot^{\cdot}} & & \vdots\\ 1 & a_1 & \cdots & a_{n-1}\end{bmatrix} \tag{A.4}$$

and

$$v_k=[\lambda_k^{n-1} \cdots \lambda_k 1]^T. \tag{A.5}$$

Next, it can easily be verified that

$$v_k^TC=\lambda_k v_k^T \tag{A.6}$$

which means that v_k is a left eigenvector of C. Left and right eigenvectors associated with different eigenvalues must be orthogonal

$$v_i^Tu_k=0 \qquad \text{for } i\neq k \tag{A.7}$$

which can be seen as follows. From (A.1) and (A.6), we get

$$\lambda_iv_i^Tu_k-\lambda_kv_i^Tu_k=v_i^TCu_k-v_i^TCu_k=0$$

which implies (A.7) since $\lambda_i \neq \lambda_k$. Writing out (A.7) for $i = 1, \cdots, n$, $i \neq k$, we obtain

$$\begin{bmatrix}\lambda_1^{n-1}\\ \vdots\\ \lambda_{k-1}^{n-1}\\ \lambda_{k+1}^{n-1}\\ \vdots\\ \lambda_n^{n-1}\end{bmatrix}+\begin{bmatrix}\lambda_1^{n-2} & \cdots & \lambda_1 & 1\\ \vdots & & \vdots & \vdots\\ \lambda_{k-1}^{n-2} & \cdots & \lambda_{k-1} & 1\\ \lambda_{k+1}^{n-2} & \cdots & \lambda_{k+1} & 1\\ \vdots & & \vdots & \vdots\\ \lambda_n^{n-2} & \cdots & \lambda_n & 1\end{bmatrix}\begin{bmatrix}u_{2,k}\\ \cdot\\ \cdot\\ \cdot\\ \cdot\\ u_{n,k}\end{bmatrix}=0. \tag{A.8}$$

Since $\{\lambda_k\}$ are distinct by assumption, the vander Monde matrix appearing in (A.5) is nonsingular. Therefore, u_k is uniquely determined by $\{\lambda_1, \cdots, \lambda_{k-1}, \lambda_{k+1}, \cdots, \lambda_n\}$ and, in particular, does not depend on λ_k. Using this property we get by differentiating (A.2) with respect to λ_k

$$\partial a/\partial\lambda_k=-u_k. \tag{A.9}$$

From (A.3) and (A.9) the Jacobian matrix $\partial a/\partial\lambda$ is equal to

$$\frac{\partial a}{\partial\lambda}=-[u_1 \cdots u_n]=-HV \tag{A.10}$$

where H was defined in (A.4) and V is the vander Monde matrix $V = [v_1 \cdots v_n]$. Expression (15) is now proven immediately from the i, k entry of the matrices in (A.10). ■

References

[1] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, MA: M.I.T. Press, 1983.

[2] T. Söderström and P. Stoica, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, to be published.

[3] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[4] B. Wahlberg and L. Ljung, "Design variables for bias distribution in transfer function estimation," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 134–144, Feb. 1986.

[5] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[6] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[7] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 716–723, Dec. 1974.

[8] A. Nehorai and B. Porat, "Adaptive comb filtering for harmonic signal enhancement," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1124–1138, Oct. 1986.

[9] A. Nehorai and D. Starer, "An adaptive SSB carrier estimator," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Dallas, TX, Apr. 1987, pp. 2109–2112.

[10] T. Söderström and P. Stoica, "Some properties of the output error method," *Automatica*, vol. 18, pp. 93–99, Jan. 1982.

[11] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[12] A. Nehorai and P. Stoica, "Adaptive constrained ARMA signals in the presence of noise," in *Proc. ICASSP*, Dallas, TX, Apr. 1987, pp. 1003–1006.

Persistency of Excitation Results for Structured Nonminimal Models

MICHAEL HEYMANN

Abstract—**It is frequently convenient to employ specially structured nonminimal models for parameter estimation such as in the case of direct**

Manuscript received February 3, 1987. This work was supported in part by the Technion Fund for promotion of research.

The author is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.

IEEE Log Number 8716170.

adaptive control where the model is parameterized in terms of the desired control law parameters. This is done, for example, in direct model reference adaptive control and in direct pole assignment adaptive control algorithms. It is shown that the parameters appearing in these nonminimal models can be uniquely estimated if and only if a certain *design identity* has a unique solution. The result is used to develop persistency of excitation results for these models.

I. Introduction

The problem of persistency of excitation for parameter convergence in estimation and adaptive control received a great deal of attention in the recent literature (see, e.g., [1]-[9]).

While in many cases minimal models (which have no redundancies through common factors) are used for estimation of parameters, there are situations where the employment of nonminimal models is necessary. This occurs typically in direct adaptive control applications where plant models are specially structured to be parameterized by the adaptive controller parameters. Examples of these instances are direct model reference adaptive control [10], direct pole assignment adaptive control [11], direct model reference adaptive pole assignment [12], as well as in multivariable systems where a simple left MFD is employed having a diagonal "denominator" matrix.

In these nonminimal models the question arises as to whether or not the parameters can be uniquely estimated from the plant input-output data. This often has important implications, e.g., in deciding whether an adaptive design is stable and convergent and in choosing suitable inputs that guarantee the desired stability and convergence.

In the present note we show that the parameters in structured (generally nonminimal) models can be uniquely determined if and only if a certain design identity has a unique solution. The latter is shown to be related to the output-reachability of an *associated-signal system* which in turn is used to develop persistency of excitation results.

The note focuses on single-input single-output systems but similar results can be developed for the multiinput multioutput case.

II. A General Parameterization Problem

We consider discrete time, time-invariant single-input single-output plants of the form

$$p(D)y(t)=r(D)u(t) \tag{2.1}$$

where $p(D)$ and $r(D)$ are real polynomials in the unit delay operator D [i.e., $D^k x(t) = x(t-k)$] of the form

$$p(D)=1+\sum_{i=1}^{n} p_i D^i \tag{2.2}$$

$$r(D)=\sum_{i=1}^{n} r_i D^i. \tag{2.3}$$

We assume that the model (2.1) of the plant is minimal, that is, the polynomials $p(D)$ and $r(D)$ are coprime.

As a key step in setting up the parameter estimation problem we replace the minimal model (2.1) by a structured nonminimal model of the form

$$g(D)y(t)=h(D)u(t) \tag{2.4}$$

where the polynomials $g(D)$ and $h(D)$ are assumed to be parameterized as follows:

$$g(D)=c(D)+\sum_{i=1}^{m_a} \alpha_i a_i(D) \tag{2.5}$$

$$h(D)=d(D)+\sum_{i=1}^{m_b} \beta_i b_i(D) \tag{2.6}$$

where m_a and m_B are positive integers, where $\alpha_1, \cdots, \alpha_{m_a}, \beta_1, \cdots, \beta_{m_b}$ are real parameters and where

$$a_i(D)=\sum_{k=1}^{l} a_{ik}D^k,\ i=1, \cdots, m_a;\ b_i(D)=\sum_{k=1}^{l} b_{ik}D^k,\ i=1, \cdots, m_b;$$

$$c(D)=\sum_{k=0}^{l} c_k D^k,\ d(D)=\sum_{k=0}^{l} d_k D^k.$$

It is assumed that the polynomials $a_i(D)$, $i = 1, \cdots, m_a$ are linearly independent over the reals, i.e., that there exists no nontrivial set of constants $\gamma_1, \cdots, \gamma_{m_a}$ such that $\Sigma_{i=1}^{m_a} \gamma_i a_i(D) = 0$ (the zero polynomial). It is similarly assumed that the $b_j(D)$, $j = 1, \cdots, m_b$ are linearly independent. These assumptions imply that max $(m_a, m_b) \le l$.

We note that for (2.4) to represent (in a generally nonminimal way) the plant (2.1) it is necessary (and sufficient) that $g(D) = k(D)p(D)$ and $h(D) = k(D)r(D)$ for some polynomial $k(D)$.

We illustrate the parameterization (2.4) for two cases of interest in adaptive control.

Example 2.1 (Direct Model Reference Adaptive Control): We assume that in (2.3) $r_i = 0$ for $i < d$, $r_d \neq 0$ [d is usually called the relative degree of (2.1)]. The reference model for the closed-loop plant is given by

$$p^*(D)y_m(t)=D^d v(t) \tag{2.7}$$

where $p^*(D) = 1 + \Sigma_{i=1}^{d} p_i^* D^i$ is a prespecified (stable) polynomial and where $y_m(t)$ and $v(t)$ are the reference model output and the command input, respectively. The control law is of the form

$$\left[\frac{1}{v_o} q(D)+E(D)\right] u(t)=-F(D)y(t)+q(D)v(t) \tag{2.8}$$

where $q(D) = 1 + \Sigma_{i=1}^{n} q_i D^i$ is a prespecified (stable) polynomial and where the polynomials $E(D) = \Sigma_{i=1}^{n} e_i D^i$ and $F(D) = \Sigma_{i=1}^{n} f_i D^i$ as well as the constant v_o are to be determined so as to satisfy the closed-loop performance specification, i.e., (2.7).

Substitution of (2.8) in (2.1) to eliminate $u(t)$ and equating the resultant expression with (2.7) yields the following nonminimal parameterization for (2.4):

$$g(D)=p^*(D)q(D)-D^d F(D) \tag{2.9}$$

$$h(D)=D^d\left[\frac{1}{v_o} q(D)+E(D)\right]. \tag{2.10}$$

Equating (2.9) and (2.10) with (2.5) and (2.6), respectively, gives $m_a = n$, $m_b = n+1$, $a_i(D) = b_i(D) = D^{d+i}$, $i = 1, \cdots, n$, $b_{n+1}(D) = D^d q(D)$, $c(D) = q(D)p^*(D)$ and $d(D) = 0$. The parameters to be estimated in the resultant model are $\alpha_i = -f_i$, $\beta_i = e_i$, $i = 1, \cdots, n$, and $\beta_{n+1} = 1/v_o$. Note that the α_i and β_i of the nonminimal model directly parameterize the model reference control law.

Example 2.2 (Direct Pole-Placement Adaptive Control): Since $p(D)$ and $r(D)$ are coprime, there exist polynomials $\gamma(D) = 1 + \Sigma_{i=1}^{n} \gamma_i D^i$, $\delta(D) = \Sigma_{i=1}^{n} \delta_i D^i$, $\rho(D) = 1 + \Sigma_{i=1}^{n} \rho_i D^i$ and $\sigma(D) = \Sigma_{i=1}^{n} \sigma_i D^i$ such that

$$\gamma(D)p(D)+\delta(D)r(D)=p^*(D)q(D) \tag{2.11}$$

$$\rho(D)p(D)+\sigma(D)r(D)=1 \tag{2.12}$$

where $p^*(D) = 1 + \Sigma_{i=1}^{n} p_i^* D^i$ and $q(D) = 1 + \Sigma_{i=1}^{n} q_i D^i$ are prespecified stable polynomials.

Multiplying (2.1) by $\sigma(D)\gamma(D)$ and using (2.11) gives

$$\sigma(D)p^*(D)q(D)y(t)=\sigma(D)r(D)[\delta(D)y(t)+\gamma(D)u(t)]. \tag{2.13}$$

Similarly, multiplying (2.1) by $\rho(D) \cdot \delta(D)$ and using (2.11) gives

$$\rho(D)p^*(D)q(D)u(t)=\rho(D)p(D)[\delta(D)y(t)+\gamma(D)u(t)]. \tag{2.14}$$

Adding (2.13) and (2.14), and making use of (2.12) gives the following nonminimal model of the plant:

$$[\delta(D)-p^*(D)q(D)\sigma(D)]y(t)=[p^*(D)q(D)\rho(D)-\gamma(D)]u(t). \quad (2.15)$$

This model is of the form (2.4) with the polynomials $g(D)$ and $h(D)$ parameterized as

$$g(D)=\delta(D)-p^*(D)q(D)\sigma(D) \quad (2.16)$$

$$h(D)=p^*(D)q(D)\rho(D)-\gamma(D). \quad (2.17)$$

Comparing (2.16) and (2.17) to (2.5) and (2.6), respectively, gives

$$c(D)=-p^*(D)q(D);\ d(D)=p^*(D)q(D) \quad (2.18)$$

$m_a = m_b = 2n$, $a_i(D) = b_i(D) = D^i$; $i = 1, \cdots, n$, $a_i(D) = b_i(D) = d^{i-n}p^*(D)q(D)$; $i = n+1, \cdots, 2n$, $\alpha_i = \delta_i$, $\beta_i = -\gamma_i$, $i = 1, \cdots, n$ and $\alpha_i = -\sigma_{i-n}$, $\beta_i = \rho_{i-n}$, $i = n+1, \cdots, 2n$.

Note that the common factor $k(D)$ of $g(D)$ and $h(D)$ in this case is the polynomial $[\sigma(D)\gamma(D) - \rho(D)\delta(D)]$. Following a procedure similar to the development in Example 2.1, it is readily verified that the nonminimal model (2.15) is obtained by using in (2.1) the control law

$$\gamma(D)u(t)=-\delta(D)y(t)+q(D)v(t) \quad (2.19)$$

which yields upon making use of (2.11) the closed-loop equation

$$p^*(D)y(t)=r(D)v(t). \quad (2.20)$$

Thus, the parameters γ_i and δ_i directly parameterize the pole assignment control law.

III. Conditions for Uniqueness of Parameters

Equation (2.4) with $g(D)$ and $h(D)$ given by (2.5) and (2.6) constitutes a model for the plant (2.1) if and only if the following *design identity* is satisfied

$$\left[c(D)+\sum_{i=1}^{m_a}\alpha_i a_i(D)\right] r(D)=\left[d(D)+\sum_{i=1}^{m_b}\beta_i b_i(D)\right] p(D). \quad (3.1)$$

Since we wish to use the (possibly nonminimal) model (2.4) to uniquely estimate the parameters $\alpha_1, \cdots, \alpha_{m_a}, \beta_1, \cdots, \beta_{m_b}$ from the plant input-output data, we conclude that a necessary and sufficient condition for existence and uniqueness of these parameters is that (3.1) is solvable by a unique vector

$$\theta^*=[\alpha_1, \cdots, \alpha_{m_a}, \beta_1, \cdots, \beta_{m_b}]^T. \quad (3.2)$$

For parameter estimation purposes it is convenient to represent the model (2.4) by a linear regression equation

$$c(D)y(t)-d(D)u(t)=\phi^T(t)\theta^* \quad (3.3)$$

where

$$\phi(t)=[-a_1(D)y(t), \cdots, -a_{m_a}(D)y(t), b_1(D)u(t), \cdots, b_{m_b}(D)u(t)]. \quad (3.4)$$

One can now use standard estimation procedures [10] to estimate the parameter vector θ^* using input-output data from the plant. For example, one could employ the recursive least squares (RLS) algorithm.

Consider now a sequence of vectors $\{\phi(i)\}_{i=i_o}^{\infty}$, $\phi(i) \in \mathbb{R}^m$ and denote by $\Phi_{m,N}(k)$ the $m \times N$ matrix

$$\Phi_{m,N}(k) := [\phi(k+1), \cdots, \phi(k+N)]. \quad (3.5)$$

We say that $\{\phi(i)\}$ is *spanning* (of order N) if there exists a number $\epsilon > 0$ and integers k and N such that

$$\lambda_{\min}(\Phi_{m,N}(k)\Phi_{m,N}^T(k))\geq\epsilon. \quad (3.6)$$

The sequence will be called *persistently spanning* (of order N) if there exists a sequence of positive numbers $\{\epsilon_i\}_{i=0}^{\infty}$ and positive integers k and N such that for all $i \geq 0$

$$\lambda_{\min}(\Phi_{m,N}(k+iN)\Phi_{m,N}^T(k+iN))\geq\epsilon_i \quad (3.7)$$

and it will be called *uniformly persistently spanning* (of order N) if for all $i \geq 0$, the ϵ_i in (3.7) can be chosen so that $\epsilon_i > \epsilon$ for some $\epsilon > 0$.

For a scalar sequence $\{u(i)\}_{i=i_o}^{\infty}$ we first construct an associated m-vector $\bar{u}_m(i)$ as follows:

$$\bar{u}_m(i)=[u(i+1), u(i+2), \cdots, u(i+m)]^T. \quad (3.8)$$

We then denote by $U_{m,N}(k)$ the $m \times N$ matrix

$$U_{m,N}(k) := [\bar{u}_m(k+1), \cdots, \bar{u}_m(k+N)]. \quad (3.9)$$

The spanning properties of $U_{m,N}(k)$ are defined just as for $\Phi_{m,N}(k)$.

Now it is known [10] that the RLS algorithm with covariance resetting yields a sequence of estimates that converges exponentially fast to θ^*, provided the sequence $\{\phi(t)\}$ of regression vectors is uniformly persistently spanning. It is further known [1], [3], [10] that the persistent spanning property of $\{\phi(t)\}$ can be generated from the input $u(t)$ provided a certain reachability property of $\phi(t)$ holds.

In Section IV we shall exhibit a linear plant in state-space form that generates the vector $\phi(t)$ as its output. We shall show that this plant, called the *associated-signal system* of (2.4), is output-reachable if and only if the design identity (3.1) has a unique solution vector θ^*.

Finally in Section V we will use this result to develop necessary and sufficient conditions on the input $u(t)$ so that $\{\phi(t)\}$ is uniformly persistently spanning.

IV. Reachability of the Regression Vector

We let the state of the associated-signal system be defined as

$$x(t)=[y(t-1), \cdots, y(t-l), u(t-1), \cdots, u(t-l)]^T. \quad (4.1)$$

It follows from (2.1) that $x(t)$ satisfies the following discrete-time linear state equation

$$x(t+1)=Ax(t)+bu(t) \quad (4.2)$$

where

$$A=\begin{bmatrix} A_{11}A_{12} \\ A_{21}A_{22}\end{bmatrix},\ b=\begin{bmatrix} b_1 \\ b_2\end{bmatrix} \quad (4.3)$$

with A_{ij} $l \times l$ matrices and b_i $l \times 1$ matrices as follows:

$$A_{11}=\begin{bmatrix} -p_1 \cdots -p_n\ 0 \cdots 0 \\ \qquad\qquad\qquad 0 \\ \qquad\qquad\qquad \cdot \\ \qquad I_{l-1} \qquad \cdot \\ \qquad\qquad\qquad 0 \end{bmatrix} \quad (4.4a)$$

$$A_{12}=\begin{bmatrix} r_1 \cdots r_n\ 0 \cdots 0 \\ 0 \\ \ \end{bmatrix} \quad (4.4b)$$

$$A_{21}=0 \quad (4.4c)$$

$$A_{22}=\begin{bmatrix} 0 \cdots\cdot 0 \\ \qquad \cdot \\ I_{l-1} \quad \cdot \\ \qquad 0 \end{bmatrix} \quad (4.4d)$$

$$b_1=\begin{bmatrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{bmatrix}\quad b_2=\begin{bmatrix}1\\0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{bmatrix}. \tag{4.4e}$$

The signal vector $\phi(t)$ is obtained by passing $x(t)$ through the following output equation:

$$\phi(t)=Cx(t) \tag{4.5}$$

where

$$C=\begin{bmatrix}C_{11} & 0\\ 0 & C_{22}\end{bmatrix} \tag{4.6}$$

with C_{11} an $m_a \times l$ matrix and C_{22} an $m_b \times l$ matrix as follows:

$$C_{11}=\begin{bmatrix}a_{11} & \cdots & a_{1l}\\ \vdots & & \vdots\\ a_{m_a1} & \cdots & a_{m_al}\end{bmatrix}\quad C_{22}=\begin{bmatrix}b_{11} & \cdots & b_{1l}\\ \vdots & & \vdots\\ b_{m_b1} & \cdots & b_{m_bl}\end{bmatrix}. \tag{4.7}$$

We note that our assumption of Section II that the $a_i(D)$ as well as the $b_i(D)$ are linearly independent sets implies that rank $C_{11}=m_a$ and rank $C_{22}=m_b$, whence rank $C=m_a+m_b$.

We recall that a linear system is output-reachable (from the input) if and only if every vector in its output space can be generated (reached) using a suitable input sequence. Stated equivalently, the system is output reachable if and only if the only vector orthogonal to all reachable outputs is the zero vector. Thus, the associated signal system (4.2)–(4.7) is output-reachable provided the only vector $\mu=[\mu_{11},\cdots,\mu_{1m_a},\mu_{21},\cdots,\mu_{2m_b}]^T$ satisfying

$$\mu^T\phi(t)=0 \tag{4.8}$$

for all $\phi(t)$ is $\mu=0$. Using the definition of $\phi(t)$ in (3.4), (4.8) becomes

$$\mu^T\phi(t)=\left[\sum_{i=1}^{m_a}\mu_{1i}a_i(D)\right]y(t)+\left[\sum_{j=1}^{m_b}\mu_{2j}b_j(D)\right]u(t)=0 \tag{4.9}$$

whence the associated-signal system is output-reachable if and only if (4.9) has as its only solution, for all signal sequences generated by (2.1), the trivial solution $\mu=0$. Multiplying (4.9) by $p(D)$ and employing (2.1), gives the (obvious) equivalent condition that the associated-signal system is output-reachable if and only if the equation

$$\left[\sum_{i=1}^{m_a}\mu_{1i}a_i(D)\right]r(D)+\left[\sum_{j=1}^{m_b}\mu_{2j}b_j(D)\right]p(D)=0 \tag{4.10}$$

has only the zero solution for μ. But (4.10) is the homogeneous equation of (3.1), whence the latter statement is equivalent to saying that (3.1) has a unique solution for $(\alpha_1,\cdots,\alpha_{m_a},\beta_1,\cdots,\beta_{m_b})$. We thus have the following elementary but important:

Theorem 4.1: Assume that (3.1) is solvable. Then this solution is unique if and only if the associated signal system is output reachable. □

We illustrate Theorem 4.1 by applying it to the second example in Section II.

Example 2.2 (Continued).

Equation (3.1) for this problem is

$$[\delta(D)-p^*(D)q(D)\sigma(D)]r(D)=[p^*(D)q(D)\rho(D)-\gamma(D)]p(D) \tag{4.11}$$

and this equation is of course solvable in view of the coprimeness assumption for $r(D)$ and $p(D)$. This equation has a unique solution if and only if the homogeneous equation

$$[\delta(D)-p^*(D)q(D)\sigma(D)]r(D)=[p^*(D)q(D)(\rho(D)-1)-(\gamma(D)-1)]p(D) \tag{4.12}$$

has no nonzero solution for the parameters $\delta_1,\cdots,\delta_n,\sigma_1,\cdots,\sigma_n,\rho_1,\cdots,\rho_n,\gamma_1,\cdots,\gamma_n$. Writing (4.12) alternatively as

$$D\left\{\left[\sum_{i=1}^{n}\delta_iD^{i-1}\right]r(D)+\left[\sum_{i=1}^{n}\gamma_iD^{i-1}\right]p(D)\right\}$$
$$-Dq(D)p^*(D)\left\{\left[\sum_{i=1}^{n}\sigma_iD^{i-1}\right]r(D)+\left[\sum_{i=1}^{n}\rho_iD^{i-1}\right]p(D)\right\}=0) \tag{4.13}$$

it is seen that (4.12) can have nontrivial solutions if degree $(q(D)p^*(D)) < 2n$. If on the other hand, degree $(q(D)p^*(D)) \geq 2n$ then every solution of (4.12) must satisfy

$$\left[\sum_{i=1}^{n}\delta_iD^{i-1}\right]r(D)+\left[\sum_{i=1}^{n}\gamma_iD^{i-1}\right]p(D)=0 \tag{4.14}$$

$$\left[\sum_{i=1}^{n}\sigma_iD^{i-1}\right]r(D)+\left[\sum_{i=1}^{n}\rho_iD^{i-1}\right]p(D)=0. \tag{4.15}$$

Finally, for (4.14) and (4.15) to have no nontrivial solutions, it is necessary and sufficient for $p(D)$ and $r(D)$ to be coprime. (This is based on the well-known fact that two polynomials $f(\lambda)$ and $g(\lambda)$ of degree n are coprime if and only if the equation $h(\lambda)f(\lambda)+k(\lambda)g(\lambda)=0$ has no nonzero solution polynomials $h(\lambda)$ and $g(\lambda)$ whose degrees are at most $n-1$.) We conclude the example with the following summary: The associated-signal system of the direct pole placement adaptive controller (as described, for example, in [11]) is output-reachable if and only if the following conditions both hold: i) $p(D)$, $r(D)$ are coprime; and ii) degree $(p^*(D)q(D)) \geq 2n$. □

V. Uniform Persistent Spanning

We consider in this section the problem of achieving uniform persistent spanning of the output from the input of linear plants

$$\left.\begin{aligned}x(k+1)&=Ax(k)+bu(k)\\ \phi(k)&=Cx(k)\end{aligned}\right\} \tag{5.1}$$

where $u(k)\in\mathbb{R}^1$, $x(k)\in\mathbb{R}^n$, and $\phi(k)\in\mathbb{R}^m$.

We first establish the following fact which extends a result of [3] to the case of output-reachability. We denote by μ the dimension of the reachable subspace of (5.1), i.e.,

$$\mu := \text{rank}\,[b, Ab, \cdots, A^{n-1}b]. \tag{5.2}$$

Lemma 5.1: Assume that the plant (5.1) is output-reachable. Then for any nonzero $m-$ vector α, there exist nonzero vectors η and ξ (dependent of α) such that for any arbitrary reachable x_k and any arbitrary $u_{k+1},\cdots,u_{k+\mu}$,

$$\eta'\bar{u}_\mu(k)=\alpha'\Phi_{m,\mu+1}(k)\xi \tag{5.3}$$

where $\Phi_{m,\mu+1}(k)$ and $\bar{u}_m(k)$ are as defined in (3.4) and (3.8).

Proof (Adapted from [3]): Define $z(k)=\alpha'\phi(k)$. The sequence $\{z_k\}$ can be regarded as an output of (5.1) evolving from x_k via the input sequence $u_{k+1},\cdots,u_{k+\mu}$. Hence, it is related to $\{u(k)\}$ by an equation of the form

$$m(D)z(k)=n(D)u(k) \tag{5.4}$$

where

$$n(D)=n_1D+\cdots+n_\mu D^\mu, \tag{5.5}$$

$$m(D)=1+m_1D+\cdots+m_\mu D^\mu. \tag{5.6}$$

The polynomial $z^\mu m(z) = D^{-\mu}m(D^{-1})$ is a divisor of the characteristic polynomial of A. Further, the polynomial $n(D)$ is nonzero by virtue of the output-reachability of (5.1).

From (5.4) we then have

$$z(k+\mu+1)+m_1z(k+\mu)+\cdots+m_\mu z(k+1) = n_1u(k+\mu)+\cdots+n_\mu u(k+1), \tag{5.7}$$

whence (5.3) follows with

$$\xi=[m_\mu, m_{\mu-1}, \cdots, m_1, 1]^T \tag{5.8}$$

$$\eta=[n_\mu, n_{\mu-1}, \cdots, n_1]^T. \tag{5.9}$$

□

With the aid of Lemma 5.1 the following result is easily established.

Proposition 5.2: Assume that the system (5.1) is output-reachable and let x_{k_o} be an arbitrary initial state. Then $\{\phi(t)\}$ will be uniformly persistently spanning of order N provided $\{\bar{u}_\mu(t)\}$ is uniformly persistently spanning of order $N-\mu$.

Proof (After [3]): With the use of (5.3) and the Cauchy-Schwarz inequality, we obtain

$$\eta^T\bar{u}_\mu(k)\bar{u}_\mu^T(k)\eta=\alpha^T\Phi_{m,\mu+1}(k)\xi\xi^T\Phi_{m,\mu+1}^T(k)\alpha \le \|\xi\|^2\alpha^T\Phi_{m,\mu+1}(k)\Phi_{m,\mu+1}^T(k)\alpha$$

hence,

$$\eta^T[U_{\mu,N-\mu}(k_o)U_{\mu,N-\mu}^T(k_o)]\eta=\eta^T\left[\sum_{k=k_o+1}^{k_o+N-\mu}\bar{u}_\mu(k)\bar{u}_\mu^T(k)\right]\eta$$

$$\le\|\xi\|^2\alpha^T\left[\sum_{k=k_o+1}^{k_o+N-\mu}\Phi_{m,\mu+1}(k)\Phi_{m,\mu+1}^T(k)\right]\alpha$$

$$=\|\xi\|^2\alpha^T\left[\sum_{k=k_o+1}^{k_o+N-\mu}\sum_{j=k+1}^{k+\mu+1}\phi(j)\phi^T(j)\right]\alpha$$

$$\le(\mu+1)\|\xi\|^2\alpha^T\sum_{k=k_o+2}^{k_o+N+1}\phi(j)\phi^T(j)\Bigg]\alpha$$

$$=(\mu+1)\|\xi\|^2\alpha^T[\Phi_{m,N}(k_o+1)\Phi_{m,N}^T(k_o+1)\Bigg]\alpha.$$

It follows that

$$\min_{\|\alpha\|=1}\alpha^T[\Phi_{m,N}(k_o+1)\Phi_{m,N}^T(k_o+1)\alpha \ge \frac{\|\eta^*\|^2}{(\mu+1)\|\xi^*\|^2}\lambda_{\min}[U_{\mu,N-\mu}(k_o)U_{\mu,N-\mu}^T(k_o)] \tag{5.10}$$

where

$$\|\eta^*\|^2:=\min_{\|\alpha\|=1}\|\eta(\alpha)\|^2 \text{ and } \|\xi^*\|^2:=\max_{\|\alpha\|=1}\|\xi(\alpha)\|^2$$

and the proof is complete. □

In typical adaptive control applications, the input $u(k)$ to (5.1) is accessible only through a feedback law of the form

$$u(k)=F_jx(k)+v(k) \tag{5.11}$$

where F_j is a periodically changing *feedback-gain matrix* (or *vector* in the case of single-input plants) with the subscript j denoting the jth period and where $v(k)$ is an external command input.

The closed-loop equation for (5.1) is then

$$\left.\begin{aligned}x(k+1)&=A_jx(k)+bv(k)\\ \phi(k)&=Cx(k)\end{aligned}\right\} \tag{5.12}$$

where $A_j = A + BF_j$ and the problem of persistent spanning becomes that of spanning the output $\phi(k)$ of (5.12) from $v(k)$.

Since the reachable subspace of a linear system is invariant under state feedback, so is the dimension μ and if (5.1) is output-reachable, so is (5.12) for every feedback gain F_j.

Suppose now that the F_j all take value in a compact region, so that the $A_j(=A + bF_j)$ lie in a compact region as well, say $\boldsymbol{D}$. Since the η^* and ξ^* of (5.10) depend continuously on the system data A, b, C, it follows that there are positive numbers v and ρ such that $\rho = \min\|\eta^*\|^2$ and $v = \max\|\xi^*\|^2$, where the minimization and maximization are performed over all systems A, b, C, with A taking value in $\boldsymbol{D}$. From (5.10) we then have that the output of every output-reachable system A, b, C, with A taking value in $\boldsymbol{D}$ satisfies the persistent-spanning inequality

$$\min_{\|\alpha\|=1}\alpha^T[\Phi_{m,N}(k_o+1)\Phi_{m,N}^T(k_o+1)]\alpha \ge\frac{\rho}{(\mu+1)v}\lambda_{\min}[V_{\mu,N-\mu}(k_o)V_{\mu,N-\mu}^T(k_o)]. \tag{5.13}$$

Inequality (5.13) implies immediately the following result which forms the basis for parameter convergence proofs in direct adaptive control algorithms employing parameter estimation in nonminimal models.

Theorem 5.3 (Uniform Persistent Spanning Under Block-Invariant State Feedback): Consider an output-reachable linear plant of the form (5.12) with $A_j(=A + bF_j)$ taking values in a compact region $\boldsymbol{D}$. Suppose that A_j is allowed to change only at time instances $t_j = k_o + jN$, $j = 0, 1, 2, \cdots$ with $N \ge 2\mu$ and is held constant otherwise. If for each index $j = 0, 1, 2, \cdots$

$$\lambda_{\min}[V_{\mu,N-\mu}(t_j)V_{\mu,N-\mu}^T(t_j)]\ge\epsilon$$

for some fixed $\epsilon > 0$, then the sequence $\{\phi(k)\}$ of (5.12) is uniformly persistently spanning (of order N) for any reachable initial state x_o. □

As a final remark it should be noted that Theorem 5.3 focuses attention on readable initial states and (in general) not on all arbitrary initial states. For the case of arbitrary initial states the reader is referred to [13].

ACKNOWLEDGMENT

The author wishes to express his thanks to G. C. Goodwin for suggesting the *structured-nonminimal-model* concept and for his help in preparing this revised manuscript.

REFERENCES

[1] G. C. Goodwin and E. K. Teoh, "Persistency of excitation in the presence of possibly unbounded signals," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 589-592, 1985.

[2] G. C. Goodwin, J. P. Norton, and M. N. Wiswanathan, "Persistency of excitation for non-minimal models of systems having pruely deterministic disturbances," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 589-590, 1985.

[3] J. B. Moore, "Persistency of excitation in extended least squares," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 60-68, 1983.

[4] B. D. O. Anderson and R. M. Johnstone, "Global adaptive pole positioning," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 11-22, 1985.

[5] B. D. O. Anderson and C. R. Johnson, Jr., "Exponential convergence of adaptive identification and control algorithms," *Automatic*, vol. 18, pp. 1-13, 1982.

[6] R. R. Bitmead, "Persistence of excitation conditions and the convergence of adaptive schemes," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 183-191, 1984.

[7] B. D. O. Anderson, "Exponential convergence and persistent excitation," in *Proc. 21st IEEE Conf. Decision Contr.*, 1982, pp. 12-17.

[8] S. Boyd and S. S. Sastry, "Necessary and sufficient conditions for parameter convergence in adaptive control," in *Proc. 22nd IEEE Conf. Decision Contr.*, 1983, pp. 1584-1587.

[9] S. Boyd and S. S. Sastry, "On parameters convergence in adaptive control," *Syst. Contr. Lett.*, pp. 311-319, 1983.

[10] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[11] H. Elliott, "Direct adaptive pole placement with application to nonminimum phase systems," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 720-722, 1982.

[12] A. Feuer, "A parametrization for model reference adaptive pole-placement," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 782-785, 1986.

[13] A. Feuer and M. Heymann, "On minimal spanning blocks of discrete linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 352-355, 1986.