Math. Control Signals Systems (1988) 1: 203–226

Mathematics of Control, Signals, and Systems

Persistency of Excitation Criteria for Linear, Multivariable, Time-Varying Systems*

I. M. Y. Mareels† and M. Gevers‡

Abstract. For continuous-time, multiple-input, multiple-output, linear systems, we present conditions under which the persistency of excitation of one regression vector implies the persistency of another regression vector derived from the first via a linear, dynamical transformation. We then introduce a definition of sufficient richness for vector input signals in the form of a persistency of excitation condition on a basis regression vector. Finally we establish input conditions which guarantee the persistency of excitation of a large class of regression vectors obtained from both time-invariant and time-varying systems.

Key words. Persistency of excitation, System identification, Adaptive control.

1. Introduction

Persistency of excitation and sufficient richness have a longstanding history. These concepts were first introduced as conditions for parameter identifiability in identification [L]. In the adaptive control context they were introduced to guarantee the exponential convergence of adaptive algorithms [A], [MN]. Their importance for the robustness of adaptive algorithms with respect to model errors was first recognized in [MN]. Later the concept of persistency of excitation was more fully developed in [IK] and [AB] where related notions such as "dominant richness" have been introduced.

A continuing research effort has already produced a wealth of results concerning the generation of persistently exciting "regression vectors" by filtering a scalar- or vector-valued input function through a dynamical linear, *time-invariant* filter. Both continuous-time [IK], [AB], [BS1] and discrete-time [IK], [AB], [BS], [GM], [GT] results are available. Our contribution is motivated by [BG], where persistency of excitation is used in a *nonstationary* environment. Also in [BG] a more

* Date received: May 18, 1987. Date revised: October 19, 1987. The work reported was performed while both authors were at the Department of Systems Engineering, Research School of Physical Sciences, Australian National University, G.P.O. Box 4, A.C.T. 2601, Australia.

† Laboratorium voor Regeltechniek, Rijksuniversiteit Gent, Grote Steenweg Noord 2, B9710 Gent, Belgium. This author is a research assistant with the National Fund for Scientific Research, Belgium, whose support is acknowledged.

‡ Laboratoire d'Automatique, de Dynamique et d'Analyse des Systèmes, Université de Louvain La Neuve, Bâtiment Maxwell, B-1348 Louvain La Neuve, Belgium.

general type of regression vector is introduced, which spurred us to consider in greater detail the relationship between different regression vectors.

Our results are presented in continuous time, but our methods of proof, which are new and rather simple, can be transliterated into the discrete-time case without difficulty. This is unlike most previous approaches which were either specifically conceived for continuous time [D], [BS2], [DAT] or discrete time [GM], [JA], [AJ], [M].

The main contributions of this paper are as follows. We introduce a new definition of sufficient richness for vector input signals in the form of the persistency of excitation of a basis regression vector, and we present conditions under which the persistency of excitation of one regression vector implies the persistency of excitation of another regression vector derived from the first by a linear dynamical transformation. This allows us, in particular, to establish the persistency of excitation of the output or state of a multiple-input, multiple-output (MIMO) system from the sufficient richness of its input vector.

Some of our results for time-invariant systems already exist, but only in the more restricted context of stationary signals (having a spectral measure) together with a definition of sufficient richness based on spectral lines [AB], [BS1], [GM], [BS2]. No stationarity assumption is required here. This important generalization allows us to treat time-varying systems.

Indeed, we derive conditions under which these latter persistency of excitation results extend to linear time-varying systems with both slow and fast time variations. To the authors' knowledge, the only other instances where persistency of excitation results for nonstationary systems are considered, are found in [D]. Our approach yields a simple proof for a more general result involving more general regressors in the presence of more general nonstationary effects; only slow time variations were considered in [D].

This paper is organized as follows. Section 2 contains the basic definition of persistency of excitation and introduces some notations. Section 3 contains some technical results, "swapping lemmas," which are used in Section 4 to infer the persistency of excitation of the output (the state) of one time-invariant MIMO system from the persistency of excitation of the output (state) of another MIMO system that is dynamically related to the first. This leads to a new operational definition of sufficient richness for vector input signals. In the case of single-input, single-output (SISO) systems a sharper result is obtained; it allows us to state exactly under what conditions persistency of excitation of a signal is or is not preserved through dynamical filtering. This is the object of Section 5. Finally, our most important results are in Section 6, where we give conditions under which the output (state) of a linear time-varying system is persistently exciting provided that the input is sufficiently rich. We present results for systems whose parameters vary slowly or rapidly compared with the dynamical behavior as excited by the input.

2. Definitions and Notations

In this section we define and comment on the concept of persistently exciting vector functions and introduce some notation.

Definition 2.1. A bounded locally square integrable vector function $\varphi: \mathbb{R} \to \mathbb{R}^n$ is said to be persistently exciting if there exist a constant s_0 and positive constants T_0 and α such that

$$\frac{1}{T_0}\int_s^{s+T_0} \varphi(t)\varphi^T(t)\,dt \geq \alpha I > 0 \qquad \text{for all} \quad s \geq s_0. \tag{2.1}$$

Comment 2.1. Notice that if (2.1) holds for a $T_0 > 0$ it also holds for all $T \geq T_0$, possibly with a different α-constant, however, not smaller than $\alpha/2$. In particular, (2.1) implies that

$$\liminf_{T\uparrow\infty} \lambda_{\min}\left[\frac{1}{T}\int_s^{s+T} \varphi(t)\varphi^T(t)\,dt\right] \geq \alpha/2 > 0 \tag{2.2}$$

uniformly in s, for all $s > s_0$. Alternatively, (2.2) obviously implies (2.1) for bounded, (locally) square integrable φ's. Notice that if (2.1) holds for a given s_0 it also holds for $s_0 = 0$, possibly with a larger T_0 and a smaller α. We choose Definition (2.1) because it offers more flexibility.

The importance of the concept "persistently exciting function" is captured by the observation that $\varphi(t)$ being persistently exciting is equivalent to the uniform (exponential) asymptotic stability of the zero solution of the differential equation

$$\dot{x}(t) = -\varphi(t)\varphi^T(t)x(t) \tag{2.3}$$

[A], [MN]. A far-reaching interpretation of this equivalence connects the persistency of excitation of $\varphi(t)$ to the concepts of uniform observability and identifiability, see, e.g., [A], [MN], and [AJ].

Throughout this paper we assume that the signal vector $u(t)$ is bounded and locally (Riemann) integrable. In addition we use the following notation: PE stands for persistently exciting or persistency of excitation; SR stands for sufficiently rich; BIBS stability stands for bounded input, bounded state stability; D denotes derivative with respect to time; $D_2 uv^T \equiv uDv^T$; $\|\cdot\|$ denotes the vector 2-norm or the corresponding induced matrix norm; and A^{adj} stands for the adjoint of the matrix A.

3. Some Swapping Lemmas

If a vector φ_1 is PE, then clearly a vector $\varphi_2 = A\varphi$, with A constant and of full row rank is also PE. In analyzing adaptive identification and/or control algorithms it is important that we can infer the persistency of excitation of one vector-valued function from the persistency of excitation of another vector-valued function which is linked to the first by some linear, dynamical transformation. For example, we may want to infer the persistency of excitation of φ_2 from that of φ_1 when φ_1 and φ_2 are as follows:

$$\varphi_1 = \left(u \quad \frac{u}{s+d} \quad \cdots \quad \frac{u}{(s+d)^{n-1}}\right)^T,$$

$$\varphi_2 = \left(u \quad \frac{u}{s+c_1} \quad \cdots \quad \frac{u}{s+c_{n-1}}\right)^T,$$

with d positive and the c_i positive and distinct.

In this section we set up some tools allowing us to do so. Consider the linear systems

$$\dot{x}_i(t) = F_i x_i(t) + G_i u(t), \tag{3.1}$$

$$y_i(t) = H_i x_i(t) + J_i u(t), \tag{3.2}$$

where u is the m-dimensional input, x_i the n_i-dimensional state, and y_i the p_i-dimensional output; F_i, G_i, H_i, J_i are (real) matrices of appropriate dimensions. Define

$$p^{(i)}(s) = \det(sI - F_i), \tag{3.3}$$

$$Q^{(i)}(s) = H_i(sI - F_i)^{\text{adj}} + J_i \det(sI - F_i) \tag{3.4}$$

and introduce the notation

$$p^{(i)}(s) = p_0^{(i)} s^{n_i} + p_1^{(i)} s^{n_i - 1} + \cdots + p_{n_i-1}^{(i)} s + p_{n_i}^{(i)} \qquad \text{with} \quad p_0^{(i)} = 1, \tag{3.5}$$

$$Q^{(i)}(s) = Q_0^{(i)} s^{n_i} + Q_1^{(i)} s^{n_i - 1} + \cdots + Q_{n_i-1}^{(i)} s + Q_{n_i}^{(i)}. \tag{3.6}$$

The following result is a direct consequence of the Cayley–Hamilton theorem—see p. 657 of [K].

Lemma 3.1. *For the linear systems* (3.1) *and* (3.2) *with the definitions* (3.3)–(3.6) *we have that*

$$Q_k^{(i)} = J_i p_k^{(i)} + \sum_{l=0}^{k-1} p_{k-1-l}^{(i)} H_i F_i^l G_i, \qquad k = 1, \ldots, n_i, \tag{3.7}$$

and

$$Q_0^{(i)} = J_i. \tag{3.8}$$

The next lemma links the inner product of y_i and a suitable test function ψ with the inner product of u and ψ.

Lemma 3.2. *For any row vector test function $\psi: \mathbb{R} \to \mathbb{R}^r$, n_i times continuously differentiable on $\mathbb{R}$ we have*

$$\int_{t_0}^{t_1} p^{(i)}(-D_2) y_i(t) \psi(t)\, dt = \int_{t_0}^{t_1} Q^{(i)}(-D_2) u(t) \psi(t)\, dt + g^{(i)}(t_0, t_1), \tag{3.9}$$

where $g^{(i)}$ is given by

$$\begin{aligned} g^{(i)}(t_0, t_1) &= \sum_{k=1}^{n_i} \sum_{l=0}^{k-1} (-1)^{k-l} p_{n_i-k}^{(i)} H_i F_i^l x_i(t) D^{k-1-l} \psi(t) \Big|_{t_0}^{t_1} \\ &= -H_i(-D_2 - F_i)^{\text{adj}} x_i(t) \psi(t) \Big|_{t_0}^{t_1}. \end{aligned} \tag{3.10}$$

Proof. Using integration by parts, and equations (3.5) and (3.6),

$$\int_{t_0}^{t_1} y_i(t) D^k \psi(t)\, dt = \int_{t_0}^{t_1} H_i x_i(t) D^k \psi(t)\, dt + \int_{t_0}^{t_1} J_i u(t) D^k \psi(t)\, dt$$

$$= H_i x_i(t) D^{k-1}\psi(t)|_{t_0}^{t_1} - \int_{t_0}^{t_1} H_i F_i x_i(t) D^{k-1}\psi(t)\,dt$$
$$+ \int_{t_0}^{t_1} J_i u(t) D^k \psi(t)\,dt - \int_{t_0}^{t_1} H_i G_i u(t) D^{k-1}\psi(t)\,dt$$
$$= \sum_{l=0}^{k-1} (-1)^l H_i F_i^l x_i(t) D^{k-1-l}\psi(t)|_{t_0}^{t_1} + (-1)^k \int_{t_0}^{t_1} H_i F_i^k x_i(t)\psi(t)\,dt$$
$$+ \int_{t_0}^{t_1} J_i u(t) D^k \psi(t)\,dt - \int_{t_0}^{t_1} \sum_{l=0}^{k-1} (-1)^l H_i F_i^l G_i u(t) D^{k-1-l}\psi(t)\,dt.$$

Hence, using the definition of $p^{(i)}(-D_2)$,

$$\int_{t_0}^{t_1} p^{(i)}(-D_2) y_i(t)\psi(t)\,dt$$
$$= \sum_{k=0}^{n_i} (-1)^k p_{n_i-k}^{(i)} \int_{t_0}^{t_1} J_i u(t) D^k \psi(t)\,dt$$
$$+ \sum_{k=1}^{n_i} (-1)^{k+1} p_{n_i-k}^{(i)} \sum_{l=0}^{k-1} (-1)^l \int_{t_0}^{t_1} H_i F_i^l G_i u(t) D^{k-1-l}\psi(t)\,dt$$
$$+ \sum_{k=1}^{n_i} (-1)^k p_{n_i-k}^{(i)} \sum_{l=0}^{k-1} (-1)^l H_i F_i^l x(t) D^{k-1-l}\psi(t)|_{t_0}^{t_1}.$$

Rearranging the double sums, and using Lemma 3.1, we arrive at

$$\int_{t_0}^{t_1} p^{(i)}(-D_2) y_i(t)\psi(t)\,dt = \int_{t_0}^{t_1} Q^{(i)}(-D_2) u(t)\psi(t)\,dt - H_i(-D_2 - F_i)^{\mathrm{adj}} x_i(t)\psi(t)|_{t_0}^{t_1}.$$

■

Expression (3.9) relates y_i to u since, roughly speaking, $p^{(i)}(D)y_i(t) = Q^{(i)}(D)u(t)$, without imposing restrictive assumptions on the differentiability of u and y_i. The following result goes a step further, establishing similar relationships between the outputs y_1, y_2 produced by the same input u but by different linear systems. We have:

Lemma 3.3. *Consider the systems* (3.1)–(3.4) *and assume there exist polynomial matrices* $T^{(1)}(s)$, $T^{(2)}(s)$ *such that*

$$T^{(1)}(s)Q^{(1)}(s) = T^{(2)}(s)Q^{(2)}(s). \tag{3.11}$$

Then for any row vector test function $\psi\colon \mathbb{R} \to \mathbb{R}^r$, *sufficiently continuously differentiable on* $\mathbb{R}$, $\psi \in C^n$, $n = \max(n_2 + \deg T^{(2)}, n_1 + \deg T^{(1)})$, *we have that*

$$\int_{t_0}^{t_1} T^{(1)}(-D_2)p^{(1)}(-D_2)y_1(t)\psi(t)\,dt$$
$$= \int_{t_0}^{t_1} T^{(2)}(-D_2)p^{(2)}(-D_2)y_2(t)\psi(t)\,dt + g^{(1,2)}(t_0, t_1), \tag{3.12}$$

where $g^{(1,2)}$ *is given by*

$$g^{(1,2)}(t_0, t_1) = -T^{(1)}(-D_2)H_1(-D_2 - F_1)^{\mathrm{adj}}x_1(t)\psi(t)|_{t_0}^{t_1} + T^{(2)}(-D_2)H_2(-D_2 - F_2)^{\mathrm{adj}}x_2(t)\psi(t)|_{t_0}^{t_1}. \quad (3.13)$$

Proof. The proof follows from a repeated application of Lemma 3.2. ■

4. Time-Invariant Systems

In this section we show that if the output of a time-invariant MIMO system is persistently exciting, the output of a "related" time-invariant MIMO system is also persistently exciting. Roughly speaking, the relationship is in terms of the zeros of the second system being a subset of the zeros of the first. A precise statement is as follows.

Theorem 4.1. *Consider two* MIMO *systems* (*defined as in* (3.1)–(3.6); $i = 1, 2$) *with the following assumptions*:

A.1. $u(t) \in L_\infty$.

A.2. $\mathrm{Re}\,\lambda_j(F_i) < 0$, $i = 1, 2$, $j = 1, \ldots, n_i$, *and* $n_1 \geq n_2$.

A.3. *There exist constants* t_1, $\alpha_1 > 0$, $\beta_1 > 0$, $T_1 > 0$ *such that*

$$\beta_1 I \geq \frac{1}{T}\int_t^{t+T} y_1(\tau)y_1^T(\tau)\,d\tau \geq \alpha_1 I \qquad \text{for all} \quad T \geq T_1, \quad t \geq t_1. \quad (4.1)$$

A.4. *There exists a constant matrix* $R \in \mathbb{R}^{p_2 \times p_1}$ *of full row rank such that*

$$Q^{(2)}(s) = RQ^{(1)}(s). \quad (4.2)$$

Then there exist constants $\alpha_2 > 0$, $\beta_2 > 0$, *and* $T_2 \geq T_1$ *such that*

$$\beta_2 I \geq \frac{1}{T_2}\int_t^{t+T_2} y_2(\tau)y_2^T(\tau)\,d\tau \geq \alpha_2 I \qquad \text{for all} \quad t \geq t_1. \quad (4.3)$$

Proof. For any $t \geq t_1$ and $T \geq T_1$, define $\varphi(\tau) \in \mathbb{R}^{p_1}$ on $(t, t + T)$ as the solution of

$$p^{(1)}(-D)\varphi(\tau) = y_1^T(\tau), \qquad \varphi(t + T) = \varphi^{(1)}(t + T) = \cdots = \varphi^{(n_1-1)}(t + T) = 0, \quad (4.4)$$

and $\psi(\tau) \in \mathbb{R}^{p_1}$ as

$$\psi(\tau) = p^{(2)}(-D)\varphi(\tau), \qquad \tau \in (t, t + T) \quad (4.5)$$

(ψ is well defined because $n_1 \geq n_2$). The Cauchy–Schwarz inequality yields

$$\int_t^{t+T} y_2(\tau)y_2^T(\tau)\,d\tau \int_t^{t+T} \psi^T(\tau)\psi(\tau)\,d\tau \geq \left(\int_t^{t+T} y_2(\tau)\psi^T(\tau)\,d\tau\right)\left(\int_t^{t+T} y_2(\tau)\psi^T(\tau)\,d\tau\right)^T. \quad (4.6)$$

By Lemma 3.3

$$\int_t^{t+T} y_2(\tau)\psi^T(\tau)\,d\tau = \int_t^{t+T} Ry_1(\tau)y_1^T(\tau)\,d\tau - g^{(1,2)}(t, t+T). \tag{4.7}$$

It follows by A.1 and A.2 that there exist K_1, K_2 such that

$$\|\psi(t)\| \le K_1 < \infty \quad \text{and} \quad \|g^{(1,2)^T}(t, t+T)\| \le K_2 < \infty \quad \text{for all } t, T, \tag{4.8}$$

where $g^{(1,2)^T}$ denotes the transpose of $g^{(1,2)}$. Here and in the sequel of this proof, the underlying norm is the vector 2-norm. Therefore, for $T \ge T_1$,

$$\frac{1}{T}\int_t^{t+T} y_2(\tau)y_2^T(\tau)\,d\tau \ge \frac{1}{K_1^2T^2}[\alpha_1^2T^2RR^T - 2\beta_1TK_2\|R\|I]. \tag{4.9}$$

Denoting

$$T^* = \frac{2\beta_1K_2\|R\|}{\alpha_1^2\lambda_{\min}(RR^T)} \tag{4.10}$$

and taking $T_2 > \max(T_1, T^*)$ yields (4.3) with $\beta_2 = \|y_2\|_\infty$ and

$$\alpha_2 = \frac{1}{K_1^2}\left[\alpha_1^2\lambda_{\min}(RR^T) - \frac{2\beta_1K_2\|R\|}{T_2}\right]. \quad \blacksquare \tag{4.11}$$

Comment 4.1. We might wonder how the bound T^* in (4.10) can be reduced, or α_2 in (4.11) increased. This will be particularly relevant to our time-varying results of Section 6. Suppose $p^{(1)}(s)$ and $p^{(2)}(s)$ are fixed, and note that $\psi = (p^{(2)}(-D)/p^{(1)}(-D))y_1$, and that $\|g^{(1,2)}\| = O(\|\psi\|\,\|y_1\|)$. Assume now that the input signal amplitude is increased by a factor C. Then it follows that α_1, β_1, K_1^2, and K_2 are increased by a factor C^2. Therefore β_1K_2/α_1^2, and hence T^*, remain essentially unchanged, and the same holds for α_1/K_1^2 and β_1/K_1^2. It follows from (4.11) that α_2 is roughly proportional to C^2, the input power. Another way to increase α_2 is by increasing α_1/β_1, i.e., by decreasing the condition number of the PE matrix (4.1).

Comment 4.2. In the case of periodic input signals ($u(t+T) = u(t)$), the result of Theorem 4.1 can be strengthened in the sense that (4.3) holds for $T_2 = T_1 = T$, provided t is large enough. This follows from the fact that, for periodic signals, the term $g^{(1,2)}(t, t+T)$ in (4.7) decays exponentially fast as t approaches infinity.

Comment 4.3. It is possible to define $\varphi(t)$ as the solution of (4.4) with the end conditions determined in such a way that $g^{(1,2)}(t, t+T)$ is zero for almost all t. The requirement $g^{(1,2)}(t, t+T) = 0$ imposes a linear relation in $\varphi^{(i)}(t)$ and $\varphi^{(i)}(t+T)$ ($i = 0, \dots, n-1$). A detailed and cumbersome analysis reveals that it is possible to satisfy this relation generically. The present approach is more straightforward, yielding more easily obtainable estimates for α_2, β_2, and T_2; however, it fails to show that $T_1 = T_2$ generically, which would follow from the above-mentioned alternative.

Corollary 4.1. *If $x_1(t)$ and $x_2(t)$ are generated by* (3.1) *with assumptions* A.1 *and* A.2, *and if $x_1(t)$ is PE and there exists a constant matrix $R \in \mathbb{R}^{n_2 \times n_1}$ such* *that*

$$(sI - F_2)^{\mathrm{adj}} G_2 = R(sI - F_1)^{\mathrm{adj}} G_1, \tag{4.12}$$

then $x_2(t)$ *is PE.*

Comment 4.4. In [BS2] (and [BS] where the discrete-time case is discussed) a result analogous to this corollary is presented. The main difference is that here the signals (u, y) are not required to be stationary which is fundamental in the approach presented in [BS] and [BS2]. It is necessary to remove this condition in order to discuss PE (or SR) in the context of linear time-varying systems.

Notice that the output vectors $y_1(t)$ and $y_2(t)$ of Theorem 4.1 can also be written (mixing time and Laplace transform notations)

$$y_i(t) = \frac{1}{p^{(i)}(s)} Q^{(i)}(s) u(t), \qquad i = 1, 2. \tag{4.13}$$

The result then says that if $p^{(1)}(s)$ and $p^{(2)}(s)$ are Hurwitz, with $\deg(p^{(1)}) \geq \deg(p^{(2)})$ and if $Q^{(2)}(s) = RQ^{(1)}(s)$, then the persistency of excitation of $y_1(t)$ implies the persistency of excitation of $y_2(t)$. It seems natural, then, to characterize the sufficient richness of an input signal $u(t)$ in terms of the PE of a particular regressor vector $\varphi(t)$ from which all regression vectors generated by systems of a given order can be derived by a PE preserving dynamical transformation.

Definition 4.1. A signal $u(t)\colon \mathbb{R} \to \mathbb{R}^m$ is called sufficiently rich of order $(n_1, \ldots, n_m/n)$ if, for any $\gamma > 0$, there exist constants t_1, $\alpha > 0$, and $T > 0$ such that

$$\frac{1}{T} \int_t^{t+T} \psi(\tau)\psi^T(\tau)\, d\tau \geq \alpha I \qquad \text{for all} \quad t \geq t_1 \tag{4.14}$$

with

$$\psi(t) = \frac{1}{(s+\gamma)^{n-1}} \left[\begin{array}{ccc} 1 & 0 & 0 \\ s & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ s^{n_1-1} & 0 & 0 \\ 0 & 1 & \vdots \\ \vdots & s & \vdots \\ \vdots & \vdots & \vdots \\ 0 & s^{n_2-1} & 0 \\ \hline & \vdots & \\ \hline 0 & 0 & 1 \\ \vdots & \vdots & s \\ \vdots & \vdots & \vdots \\ 0 & 0 & s^{n_m-1} \end{array} \right] u(t), \tag{4.15}$$

where $n \geq \max n_i$.

The following result now follows immediately from Definition 4.1 and Theorem 4.1.

Theorem 4.2. *Let $u(t)$: $\mathbb{R} \to \mathbb{R}^m$ be sufficiently rich of order* $(n_1 + 1, \ldots, n_m + 1/n + 1)$ *and let*

$$y(t) = \frac{1}{p(s)} Q(s)u(t) \tag{4.16}$$

with the following assumptions:

(1) *$p(s)$ is scalar and Hurwitz with* $\deg(p(s)) \leq n$.
(2) *$Q(s)$ is a polynomial matrix of dimension $p \times m$ with column degrees $n_1, \ldots, n_m$, and* $\deg q_{ij}(s) \leq \deg p(s)$ *for all i, j.*
(3) *The system* (4.16) *is output reachable, i.e., there exists no $\alpha \neq 0$, $\alpha \in \mathbb{R}^p$ such that*

$$\alpha^T Q(s) = 0 \quad \textit{for all } s.$$

Then $y(t)$ is PE.

Proof. With $\psi(t)$ as defined by (4.15), it is easy to see that $y(t)$, given by (4.16), can be generated as

$$y(t) = \frac{(s + \gamma)^n}{p(s)} R\psi(t)$$

for some $R \in \mathbb{R}^{p \times \sum n_i}$. By assumption A.3, R has full row rank. The result then follows from Theorem 4.1. ■

Comment 4.5. Theorem 4.2 allows us to establish easily the persistency of excitation of a number of regression vectors that commonly arise in adaptive estimation and adaptive control problems. In particular, if $u(t)$: $\mathbb{R} \to \mathbb{R}^m$ is PE of order $(n, n, \ldots, n/n)$, then the following commonly used regression vectors are PE:

$$\varphi_1(t) = \left(u \quad \frac{u}{s + d} \quad \cdots \quad \frac{u}{(s + d)^{n-1}}\right)^T,$$

$$\varphi_2(t) = \left(u \quad \frac{u}{s + c_1} \quad \cdots \quad \frac{u}{s + c_{n-1}}\right)^T,$$

$$\varphi_3(t) = \frac{1}{p(s)}(u \quad su \quad \cdots \quad s^{n-1}u)^T,$$

provided $d > 0$ for $\varphi_1(t)$, the c_i are positive and distinct for $\varphi_2(t)$, and $p(s)$ is a Hurwitz polynomial of degree $n - 1$ for $\varphi_3(t)$. To prove this, the only task is to establish the output reachability of the system from u to φ, which is easily done.

In some problems the regression vector is of the following type:

$$\varphi(t) = H(s)\begin{bmatrix} y(t) \\ u(t) \end{bmatrix}, \qquad y(t)\colon \mathbb{R} \to \mathbb{R}^p, \quad u(t)\colon \mathbb{R} \to \mathbb{R}^m, \quad \varphi(t)\colon \mathbb{R} \to \mathbb{R}^r, \tag{4.17}$$

where $y(t)$ is the output of a time-invariant system driven by $u(t)$ (see, e.g., [BG]):

$$y(t) = G(s)u(t). \tag{4.18}$$

To infer persistency of excitation of $\varphi(t)$ from a sufficiently rich $u(t)$, the main difficulty is to prove that the system (4.17)–(4.18) is output reachable. We have the following result.

Theorem 4.3. *Consider the system* (4.17)–(4.18) *with $H(s)$ and $G(s)$ proper. Let $H(s) = Q(s)/p(s)$ where $p(s)$ is the least common denominator of the entries of $H(s)$, and $G(s) = A^{-1}(s)B(s)$, a left coprime polynominal matrix fraction description* (PMFD). *Then $\varphi(t)$ is output reachable from $u(t)$ if and only if*

$$\begin{aligned} &\alpha^T Q(s) = \beta^T(s)[A(s) \,\vdots\, -B(s)] \qquad \textit{for all } s, \\ &\textit{for some} \quad \alpha \in \mathbb{R}^r \quad \textit{and some} \quad \beta(s) \in \mathbb{R}^p[s], \\ &\textit{implies} \quad \alpha = 0, \quad \textit{and} \quad \beta(s) \equiv 0. \end{aligned} \tag{4.19}$$

Proof. Let $G(s) = M(s)R^{-1}(s)$, where $M(s)R^{-1}(s)$ is a right coprime PMFD, and note that

$$\varphi(t) = H(s)\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = H(s)\begin{bmatrix} G(s) \\ I_m \end{bmatrix} u(t) \triangleq T(s)u(t). \tag{4.20}$$

Now $\varphi(t)$ is output reachable from $u(t)$ if

$$\alpha^T T(s) = 0 \quad \text{for all } s, \quad \text{with} \quad \alpha \in \mathbb{R}^r \quad \text{implies} \quad \alpha = 0 \tag{4.21}$$

or, equivalently, if,

$$\alpha^T Q(s)\begin{bmatrix} M(s) \\ R(s) \end{bmatrix} = 0 \quad \text{for all } s, \quad \text{with} \quad \alpha \in \mathbb{R}^r \quad \text{implies} \quad \alpha = 0. \tag{4.22}$$

Next we note that, since $A^{-1}(s)B(s) = M(s)R^{-1}(s)$, it follows that $[A(s) \,\vdots\, -B(s)]$ is in the left nullspace of $\begin{bmatrix} M(s) \\ R(s) \end{bmatrix}$. Since $\begin{bmatrix} M(s) \\ R(s) \end{bmatrix}$ has dimension $(p + m) \times m$ and rank m, and since the p rows of $[A(s) \,\vdots\, -B(s)]$ are linearly independent for all s by coprimeness, it follows that they span this nullspace.

a. *Sufficiency.* Suppose (4.19) holds, and suppose that there exists a vector $\alpha \in \mathbb{R}^r$ such that

$$\alpha^T Q(s)\begin{bmatrix} M(s) \\ R(s) \end{bmatrix} = 0 \quad \text{for all } s. \tag{4.23}$$

Then $\alpha^T Q(s)$ is in the left nullspace of $\begin{bmatrix} M(s) \\ R(s) \end{bmatrix}$. Hence $\alpha^T Q(s) = \beta^T(s)[A(s) \,\vdots\, -B(s)]$ for some $\beta(s)$, and this implies $\alpha = 0, \beta(s) \equiv 0$ by (4.19); hence $\varphi(t)$ is output reachable from $u(t)$.

b. *Necessity.* Suppose (4.22) holds, and assume that $\alpha^T Q(s) = \beta^T(s)[A(s) \,\vdots\, -B(s)]$

for all s, for some $\alpha \in \mathbb{R}^r$ and $\beta(s) \in \mathbb{R}^p[s]$. Then

$$\alpha^T Q(s)\begin{bmatrix} M(s) \\ R(s) \end{bmatrix} = \beta^T(s)[A(s) \ \vdots \ -B(s)]\begin{bmatrix} M(s) \\ R(s) \end{bmatrix} = 0 \quad \text{for all } s. \tag{4.24}$$

Hence $\alpha = 0$ by (4.22), and therefore $\beta^T(s)[A(s) \ \vdots \ -B(s)] = 0$ for all s. This implies $\beta(s) \equiv 0$ by coprimeness of $[A(s), B(s)]$. ■

Comment 4.6. We note that condition (4.19) implies in particular that $H(s)$ must be output reachable.

Corollary 4.2. *Consider the system* (4.17)–(4.18) *with* (4.18) SISO. *Let* $H(s) = Q(s)/p(s)$ *as before. Then a sufficient condition for output reachability of* $\varphi(t)$ *from* $u(t)$ *is*:

(1) $H(s)$ *is output reachable.*
(2) $\deg Q(s) < \deg G(s)$, *where* $\deg Q(s) = \max\{\deg q_{ij}(s)\}$ *and* $\deg G(s)$ *is the McMillan degree of* $G(s)$.

Proof. Let $G(s) = b(s)/a(s)$ where $a(s)$ and $b(s)$ are coprime polynominals in s. Then $\varphi(t)$ is output reachable if

$$\alpha^T Q(s)\begin{bmatrix} b(s) \\ a(s) \end{bmatrix} = 0 \qquad \text{for all } s \text{ implies} \quad \alpha = 0. \tag{4.25}$$

By (1), $\alpha^T Q(s) = 0$ for all s implies $\alpha = 0$, and by the degree condition and the coprimeness of $[a(s), b(s)]$, $\alpha^T Q(s)(\neq 0)$ cannot be in the left nullspace of $\begin{bmatrix} b(s) \\ a(s) \end{bmatrix}$. ■

It follows immediately that if $u(t)$ is sufficiently rich of order $(2n/2n)$ and if (4.18) is SISO with McMillan degree n, then the following commonly used regression vectors (see, e.g., [AB] and [BG]) are PE:

$$\varphi_4(t) = \frac{1}{(s+\alpha)^{n-1}}[y \quad sy \quad \cdots \quad s^{n-1}y \quad u \quad su \quad \cdots \quad s^{n-1}u]^T,$$

$$\varphi_5(t) = \left[y \quad \frac{y}{s+c_1} \quad \cdots \quad \frac{y}{s+c_{n-1}} \quad u \quad \frac{u}{s+c_1} \quad \cdots \quad \frac{u}{s+c_{n-1}}\right]^T,$$

with $\alpha > 0$ in $\varphi_4(t)$ and the c_i positive and distinct in $\varphi_5(t)$. This follows from the fact that these regression vectors can be written as $\varphi(t) = (Q(s)/p(s))\begin{bmatrix} y(t) \\ u(t) \end{bmatrix}$ with $Q(s)/p(s)$ output reachable by Theorem 4.3 and with $\deg Q(s) \leq n - 1$ and $\deg(p(s)) = 2n - 1$.

5. A Special Case: SISO Systems

In the previous section we have related the persistency of excitation of the output of one linear time-invariant system to the persistency of excitation of the output of another linear time-invariant system driven by the same input. Restricting ourselves to SISO systems it is possible to derive a more complete result, describing how persistency of excitation is propagated through linear systems.

Lemma 5.1. *Consider the following time-invariant* SISO *systems with transfer functions*

$$y_i(t) = H_i(s)u(t), \qquad i = 1, 2, \tag{5.1}$$

$$H_i(s) = \frac{q_i(s)}{p_i(s)}. \tag{5.2}$$

Assume that the $H_i(s)$'s are strictly stable, proper but not strictly proper, and have the same zeros with zero real part. Under these conditions, if y_1 is PE, *i.e., if there exist $\alpha_1, \beta_1, T_1 > 0$ and t_1 such that*

$$\alpha_1 \leq \frac{1}{T_1}\int_t^{t+T_1} y_1^2(\tau)\,d\tau \leq \beta_1 \qquad \text{for} \quad t \geq t_1, \tag{5.3}$$

then y_2 is PE, *i.e., there exist $\alpha_2, \beta_2, T_2 > 0$ and t_2 such that*

$$\alpha_2 \leq \frac{1}{T_2}\int_t^{t+T_2} y_2^2(\tau)\,d\tau \leq \beta_2 \qquad \text{for} \quad t \geq t_2. \tag{5.4}$$

Proof. Under the assumptions we have

$$2(\deg p_1 - \deg q_1) \geq \deg p_2 - \deg q_2. \tag{5.5}$$

In fact, both sides of (5.5) are zero. The Cauchy–Schwarz inequality yields

$$\int_t^{t+T_2} y_2^2(\tau)\,d\tau \geq \left(\int_t^{t+T_2} y_2(\tau)\varphi(\tau)\,d\tau\right)^2 \Big/ \int_t^{t+T_2} \varphi^2(\tau)\,d\tau.$$

Defining $\varphi(t)$ and $\psi(t)$ as the unique bounded solutions of

$$\varphi(t) = p_2(-D)\bar{q}_1(-D)\psi(t), \tag{5.6}$$

$$y_1(t) = p_1(-D)\bar{q}_2(-D)\psi(t), \tag{5.7}$$

where $q_i(D) = \bar{q}_i(D)\bar{q}(D)$, $\bar{q}$ contains the common zeros of q_1 and q_2, $\bar{q}_i$ has no zeros with zero real part (by assumption), hence guaranteeing the existence of ψ. Notice that (5.5) implies that φ is well defined via (5.6). Using the swapping Lemma 3.3 we then have

$$\int_t^{t+T_2} y_2^2(\tau)\,d\tau \geq \left(\int_t^{t+T_2} y_1^2(\tau)\,d\tau + g(t, t+T_2)\right)^2 \Big/ \int_t^{t+T_2} \varphi^2(\tau)\,d\tau.$$

The proof continues along the lines of the proof of Theorem 4.1. ■

Lemma 5.2. *Consider the two time-invariant* SISO *systems* (5.1) *and* (5.2). *Assume now that the transfer functions $H_i(s)$ are strictly stable, strictly proper, and have the same zeros with zero real part. Then y_2 is* PE *if y_1 is* PE.

Proof. If H_1 and H_2 do not satisfy condition (5.5), then a repeated application of the proof of Lemma 5.1 with condition (5.5) yields the desired result. Indeed, it is possible to construct a finite sequence of $H_2^{(j)}$, $j = 1, \ldots, m$, such that each pair $(H_1, H_2^{(1)}), \ldots, (H_2^{(j)}, H_2^{(j+1)}), \ldots, (H_2^{(m)}, H_2)$ satisfies the assumptions of Lemma 5.2 and condition (5.5), hence obtaining the sequence of implications:

$$y_1 \text{ PE} \Rightarrow H_2^{(1)}u \text{ PE} \Rightarrow \cdots \Rightarrow H_2^{(m)}u \text{ PE} \Rightarrow y_2 \text{ PE}. \qquad \blacksquare$$

Comment 5.1. It is clear that PE properties can only be lost by actual canceling of the effect of certain frequency components present in the input signal. There is also a very simple frequency domain interpretation of Lemma 5.1. If a scalar signal u has a spectrum and if $y = Hu$, with H scalar, stable, and output reachable and with $j\omega$ axis zeros $j\omega_1, \ldots, j\omega_n$, then y being PE is equivalent to the spectral measure of u being supported on at least one point different from $\omega_1, \ldots, \omega_n$ (see, e.g., [AB], [BS], and [BS2]).

6. Time-Varying Systems

6.1. Slow Time Variations

In this section we derive persistency of excitation conditions for time-varying systems whose parameter variations are sufficiently slow. In the first theorem we compare the persistency of excitation for the output of a slowly time-varying system with that of an approximating time-invariant system.

Theorem 6.1. *Consider the time-varying system*

$$\begin{cases} \dot{x}(t) = F(t)x(t) + G(t)u(t), \\ y(t) = H(t)x(t) + J(t)u(t), \end{cases} \tag{6.1}$$

with $x\colon \mathbb{R} \to \mathbb{R}^n$, $u\colon \mathbb{R} \to \mathbb{R}^m$, $y\colon \mathbb{R} \to \mathbb{R}^p$, *and with the following assumptions*:

- A.1. $F(t)$, $G(t)$, $H(t)$, *and* $J(t)$ *are bounded, regulated matrix functions of* t.
- A.2. *There exist* $K > 0$ *and* $a > 0$ *such that* $\|\Phi(t, t_0)\| \leq Ke^{-a(t-t_0)}$ *for all* t, t_0, *where* $\Phi(t, t_0)$ *is the state transition matrix of* $\dot{z}(t) = F(t)z(t)$.
- A.3. *There exist* $\varepsilon > 0$ *and* $T > 0$ *such that, for all* t *and for all* $s, \tau \in (t, t + T)$,

$$\|F(s) - F(\tau)\| < \varepsilon,$$

$$\|G(s) - G(\tau)\| < \varepsilon,$$

$$\|H(s) - H(\tau)\| < \varepsilon,$$

and

$$\|J(s) - J(\tau)\| < \varepsilon. \tag{6.2}$$

- A.4. *There exist* $\alpha_2 > 0$ *and* T_1 *with* $0 < T_1 \leq T$, *and, for all* t_0, *there exists* $\sigma \in (t_0, t_0 + T_1)$ *such that*:
  (i) $\text{Re}\, \lambda_i(F(\sigma)) < -\alpha_2$, $i = 1, \ldots, n$.

(ii) *The output* $\bar{y}(t)$ *of the frozen system*

$$\begin{cases} \dot{\bar{x}}(t) = F(\sigma)\bar{x}(t) + G(\sigma)u(t), & \bar{x}(t_0) = x(t_0), \\ \bar{y}(t) = H(\sigma)\bar{x}(t) + J(\sigma)u(t) \end{cases} \tag{6.3}$$

is PE *over an interval of length* T_1, *i.e., there exist* $\alpha_3 > 0$ *and* $t_- \geq 0$ *such that*

$$\frac{1}{T_1}\int_{t_0}^{t_0+T_1} \bar{y}(\tau)\bar{y}^T(\tau)\,d\tau \geq \alpha_3 I \qquad \text{for all} \quad t_0 \geq t_- > 0. \tag{6.4}$$

Then $y(t)$ *is* PE *for* ε *sufficiently small, i.e., there exists* $\varepsilon_1 > 0$ *such that if* $\varepsilon < \varepsilon_1$, *then there exists* $\alpha_4 > 0$ *such that*

$$\frac{1}{T_1}\int_{t_0}^{t_0+T_1} y(\tau)y^T(\tau)\,d\tau \geq \alpha_4 I \qquad \text{for all} \quad t_0 \geq t_- > 0. \tag{6.5}$$

Proof. Consider an arbitrary $t_0 \geq t_-$ and a $\sigma \in (t_0, t_0 + T_1)$ satisfying A.4. Denote $e(t) = x(t) - \bar{x}(t)$, $t \in (t_0, t_0 + T_1)$. Then

$$e(t) = \int_{t_0}^{t} e^{F(\sigma)(t-\tau)}\{[F(\tau) - F(\sigma)]x(\tau) + [G(\tau) - G(\sigma)]u(\tau)\}\,d\tau.$$

It follows by A.2, A.3, and A.4 that

$$\sup_{t_0 \leq t \leq t_0+T_1} \|e(t)\| \leq K_1\varepsilon\|u\|_\infty \tag{6.6}$$

for some finite $K_1 > 0$. Now

$$y(t) - \bar{y}(t) = H(\sigma)e(t) + [H(t) - H(\sigma)]x(t) + [J(t) - J(\sigma)]u(t).$$

Therefore, uing A.1, A.3, and (6.6)

$$\sup_{t_0 \leq t \leq t_0+T_1} \|y(t) - \bar{y}(t)\| \leq K_2\varepsilon\|u\|_\infty \tag{6.7}$$

for some finite $K_2 > 0$. Now for all $t_0 \geq t_-$

$$\begin{aligned} &\frac{1}{T_1}\int_{t_0}^{t_0+T_1} y(\tau)y^T(\tau)\,d\tau \\ &\geq \frac{1}{2T_1}\int_{t_0}^{t_0+T_1} \bar{y}(\tau)\bar{y}(\tau)^T\,d\tau - \frac{1}{T_1}\int_{t_0}^{t_0+T_1} [y(\tau) - \bar{y}(\tau)][y(\tau) - \bar{y}(\tau)]^T\,d\tau \\ &\geq \left(\frac{\alpha_3}{2} - K_2^2\varepsilon^2\|u\|_\infty\right)^2 I. \end{aligned} \tag{6.8}$$

Let $\varepsilon_1 = \dfrac{(\sqrt{\alpha_3/2})}{K_2\|u\|_\infty}$. Then, for $\varepsilon < \varepsilon_1$, there exists $\alpha_4 > 0$ such that (6.5) holds. ■

Comment 6.1. The main assumption in Theorem 6.1 is A.3 which characterizes the slow time variations. Note also that assumptions A.2 and A.4 are not completely independent given A.3, i.e., given that the system is slowly time-varying: see Chapter 1 of [C].

Our main result for slowly time-varying systems shows that if the input is sufficiently rich of prescribed order and if the variation of the parameters is slow enough, then a regressor vector containing filters of the input and output is persistently exciting provided a uniform output reachability condition is satisfied.

Theorem 6.2. *Consider the time-varying system* (6.1), *and the regression vector*

$$\varphi(t) = \frac{1}{p(s)} Q(s) \begin{bmatrix} y(t) \\ u(t) \end{bmatrix}, \tag{6.9}$$

where $p(s)$ is a polynomial and $Q(s)$ is an $r \times (p + m)$ polynomial matrix, with the following assumptions:

A.1. *$F(t)$, $G(t)$, $H(t)$, and $J(t)$ are bounded, regulated matrix functions of t.*

A.2. *There exist $K > 0$ and $a > 0$ such that $\|\Phi(t, t_0)\| \le K e^{-a(t-t_0)}$ for all t, t_0, where $\Phi(t, t_0)$ is the state transition matrix of $\dot{z}(t) = F(t)z(t)$.*

A.3. *There exist $\varepsilon > 0$ and $T > 0$ such that, for all t and for all $s, \tau \in (t, t + T)$,*

$$\|F(s) - F(\tau)\| < \varepsilon,$$

$$\|G(s) - G(\tau)\| < \varepsilon,$$

$$\|H(s) - H(\tau)\| < \varepsilon,$$

and

$$\|J(s) - J(\tau)\| < \varepsilon.$$

A.4. *$p(s)$ is Hurwitz and* $\deg p(s) \triangleq q \ge \max_{i,j} \deg q_{ij}(s)$.

A.5. *$u(t)$ is bounded and sufficiently rich of order $(n + q + 1, \ldots, n + q + 1/ n + q + 1)$, i.e., there exist t_1, $\alpha_1 > 0$, $\alpha_2 > 0$, and $T_1 > 0$, and, with $T_1 \le T$, such that*

$$\alpha_1 I \le \frac{1}{T_1} \int_t^{t+T_1} \psi(\tau)\psi^T(\tau)\, d\tau \le \alpha_2 I \qquad \text{for} \quad t \ge t_1, \tag{6.10}$$

where $\psi(\tau)$ is defined as in (4.15) *with n_1, n_2, …, n_m and n replaced by $n + q + 1$.*

A.6. *There exist $\alpha_3 > 0$ and $\alpha_4 > 0$, and, for all $t > 0$ and t_0, there exists $\sigma \in (t_0, t_0 + t)$* such that:

(i) $$\operatorname{Re} \lambda_i(F(\sigma)) < -\alpha_3. \tag{6.11}$$

(ii) *With $H(\sigma)[sI - F(\sigma)]^{-1}G(\sigma) + J(\sigma) \triangleq M_\sigma(s) = B_\sigma(s)/a_\sigma(s)$, where $a_\sigma(s)$ is the least common denominator of the entries of $M_\sigma(s)$, the system $\frac{1}{p(s)} Q(s) \begin{bmatrix} M_\sigma(s) \\ I_m \end{bmatrix}$ is uniformly output reachable, i.e.,*

$$c^T Q(s) \begin{bmatrix} B_\sigma(s) \\ a_\sigma(s) I_m \end{bmatrix} \begin{bmatrix} B_\sigma(s) \\ a_\sigma(s) I_m \end{bmatrix}^T Q^T(s) c \ge \alpha_4 \tag{6.12}$$

for all such σ, all s, and all $c \in \mathbb{R}^r$ with $\|c\| = 1$.

Then $\varphi(t)$ *is* PE *for sufficiently small* ε, *i.e., there exist* $\varepsilon^* > 0$, $\alpha_5 > 0$, *and* $T_2 \geq T_1$ *such that*

$$\frac{1}{T_2}\int_t^{t+T_2} \varphi(\tau)\varphi^T(\tau)\,d\tau \geq \alpha_5 I \qquad \textit{for all} \quad t \geq t_1 \quad \textit{and} \quad 0 < \varepsilon < \varepsilon^*. \tag{6.13}$$

Proof.

Step 1. We generate $\bar{\varphi}(\tau)$ from a frozen system and establish an upper bound for $\|\varphi(\tau) - \bar{\varphi}(\tau)\|$. Consider arbitrary t_0 and t with $t_1 \leq t_0$, $0 < t \leq T$ and a $\sigma \in (t_0, t_0 + t)$ satisfying A.6 and let $\bar{y}(t)$ be generated by (6.3) for that fixed σ. Define

$$\bar{\varphi}(t) = \frac{Q(s)}{p(s)}\begin{bmatrix}\bar{y}(\tau)\\ u(\tau)\end{bmatrix}. \tag{6.14}$$

Then, by assumptions A.1–A.4 and the fact that $t \leq T$, it follows that

$$\sup_{t_0 \leq \tau \leq t_0 + t} \|\varphi(\tau) - \bar{\varphi}(\tau)\| \leq \varepsilon C \|u\|_\infty \tag{6.15}$$

for some finite constant C, a function of the system (6.1) and the filter $Q(s)/p(s)$. Notice that C is proportional to the operator gain K/a of the system (6.1).

Step 2. We show that the regression vector of the frozen system is PE. Notice that $\bar{\varphi}(\tau)$ can be written as

$$\bar{\varphi}(\tau) = \frac{1}{p(s)a_\sigma(s)}Q(s)\begin{bmatrix}B_\sigma(s)\\ a_\sigma(s)I_m\end{bmatrix}u(\tau) = \frac{(s+\gamma)^{n+q}}{p(s)a_\sigma(s)}R_\sigma\psi(\tau) \tag{6.16}$$

for some $R_\sigma \in \mathbb{R}^{r\times(m(n+q+1))}$. We can therefore apply Theorem 4.1 to ψ and $\bar{\varphi}$, with the following identifications: $\psi = y_1$, $\bar{\varphi} = y_2$, and $R_\sigma = R$, noting that $\lambda_{\min}(R_\sigma R_\sigma^T) \geq \alpha_4$ for all σ, by A.6(ii). Therefore there exists $\alpha_6 > 0$ such that

$$\frac{1}{T_2}\int_{t_0}^{t_0+T_2} \bar{\varphi}(t)\bar{\varphi}^T(\tau)\,d\tau \geq \alpha_6 I \qquad \textit{for} \quad T_2 > \max(T_1, T^*), \tag{6.17}$$

where

$$T^* = \frac{2\alpha_2 K_2 \|R\|_\infty}{\alpha_1^2\alpha_4} \quad \text{and} \quad \|R\|_\infty = \sup_\sigma \|R_\sigma\| \tag{6.18}$$

and

$$\alpha_6 = \frac{1}{2K_1^2}\left[\alpha_1^2\alpha_4 - \frac{2\alpha_2 K_2\|R\|_\infty}{T_2}\right]. \tag{6.19}$$

Here K_1 and K_2 have the same meaning as in (4.8). They depend on γ, α_1, α_2, $p(s)$, $Q(s)$, and the unknown system via $a_\sigma(s)$ and $B_\sigma(s)$: see (4.4), (4.5), and (3.13), and recall Comment 4.1.

Step 3. By combining the PE property of $\bar{\varphi}(\tau)$ and the closeness of $\varphi(\tau)$ and $\bar{\varphi}(\tau)$, we show that $\varphi(\tau)$ is PE for ε sufficiently small. For $t_0 \geq t_1$, and provided $T_2 \leq T$

(see (6.18) and A.3), we have, using (6.17) and (6.15),

$$\frac{1}{T_2}\int_{t_0}^{t_0+T_2} \varphi(\tau)\varphi^T(\tau)\,d\tau$$

$$\geq \frac{1}{2T_2}\int_{t_0}^{t_0+T_2} \bar{\varphi}(\tau)\bar{\varphi}^T(\tau)\,d\tau - \frac{1}{T_2}\int_{t_0}^{t_0+T_2} [\varphi(\tau)-\bar{\varphi}(\tau)][\varphi(\tau)-\bar{\varphi}(\tau)]^T\,d\tau$$

$$\geq \left(\frac{\alpha_6}{2} - C^2\varepsilon^2\|u\|_\infty^2\right)I. \tag{6.20}$$

Therefore $\varphi(\tau)$ is PE, i.e., (6.13) is satisfied for some $\alpha_5 > 0$ provided in A.3, with $T > \max(T_1, T^*)$, $\varepsilon \leq \varepsilon^*$, where

$$\varepsilon^* = \frac{1}{C\|u\|_\infty}\sqrt{\frac{\alpha_6}{2} - \alpha_2}. \quad \blacksquare \tag{6.21}$$

Comment 6.2. In A.5 we have assumed that $u(t)$ is SR of order $(n+q+1, \ldots, n+q+1/n+q+1)$ rather than $(n_1, \ldots, n_m/n+q+1)$ with different n_i. This is because we have not made any assumption on the structural indices of B_σ in A.6. Of course, if more knowledge about the structure of the time-varying system is available, it can be used to weaken the SR requirement on $u(t)$ in A.5.

Comment 6.3. It follows from Comment 4.1 that α_6 increases in proportion to the input signal power $\|u\|_\infty^2$. Therefore the ratio $\sqrt{\alpha_6}/\|u\|_\infty$ is essentially unaffected by the amplitude of the input signal, while C is a function of the actual system and is proportional to K/a (see A.2). It follows that ε^* in (6.21) can only be increased, within a limited margin, by decreasing the condition number α_2/α_1 in the PE condition on $u(t)$: see (6.10). Some very preliminary results on how to choose $u(t)$ to minimize the condition number of the PE matrix in (6.10) are discussed in [MB].

6.2. Fast Time Variations

The result of Section 6.1 shows that the persistency of excitation condition derived for linear time-invariant systems is robust with respect to sufficiently slow time variations in the system description. We now extend this result to include also the effects of sufficiently fast time variations. (Actually we prove a more general result, concerning the effect of integral small perturbations on PE of which fast time variations are a special case.)

The main result of this section relies on the following lemmas, which discuss the effect of "integral small" perturbations on the dynamics of linear time-varying, stable systems.

Lemma 6.1 [C, p. 6]. *Consider the time-varying, linear system*

$$\dot{x}_1(t) = F(t)x_1(t) \tag{6.22}$$

and the perturbed linear system

$$\dot{x}_2(t) = (F(t) + F_1(t))x_2(t). \tag{6.23}$$

Assume that:

A.1. *Equation* (6.22) *is uniformly asymptotically stable, i.e., there exist* $K \geq 1$ *and* $a > 0$, *such that its transition matrix* $\Phi_1(t, \tau)$ *satisfies*

$$\|\Phi_1(t, \tau)\| \leq Ke^{-a(t-\tau)} \quad \text{for all } t, \tau. \tag{6.24}$$

A.2. $F(t)$ *and* $F_1(t)$ *are uniformly bounded, regulated matrix functions:*

$$\|F\|_\infty = M, \qquad \|F_1\|_\infty = M.$$

A.3. $F_1(t)$ *is integral small, i.e., there exist* $h > 0$ *and* $\delta > 0$ *such that for all* $|t_1 - t_2| \leq h$

$$\left| \int_{t_1}^{t_2} F_1(t)\, dt \right| < \delta.$$

Under these assumptions the perturbed system has a transition matrix $\Phi_2(t, \tau)$ *satisfying*

$$\|\Phi_2(t, \tau)\| \leq K(1 + \delta)e^{-b(t-\tau)} \quad \text{for all } t, \tau,$$

where

$$b = a - \left(3MK\delta + \frac{1}{h}\log_e[(1 + \delta)K]\right). \tag{6.25}$$

Comment 6.4. Lemma 6.1 states that stability is preserved for sufficiently small δ and sufficiently large h, e.g., $\delta \leq a/12MK$ and $h \geq (4 \log_e(1 + \delta)K)/a$ guarantees $b \geq a/2$.

The next lemma considers the effect of an integral small input on the response of an asymptotically stable linear system.

Lemma 6.2. *Consider the time-varying, linear system*

$$\dot{x}(t) = F(t)x(t) + B(t), \qquad x(t_0) = x_0. \tag{6.26}$$

Assume that:

A.1. *The homogeneous system has a transition matrix satisfying* (6.24) *for some* $K \geq 1$ *and* $a > 0$.

A.2. $F(t)$ *and* $B(t)$ *are regulated, bounded matrix functions of* t *with* $\max(\|F\|_\infty, \|B\|_\infty) = M$.

A.3. $B(t)$ *is integral small, i.e., there exist* $h > 0$ *and* $\delta > 0$ *such that for all* $|t_1 - t_2| < h$

$$\left| \int_{t_1}^{t_2} B(t)\, dt \right| < \delta.$$

Under these assumptions the solution of (6.26) *satisfies*

$$\|x(t; t_0, x_0)\| \leq Ke^{-a(t-t_0)}\|x_0\| + C\delta \quad \text{for all } t, t_0$$

where

$$C = K\left(1 + \frac{M}{a}\right)\Big/(1 - e^{-ah}). \tag{6.27}$$

Proof. The solution of (6.26) can be written as

$$x(t; t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)\, d\tau. \tag{6.28}$$

Define

$$C(\tau) = \int_{t_1}^{\tau} B(s)\, ds \quad \text{for some arbitrary } t_1.$$

Integrating by parts, we then have

$$\int_{t_1}^{t_2} \Phi(t, \tau)B(\tau)\, d\tau = \Phi(t, \tau)C(\tau)\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} F(\tau)\Phi(t, \tau)C(\tau)\, d\tau.$$

Using A.3, for all $|t_2 - t_1| \le h$

$$\begin{aligned}\left|\int_{t_1}^{t_2} \Phi(t, \tau)B(\tau)\, d\tau\right| &\le Ke^{-a(t-t_2)}\delta + MK\delta \frac{1}{a} e^{-a(t-t_2)} \\ &\le \delta K\left(1 + \frac{M}{a}\right)e^{-a(t-t_2)}.\end{aligned} \tag{6.29}$$

For the last term in (6.28) we obtain

$$\left|\int_{t_0}^{t} \Phi(t, \tau)B(\tau)\, d\tau\right| \le \sum_{k=0}^{n-1} \left|\int_{t_0+kh}^{t_0+(k+1)h} \Phi(t, \tau)B(\tau)\, d\tau\right| + \left|\int_{t_0+nh}^{t} \Phi(t, \tau)B(\tau)\, d\tau\right|, \tag{6.30}$$

where n is such that

$$t_0 + nh < t \le t_0 + (n + 1)h.$$

Using the estimate (6.29) in (6.30) and using A.1 for the first term in (6.28) yields the desired result. ■

Integral smallness captures a large class of perturbations, e.g., it is preserved under multiplication with a signal having a "band limited" spectrum:

Lemma 6.3. *Provided $v(t)$ and $\dot{v}(t)$ are uniformly bounded,*

$$\|v\|_\infty = V_0, \qquad \|\dot{v}\|_\infty = V_1,$$

and that $B(\tau)$ is integral small, i.e., there exist $h > 0$ and $\delta > 0$ such that for all $|t_1 - t_2| < h$

$$\left|\int_{t_1}^{t_2} B(\tau)\, dt\right| < \delta, \tag{6.31}$$

then $B(\tau)v(\tau)$ is also integral small, and for all $|t_1 - t_2| < h$

$$\left|\int_{t_1}^{t_2} B(\tau)v(\tau)\, d\tau\right| < \delta(V_0 + hV_1). \tag{6.32}$$

Proof. By partial integration we have

$$\left|\int_{t_1}^{t_2} B(t)v(t)\,dt\right| = \left|\int_{t_1}^{t_2} B(t)\,dt\,v(t_2) - \int_{t_1}^{t_2}\int_{t_1}^{t} B(\tau)\,d\tau\,\dot v(t)\,dt\right| \le \delta V_0 + \delta V_1 h. \quad \blacksquare$$

The previous lemmas allow us to establish our main result on preservation of persistency of excitation under integral small perturbations. The next theorem is a counterpart of Theorem 6.1. It establishes that if the output of a time-varying system is PE, then it remains so if the system dynamics and its input are perturbed by integral small perturbations.

Theorem 6.3. *Consider the linear time-varying system*

$$\dot x_1(t) = F(t)x_1(t) + G(t)u(t), \qquad x(t_0) = x_0, \tag{6.33}$$

$$y_1(t) = H(t)x_1(t) + J(t)u(t), \tag{6.34}$$

and the perturbed system

$$\dot x_1(t) = (F(t) + F_1(t))x(t) + G(t)u(t) + G_1(t), \qquad x(t_0) = x_0, \tag{6.35}$$

$$y(t) = H(t)x(t) + J(t)u(t). \tag{6.36}$$

Assume that:

A.1. *$F(t)$, $F_1(t)$, $G(t)$, $G_1(t)$, $H(t)$, and $J(t)$ are bounded, regulated matrix functions of t:*

$$\max(\|F\|_\infty, \|F_1\|_\infty, \|G\|_\infty, \|G_1\|_\infty, \|H\|_\infty, \|J\|_\infty) = M$$

and $u(t)$ is bounded: $\|u\|_\infty = U$.

A.2. *The transition matrix $\Phi(t,\tau)$ for $\dot\zeta(t) = F(t)\zeta(t)$ is exponentially stable, i.e., it satisfies* (6.24) *for some $K \ge 1$ and $a > 0$.*

A.3. *$y_1(t)$ is* PE, *i.e., there exist α_1, β_1, $T > 0$, and t_1 such that*

$$\alpha_1 I \le \frac{1}{T}\int_t^{t+T} y_1(\tau)y_1(\tau)^T\,d\tau \le \beta_1 I \qquad \textit{for all} \quad t \ge t_1.$$

A.4. *$F_1(t)$ and $G_1(t)$ are integral small, i.e., there exist δ, $h > 0$ such that for all $|\tau_1 - \tau_2| \le h$*

$$\left|\int_{\tau_1}^{\tau_2} F_1(\tau)\,d\tau\right| < \delta,$$

$$\left|\int_{\tau_1}^{\tau_2} G_1(\tau)\,d\tau\right| < \delta.$$

There exist positive constants δ_1^*, δ_2^*, and h^* depending on the system and on α_1 such that if A.4 holds with $h \ge h^*$ and $0 \le \delta < \min(\delta_1^*, (h^*/h)\delta_2^*)$ then the perturbed system is BIBS *stable and the output $y(t)$ is* PE.

Proof. We first establish the existence of positive constants δ_1, h^*, and C_1 and a time t_3 such that

$$\|y(t) - y_1(t)\| \le C_1\delta \qquad \text{for all} \quad t \ge t_3 \tag{6.37}$$

provided A.4 holds with $\delta \le \delta_1^*$ and $h \ge h^*$. The constants δ_1 and h^* only depend on the system parameters, and C_1 depends on the system parameters and h. From this it follows that for all c with $\|c\| = 1$ and for all $t \ge t_3$

$$\begin{aligned}\frac{1}{T}\int_t^{t+T} (c^T y(\tau))^2\, d\tau &\ge \frac{1}{T}\int_t^{t+T} (c^T y_1(\tau))^2\, d\tau - \sup_{t\ge t_3} \|y(t) - y_1(t)\|^2 \\ &\ge \alpha_1 - C_1^2\delta^2 \\ &\ge \frac{\alpha_1}{2}.\end{aligned}$$

The last inequality holds for $0 \le \delta \le \min(\delta_1^*, (\sqrt{\alpha_1/2})\cdot(1/C_1(h)))$, $h \ge h^*$.

We now establish (6.37). With $e(t) \triangleq x(t) - x_1(t)$, we have

$$\dot{e}(t) = (F(t) + F_1(t))e(t) + F_1(t)x_1(t) + G_1(t), \qquad e(t_0) = 0.$$

From assumptions A.1 and A.2 it follows that

$$\|x_1(t)\| \le Ke^{-a(t-t_0)}\|x_1(t_0)\| + \frac{KM}{a}U \quad \text{for all } t \text{ and } t_0.$$

Choosing $t \ge t_2$, $t_2 \triangleq t_0(\log_e(MU/a\|x_1(t_0)\|)/a)$, and using (6.33) we have

$$\|x_1(t)\| \le \frac{2KM}{a}U \qquad \text{for all} \quad t \ge t_2,$$

$$\|\dot{x}_1(t)\| \le \left(\frac{2KM}{a} + 1\right)MU \qquad \text{for all} \quad t \ge t_2.$$

From Lemma 6.3 (equation (6.3.2)) applied to $F_1(t)x_1(t)$ we obtain that $F_1(t)x_1(t)$ is integral small, i.e., for all $t, s \ge t_2$ with $|t - s| \le h$,

$$\left|\int_s^t F_1(\tau)x_1(\tau)\, d\tau\right| \le \delta\left(\frac{2KM}{a}U + \left(2\frac{KM}{a} + 1\right)MUh\right).$$

Define

$$h^* = 4\log_e\left[\left(1 + \frac{a}{12MK}\right)K\right]\Big/ a,$$

$$\delta_1^* = a/(12MK).$$

This guarantees that the perturbed system (6.35) is BIBS stable for $\delta \le \delta_1^*$ and $h \ge h^*$ as explained in Comment 6.4.

Now use Lemmas 6.1 and 6.2 to obtain

$$\begin{aligned}\|e(t)\| \le K(1 + \delta)e^{-(a/2)(t-t_2)}\|e(t_2)\| + \left[K(1 + \delta)\left(1 + \frac{2M'}{a}\right)\Big/(1 - e^{-(a/2)h})\right]\delta \\ \times\left[1 + UM\left(\frac{2K}{a} + \left(\frac{2KM}{a} + 1\right)h\right)\right] \qquad \text{for} \quad t \ge t_2,\end{aligned}$$

with $M' = \max(2M, M(1 + (2KM/a)U))$. Define

$$t_3 = t_2 - \frac{2}{a}\log_e \frac{(1 + 2M'/a)(1 + UM(2K/a + (2KM/a + 1)h))\delta}{(1 - e^{-(a/2)h})\|e(t_2)\|},$$

$$C_3(h) = 2\frac{K(1 + \delta_1^*)(1 + 2M'/a)(1 + UM(2K/a + (2KM/a + 1)h))}{(1 - e^{-(ah^*/2)})}.$$

We then have, for all $t \geq t_3$,

$$\|e(t)\| \leq C_3(h)\delta.$$

This establishes (6.37) by assumption A.1, with $C_1(h) = MC_3(h)$. The result then follows by identifying δ_2^* as

$$\delta_2^* = \inf_{h \geq h^*} \sqrt{\frac{\alpha_1}{2}} \frac{h}{C_1(h)} \frac{1}{h^*} = \sqrt{\frac{\alpha_1}{2}} \frac{1}{C_1(h^*)}. \quad \blacksquare \tag{6.38}$$

Comment 6.5. Assumption A.3 stating that the output of the time-varying system (6.33)–(6.34) is PE might be established using Theorem 6.2. Of course, the unperturbed system (6.33)–(6.34) could be time-invariant as well and then any PE criterion (e.g., those of Sections 4 or 5) could be used to guarantee A.3.

Combining Theorems 6.3 and Theorem 6.2 we deduce that the PE property is robust with respect to the (combined) effect of slow time variations and integral small perturbations in the description of the dynamics of the underlying systems/filters. A corollary in the style of Theorem 6.2, including both slow time variations and integral small perturbations is obvious.

Comment 6.6. Notice that given the stability properties of the unperturbed system as well as α_1, a lower bound for the minimum eigenvalue in the PE condition on the output of the unperturbed system, we can compute the δ_1^*, δ_2^*, h^* characterizing how much the system can be perturbed without losing stability or PE of the output. On the other hand, given δ_1^*, δ_2^*, h^* it is not always possible to preserve PE after perturbation, as δ_2^* (see (6.38)) varies in inverse proportion to the condition number of the excitation, α_1/β_1 (notice that $\beta_1 = O(((KM/a)U)^2)$! Recall Comment 6.1.

Comment 6.7. We comment upon the class of integral small perturbations. As pointed out in [C, p. 8], a subclass of the integral small perturbations are zero mean, fast time-varying, almost-periodic perturbations. Indeed, it is not difficult to show that for any $F(t)$ almost-periodic, zero mean, and for any positive δ and h there exists an $\omega_c = \omega(h, \delta)$ such that for all $\omega > \omega_c$ the signal $F(\omega t)$ is integral small, i.e.,

$$\left|\int_{t_1}^{t_2} F(\omega t)\,dt\right| < \delta \qquad \text{for all} \quad |t_1 - t_2| < h.$$

Comment 6.8. The meaning of slow and integral small signals can be deduced from the bounds (6.21) and (6.38), respectively. For slow signals of the form $\sin \omega t$, ε in A.3 of Theorem 6.2 is of the order of ω. Recalling that $C \sim K/a$, where K/a is the gain of the unperturbed system, it follows that (6.21) requires that $\omega \ll a$. This is completely in agreement with intuition; slow means frequencies well within the pass band of the system. For integral small signals of the form $M \sin \omega t$, δ and h in A.4 of Theorem 6.3 are of the order M/ω and 1, respectively. Bound (6.38) on δ requires then that $\omega a \gg M^2$. Hence the larger the perturbations are in magnitude the faster these should be. It is clear from this that integral small perturbations capture the significant effects of fast perturbations.

7. Conclusions

A major stumbling block in establishing exponential convergence of adaptive estimation or control schemes is often in demonstrating the persistency of excitation of a regression vector somewhere in the adaptive loop given that a dynamically related regression vector is PE or, better still, that an input signal (or reference signal) is sufficiently rich. We believe that this paper significantly contributes to the removal of this stumbling block in two major ways.

First we have produced conditions under which the output (or state) of a time-invariant MIMO system is PE when it is dynamically related to the output (or state) of another MIMO system. Defining a "sufficiently rich" vector input signal in terms of the persistency of excitation of a basis vector (output or state of a first MIMO system), and realizing that most commonly used regression vectors can be described as the output (or state) of a MIMO system that is dynamically related to the first one in a way that satisfies our conditions, we thereby solve the problem of establishing PE conditions for a large class of regression vectors arising in time-invariant MIMO systems.

Our other major contribution is to extend these results to a large class of MIMO, linear, time-varying systems, where the time variations must be either sufficiently slow or integral small, or a combination of these. We have observed that sufficiently fast time variations are contained in the class of integral small ones. We have also given a rough quantitative description of the required frequency separation, in that the slow time variations must be much slower than the dominant (i.e., slowest) frequencies of the system, while the fast time variations must be much faster than these. We finally note that by combining the results of Sections 6.1 and 6.2, we have produced conditions on the time variations of a time-varying system (containing both slow and fast parameter variations) that will guarantee the persistency of excitation of a regressor derived from that system provided the input vector is sufficiently rich in a well-defined sense.

References

[A] B. D. O. Anderson, Exponential stability of linear equations arising in adaptive identification, *IEEE Trans. Automat. Control*, **22** (1977), 83–88.

[AB] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, Cambridge, MA, 1986.

[AJ] B. D. O. Anderson and C. R. Johnson, Jr., Exponential convergence of adaptive identification and control algorithms, *Automatica*, **18** (1982), 1–13.

[BG] G. Bastin and M. Gevers, Stable adaptive observers for nonlinear time-varying systems, *IEEE Trans. Automat. Control*, 33 (7) (1988) (to appear).

[BS] E. W. Bai and S. Sastry, Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control, *Systems Control Lett.*, **6** (1985), 153–163.

[BS1] S. Boyd and S. Sastry, On parameter convergence in adaptive control, *Systems Control Lett.*, 3 (1983), 311–319.

[BS2] S. Boyd and S. Sastry, Necessary and sufficient conditions for parameter convergence in adaptive control, *Proceedings of the American Control Conference*, San Diego, 1984, pp. 1584–1588.

[C] W. A. Coppel, *Dichotomies in Stability Theory*, Lecture Notes in Mathematics, Vol. 629, Springer-Verlag, Berlin, 1978.

[D] S. Dasgupta, Adaptive Identification and Control, Ph.D. Dissertation, Australian National University, 1984.

[DAT] S. Dasgupta, B. D. O. Anderson, and A. C. Tsoi, Input conditions for continuous time adaptive system problems, *Proceedings of the 22nd IEEE Conference on Decision and Control*, San Antonio, Tx, 1983, pp. 211–216.

[GM] M. Green and J. B. Moore, Persistence of excitation in linear systems, *Proceedings of the American Control Conference*, Boston, 1984, pp. 412–417; *Systems Control Lett.*, 7 (1986), 351–360.

[GT] G. C. Goodwin and E. K. Teoh, Persistency of excitation in the presence of possibly unbounded signals, *IEEE Trans. Automat. Control*, **30** (1985), 595–597.

[IK] P. A. Ioannou and P. V. Kokotovic, *Adaptive Systems with Reduced Models*, Lecture Notes in Control and Information Sciences, Vol. 47, Springer-Verlag, Berlin, 1983.

[JA] R. M. Johnstone and B. D. O. Anderson, Exponential convergence of recursive least squares with exponential forgetting factor–adaptive control, *Systems Control Lett.*, **2** (1982), 69–76.

[K] T. Kailath, *Linear Systems*, p. 657, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[L] L. Ljung, Characterization of the Concept of Persistently Exciting Inputs in the Frequency Domain, Technical Report 7119, Department of Automatic Control, Lund Institute of Technology, 1971.

[MB] I. M. Y. Mareels, R. R. Bitmead, M. Gevers, C. R. Johnson, Jr., R. L. Kosut, and M. A. Poubelle, How exciting can a signal really be?, *Systems Control Lett.*, **8** (1987) 197–204.

[M] J. B. Moore, Persistence of excitation in extended least squares, *IEEE Trans. Automat. Control*, **28** (1983), 60–68.

[MN] A. P. Morgan and K. S. Narendra, On the stability of non-autonomous differential equations $\dot{x} = [A + B(t)]x$ with skew-symmetric matrix $B(t)$, *SIAM J. Control Optim.*, **15** (1977), 163–176.