[10] G. C. Goodwin and R. S. Long, "Generalization of results on multivariable adaptive control," *IEEE Trans. Automat. Contr.*, vol. AC-25, Dec. 1980.
[11] H. Elliott and W. A. Wolovich, "A parameter adaptive control structure for linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 340–351, Apr. 1982.
[12] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley-Interscience, 1972.

On the Necessity of "Block Invariance" for the Convergence of Adaptive Pole-Placement Algorithm with Persistently Exciting Input

NAHUM SHIMKIN AND ARIE FEUER

Abstract—**Global convergence of a number of discrete-time adaptive control algorithms with persistently exciting inputs has been established in the literature. A key element of the convergence proof is the "block-invariance" feature of the algorithm. Namely, while the parameters are estimated at every sample, the controller parameters remain invariant throughout a "block" of samples and are updated only at the end of each such block. The necessity of this "block-invariance" feature to guarantee convergence is established in this note by a counterexample.**

I. Introduction

Persistency of excitation (PE) plays a key role in the convergence of some adaptive control schemes. The central problem in the analysis of these schemes is the PE property of the input to the controlled system. This input consists of an external input term which can be guaranteed to be PE and a time-varying feedback term which may destroy this property for the system input.

For certain adaptive control algorithms, closed-loop stability can be established without recourse to parameter convergence. It has been shown that in this case, the PE of the external and system inputs are equivalent [8], [9]. In general, however, signal boundedness cannot be assumed *a priori*. Then, to ensure the PE property for the system input the following feature, sometimes referred to as "block-invariance," was proposed. While the parameter estimates in these schemes are updated every sample unit, the controller parameters are updated once every "block" of samples and kept constant in between. It has been established that with a PE external input, this feature, with a properly chosen block size, will guarantee the PE property for the system input. This, in turn, guarantees global parameter convergence and stability of the adaptive scheme. The block-invariance approach is applicable to various schemes but has been mostly associated with adaptive pole placement algorithms (e.g., [1], [3], [4], [10], [11]).

In light of these results a key question remained open: is the block-invariance feature necessary for convergence?

In this note, we provide a definite affirmative answer for the direct adaptive pole-placement algorithm by presenting a counterexample. The example is for a simple case but is clearly sufficient as a proof of the "block-invariance" necessity for global convergence. To support our analysis, we have run some simulations which are presented here.

II. The Adaptive System Equations

Consider the following discrete-time, first-order system

$$y(t+1)=ay(t)+bu(t) \tag{2.1}$$

Manuscript received May 20, 1987.
The authors are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, 32000, Israel.
IEEE Log Number 8718569.

where

$$a>1.$$

We assume that the value of a is known while the value of b is *unknown*. The control law we want to design should place the closed loop pole at a^* (with $|a^*|<1$). The nominal control law required to accomplish this is given by

$$u^*(t)=\left[\frac{1}{b}\right]\Delta ay(t)+v(t) \tag{2.2}$$

where $v(t)$ is the reference input and $\Delta a = a^* - a$. Simple substitution of (2.2) in (2.1) will confirm the desired result:

$$y(t+1)=a^*y(t)+bv(t). \tag{2.3}$$

However, since b is not known, the control cannot be generated by (2.2). A direct adaptive controller will be used to accomplish the desired pole-placement.

To get the adaptive control, let us first rewrite (2.1) in the form

$$u(t)=\theta^*\phi(t) \tag{2.4}$$

where

$$\theta^*=\frac{1}{b}$$
$$\phi(t)=y(t+1)-ay(t). \tag{2.5}$$

Based on (2.4), a recursive least-squares (RLS) algorithm can be used to estimate θ^*; its equations are given (see, e.g., [1]) by

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{P(t-1)\phi(t-1)[u(t-1)-\phi(t-1)\hat{\theta}(t-1)]}{1+\phi(t-1)^2P(t-1)} \tag{2.6}$$

$$P(t)=P(t-1)-\frac{P(t-1)^2\phi(t-1)^2}{1+\phi(t-1)^2P(t-1)} \tag{2.7}$$

with initial conditions $P(0)=P_0$, $\hat{\theta}(0)=\hat{\theta}_0$.

Two possible ways of using the estimate $\hat{\theta}(t)$ to generate the control signal have been considered in the literature. The first,

$$u(t)=\hat{\theta}(t-1)\Delta ay(t)+v(t) \tag{2.8}$$

and the second (termed sometimes "block-invariant feedback")

$$\left.\begin{aligned} &u(t)=\hat{\rho}(t-1)\Delta ay(t)+v(t)\\ &\text{where}\\ &\hat{\rho}(t)=\begin{cases}\hat{\theta}(t) \text{ for } t=kN,\ k=0,1,2,\cdots\\ \hat{\rho}(t-1) \text{ otherwise}\end{cases}\end{aligned}\right\} \tag{2.9}$$

and N an integer to be defined in the sequel.

While the first control law (2.8) is a natural substitute for the nominal control (2.2), the second choice has been considered for the simple reason that convergence proofs could be derived for it.

We will next outline the convergence results which exist for the above block-invariant feedback control.

III. Convergence with Block-Invariant Feedback

Using existing results (see, e.g., [2], [3], [4], [11]), the following convergence result can be proved for the system we consider.

Lemma 3.1: Consider system (2.1), estimation algorithm (2.6), (2.7), and control (2.9) with $N \geq 4$. Assume that the input $v(t)$ is bounded and satisfies the persistency of excitation (PE) condition:

C1): There exist positive ϵ, t_0 such that for all $t \geq t_0$

$$\sum_{j=t+1}^{t+(N-2)} \begin{bmatrix} v(j+1) \\ v(j) \end{bmatrix} [v(j+1),\ v(j)] \geq \epsilon I. \tag{3.1}$$

Then *for all initial* conditions the parameter $\hat{\theta}(t)$ converges to $\theta^* = 1/b$ and all signals are bounded.

Proof Outline:

We begin by constructing the associated signal system (ASS) for our estimation problem.

$$\begin{bmatrix} y(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y(t) \\ y(t-1) \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} u(t) \tag{3.2a}$$

$$\phi(t-1) = [1,\ -a] \begin{bmatrix} y(t) \\ y(t-1) \end{bmatrix}. \tag{3.2b}$$

Using PE results (e.g., [4], [6]) for LTI systems it then follows that there exists a $\delta > 0$ such that

$$\sum_{t=kN+1}^{(k+1)N} \phi(t)^2 > \delta. \tag{3.3}$$

The key in establishing (3.3) is the fact that because of the way $u(t)$ is defined in (2.9) it can be viewed as a constant state feedback control for the ASS system (3.2) for $t \in [kN+1, (k+1)N]$.

Equation (3.3), in turn, is a sufficient condition for the convergence of algorithm (2.6), (2.7) to the correct parameter. Signal boundedness then follows readily. □

IV. The Necessity of the Block Invariance

With the convergence result established in Section III, an obvious question arises: is the "block invariance" necessary? The main result of our note is the answer to this question.

The control law we investigate is given in (2.8). We first note (see [5]) that the sequence $\{\hat{\theta}(t)\}$ generated by (2.6) and (2.7) converges, for any $u(t)$, to a constant. Denoting this constant by $\hat{\theta}_\infty$, we may rewrite (2.8) in the form

$$u(t) = \hat{\theta}_\infty \cdot \Delta a \cdot y(t) + v(t) + [\hat{\theta}(t-1) - \hat{\theta}_\infty]\Delta a \cdot y(t). \tag{4.1}$$

In the framework of the ASS (3.2), this can be viewed as the sum of a constant feedback term, the reference input and a time-varying feedback term with diminishing gain. The problem is how to deal with the time-varying term. If one could assure that $\{y(t)\}$ is bounded, this term would decay to zero and, for large enough t, $u(t)$ would become PE. However, for an unbounded $\{y(t)\}$ the time-varying term can potentially prevent $u(t)$ from being PE even if $v(t)$ is. The example we present here shows that this does in fact happen, in which case the parameter estimate does not converge to its correct value.

Consider the system which consists of (2.1), (2.5), (2.6), (2.7), and (2.8):

$$y(t) = ay(t-1) + bu(t-1) \tag{4.1a}$$

$$u(t) = \hat{\theta}(t-1) \cdot \Delta a \cdot y(t) + v(t) \tag{4.1b}$$

$$\phi(t-1) = y(t) - ay(t-1) \tag{4.1c}$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{P(t-1)\phi(t-1)[u(t-1) - \phi(t-1)\hat{\theta}(t-1)]}{1 + \phi(t-1)^2 P(t-1)} \tag{4.1d}$$

$$P(t) = P(t-1) - \frac{P(t-1)^2 \phi(t-1)}{1 + \phi(t-1)^2 P(t-1)}. \tag{4.1e}$$

This is clearly a nonlinear system with $v(t)$ as its input. Now, we state the following.

Proposition 4.1: Let

$$v(t) = \omega(t) - \Delta a \cdot \varphi(t-1) z(t) \qquad (t \geq 1) \tag{4.2}$$

where

$$z(t) = a^{t+1} - (a^2+\gamma^2)^{-1} \gamma^{t+1} \left(\gamma \cos \frac{\pi t}{2} + a \sin \frac{\pi t}{2} \right) \tag{4.3a}$$

$$\omega(t) = b^{-1} \gamma^{t+1} \sin \frac{\pi t}{2} \tag{4.3b}$$

$$\varphi(t) = Q(t) \left[\frac{b^{-1}}{1-a^2} - b^{-1} \right] + b^{-1} \tag{4.3c}$$

$$Q(t) = \left[1 + \sum_{k=0}^{t-1} \left(\gamma^{k+1} \sin \frac{\pi k}{2} \right)^2 \right]^{-1} \tag{4.3d}$$

$$\gamma = a^{-1/2}. \tag{4.3e}$$

Then,

$$y(t) = z(t) \tag{4.4a}$$

$$u(t) = \omega(t) \tag{4.4b}$$

$$\hat{\theta}(t) = \varphi(t) \tag{4.4c}$$

$$P(t) = Q(t) \tag{4.4d}$$

is the corresponding state of (4.1) with initial conditions

$$y(0) = a - (a^2+\gamma^2)^{-1}\gamma^2, \qquad u(0) = 0$$

$$\hat{\theta}(0) = \frac{b^{-1}}{1-a^2}, \qquad P(0) = 1. \tag{4.5}$$

The proof of this proposition is accomplished by verifying that (4.4) does satisfy (4.1) with $v(t)$ given by (4.2) and (4.3). This had been done in the Appendix for the skeptical reader.

Let us now observe the significance of the particular solution presented in (4.4). First, we concentrate on $v(t)$. Since we chose $a < 1$ it follows from (4.3e) that $0 < \gamma < 1$ and $\gamma^2 a = 1$. This can be used in (4.2) to show that

$$v(t) = b^{-1} \Delta a \left[\frac{a+1}{2} + \frac{a-1}{2} \cos \pi t \right] + \epsilon(t) \tag{4.6}$$

where $\epsilon(t)$ is a linear combination of exponentially decaying terms. The details are again differed to the Appendix.

Two properties of $v(t)$ can be observed from (4.6). 1) $v(t)$ is *bounded,* and 2) $v(t)$ is *persistently exciting (PE),* namely, satisfies (3.1). The last property follows from results in the literature (e.g., [6]) and the fact that $v(t)$ has two spectral lines at $\omega = 0$ and $\omega = \pi$. Next, we turn our attention to the solution (4.4) corresponding to the particular initial conditions and input. From (4.4) and (A.3) in the Appendix it follows readily that

$$\lim_{t\to\infty} y(t) = \infty \tag{4.7a}$$

$$\lim_{t\to\infty} P(t) = \frac{a-1}{a^2} \tag{4.7b}$$

$$\lim_{t\to\infty} \hat{\theta}(t) = 0 \tag{4.7c}$$

$$\lim_{t\to\infty} u(t) = 0. \tag{4.7d}$$

This means that despite the two observed properties of $v(t)$ the adaptive system is neither globally stable nor does it converge globally to the correct parameter. In light of the results regarding Nussbaum's gain (see, e.g., [7]), it is not surprising that the boundedness of $v(t)$ does not guarantee stability. However, the fact that the PE property of $v(t)$ does not guarantee convergence of the parameter estimation algorithm is new. The conclusion we draw from the example considered here is that *"block invariance" is necessary to guarantee global convergence in direct adaptive pole placement with PE input.*

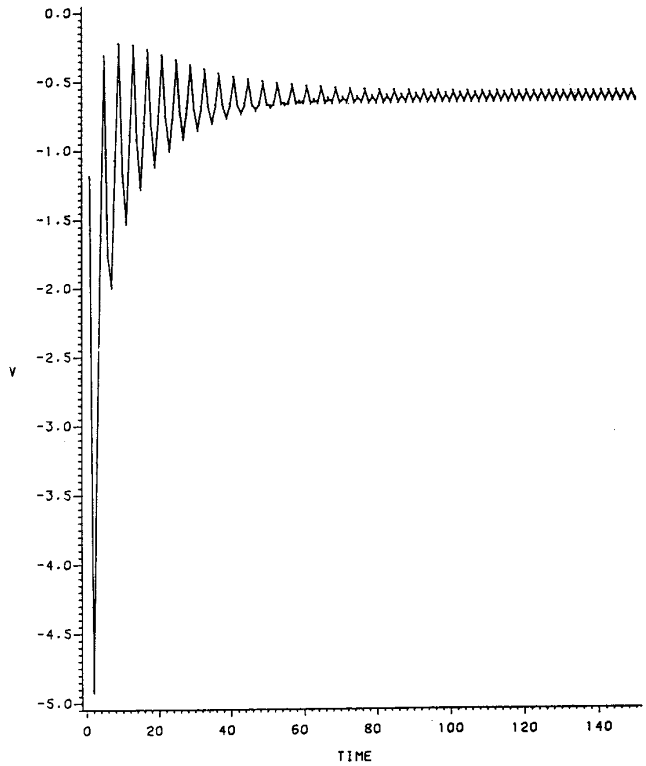

Fig. 1. External input (simple RLS).

Remark: The example was constructed by first choosing $y(t)$ and then calculating the remaining variable accordingly. By replacing the choice in (4.4) by $y(t) = a^{t+1} - \gamma^{t+1} g(t)$ where $g(t)$ is any sum of sinusoids, the resulting $v(t)$ will asymptotically resemble $g(t)^2$. Hence, $v(t)$ can be made as "rich" as one desires (namely, with arbitrarily many spectral lines) and still not guarantee convergence. Similarly, examples can be constructed for more complicated systems (higher dimension) and algorithms.

Replacing the simple RLS in the system we considered by "RLS with covariance resetting" results in a new adaptive system. A similar counterexample has been constructed for this system too—its simulation results are presented in the next section.

V. Simulation Results

To verify the results of the previous sections, we have simulated the adaptive system on a digital computer. The results with the choice $a = 1.1$, $a^* = 0.5$ and $b = 1$ are presented in the sequel. In Fig. 1, we see the shape of $v(t)$, which clearly agrees with (4.6). In Figs. 2 and 3, we show the results of four experiments. In the first one, the nominal initial conditions of Proposition 4.1 were used and the results clearly agree with our analysis. The second experiment, again with the nominal initial conditions, but this time with "block invariance" adaptive control. This time the output remains bounded and the parameter converges to its correct value as expected. In the other two experiments, we have changed the initial condition for the parameter estimation by ± 1 percent. With this change and no block-invariance, we still got convergence and stability. This indicates that the instability occurs at singular points and the algorithm may otherwise be well behaved.

The four experiments were repeated for RLS with covariance resetting. We had to recompute the external input and initial conditions to get the desired results which are presented in Figs. 4, 5, and 6 and are similar to those for the simple RLS.

In conclusion, we have shown through an example that the "block invariance" feature is necessary for global convergence. A very simple

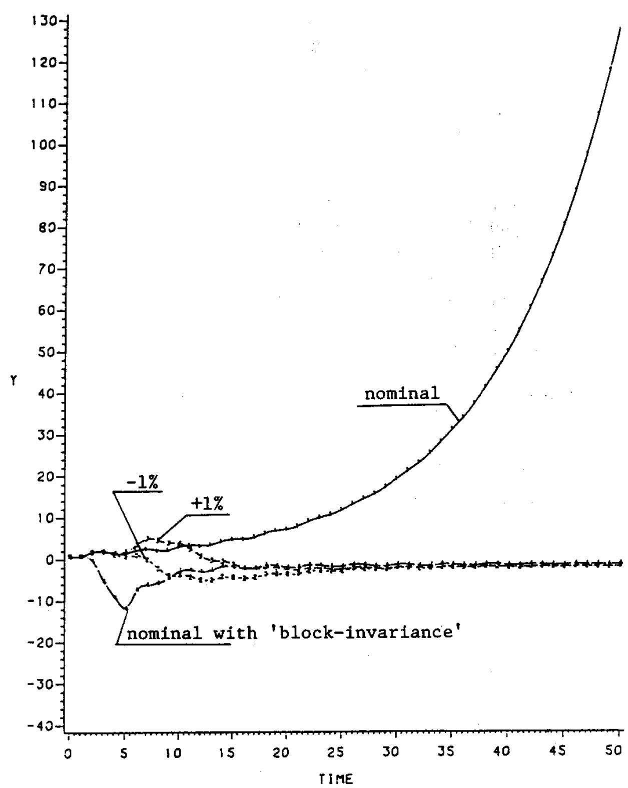

Fig. 2. Output behavior (simple RLS).

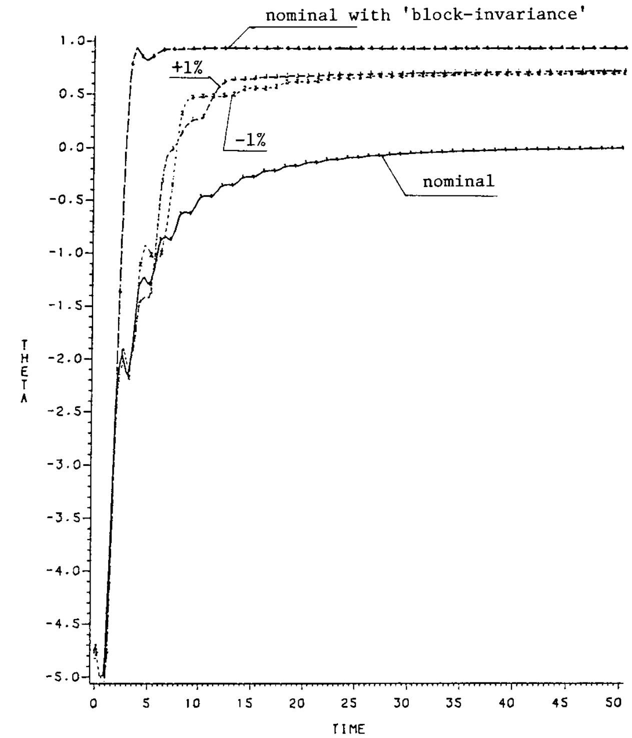

Fig. 3. Estimated parameter (simple RLS).

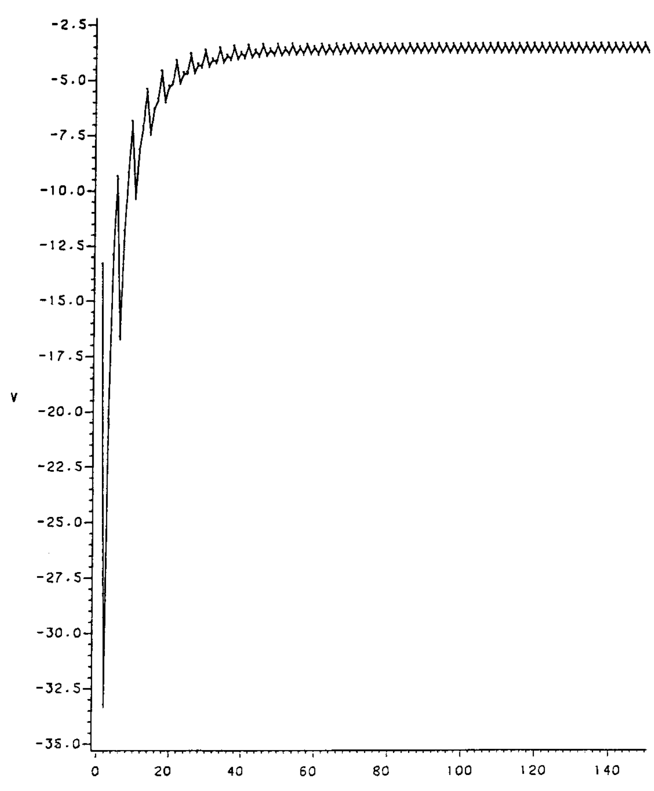

Fig. 4. External input (RLS with covariance resetting).

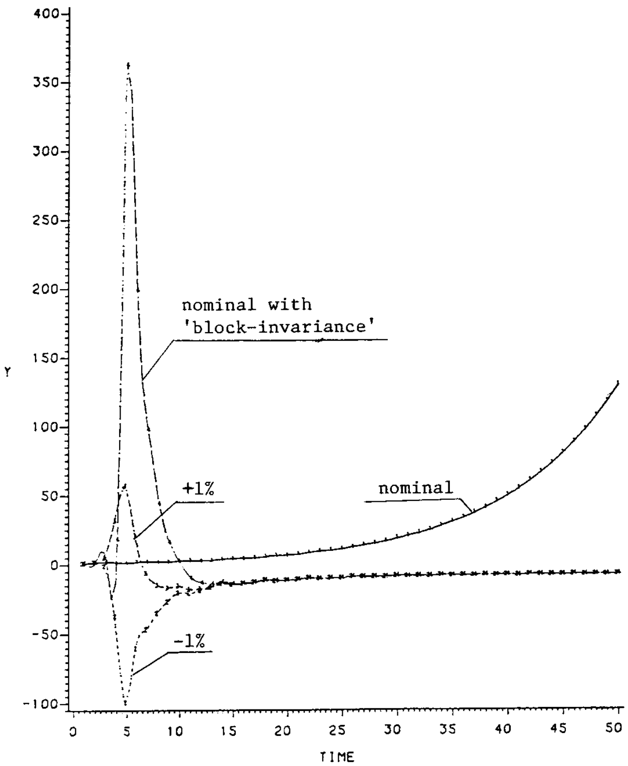

Fig. 5. Output (RLS with covariance resetting).

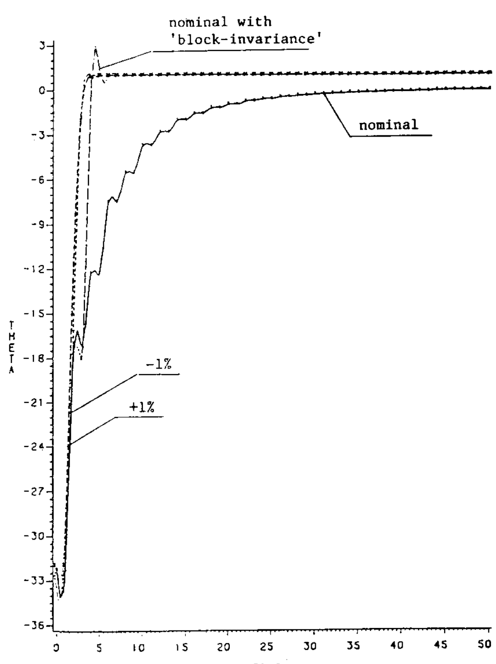

Fig. 6. Estimated parameter (RLS with covariance resetting).

control algorithm was used. It is, however, representative of most direct adaptive schemes. As for the estimation algorithm, both the "simple RLS" and the "RLS with covariance resetting" were considered, with similar results.

APPENDIX

A. Proof of Proposition 4.1:

To prove the proposition, we substitute the state as given by (4.4) at time $(t-1)$ on the right-hand side of (4.1) to get the same state shifted by one time unit to t. We begin with (4.1a):

$$\begin{aligned}
&ay(t-1)+bu(t-1)\\
&\quad=az(t-1)+b\omega(t-1)\\
&\quad=a\cdot a^t-a(a^2+\gamma^2)^{-1}\gamma^t\left[\gamma\cos\left(\frac{\pi t}{2}-\frac{\pi}{2}\right)+a\sin\left(\frac{\pi t}{2}-\frac{\pi}{2}\right)\right]\\
&\qquad+\gamma^t\sin\left(\frac{\pi t}{2}-\frac{\pi}{2}\right)\\
&\quad=a^{t+1}-(a^2+\gamma^2)^{-1}\gamma^{t+1}\left(\gamma\cos\frac{\pi t}{2}+a\sin\frac{\pi t}{2}\right)\\
&\quad=z(t)\\
&\quad=y(t).
\end{aligned}$$

Next, (4.1b):

$$\begin{aligned}
\hat{\theta}(t-1)\Delta ay(t)+v(t)&=\varphi(t-1)\Delta az(t)+\omega(t)-\varphi(t-1)\Delta az(t)\\
&=\omega(t)\\
&=u(t).
\end{aligned}$$

Notice now that using (4.1a), (4.1d) and (4.1e) can be rewritten in the form (see [1])

$$\hat{\theta}(t)=P(t)P(0)^{-1}(\hat{\theta}(0)-b^{-1})+b^{-1} \tag{A.1}$$

$$P(t)^{-1}=P(t-1)^{-1}+\phi(t-1)^2. \tag{A.2}$$

Substitution of (4.4d) and (4.5) in (A.1) will immediately verify (4.4c). Also, since by (4.1a), (4.1c), (4.3b), and (4.4b), we have

$$\begin{aligned}
\phi(t-1)&=bu(t-1)\\
&=b\omega(t-1)\\
&=\gamma^t\sin\frac{\pi(t-1)}{2};
\end{aligned}$$

hence,

$$\begin{aligned}
P(t-1)^{-1}+\phi(t-1)^2&=Q(t-1)^{-1}+\phi(t-1)^2\\
&=1+\sum_{k=0}^{t-2}\left(\gamma^{k+1}\sin\frac{\pi k}{2}\right)^2+\left(\gamma^t\sin\frac{\pi(t-1)}{2}\right)^2\\
&=Q(t)^{-1}\\
&=P(t)^{-1}
\end{aligned}$$

and (A.2) is verified and so is (4.4d). □

B. Derivation of (4.6):

We first evaluate the series in (4.3d) and substitute $\gamma^2=a^{-1}$. This gives

$$Q(t-1)=\frac{a^2-1}{a^2}\,[1-a^t\alpha(t)]^{-1} \tag{A.3}$$

where

$$\alpha(t)=a^{-1}\left[a\cos^2\frac{\pi t}{2}+\sin^2\frac{\pi t}{2}\right]$$
$$=a^{-1}\left[\frac{a+1}{2}+\frac{a-1}{2}\cos\pi t\right].$$

Therefore, from (4.3c)

$$\varphi(t-1)=b^{-1}\frac{a^2}{1-a^2}Q(t-1)+b^{-1}$$
$$=-b^{-1}\frac{a^{-t}\alpha(t)}{1-a^{-t}\alpha(t)}.$$

Substituting in (4.2) and recalling that $a > 1$, $0 < \gamma < 1$ gives immediately

$$v(t)=b^{-1}a^{-t}\alpha(t)\Delta a a^{t+1}+\epsilon(t)$$
$$=b^{-1}\Delta a\cdot a\cdot\alpha(t)+\epsilon(t)$$

with $\epsilon(t)$ a linear combination of exponentially decaying terms. □

REFERENCES

[1] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.
[2] H. Elliott, "Direct adaptive pole placement with application to non-minimum phase systems," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 720–722, 1982.
[3] 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.
[4] M. Heymann, "Persistency of excitation results for structured nonminimal models," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 112–116, 1988.
[5] R. L. Leal and G. C. Goodwin, "A globally convergent adaptive pole placement algorithm without a persistency of excitation requirement," in *Proc. 23rd Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984.
[6] I. Mareels, "Sufficiency of excitation," *Syst. Contr. Lett.*, vol. 5, pp. 159–163, 1984.
[7] R. D. Nussbaum, "Some remarks on a conjecture in adaptive control," *Syst. Contr. Lett.*, vol. 3, pp. 243–246, Nov. 1983.
[8] S. Boyd and S. S. Sastry, "Necessary and sufficient conditions for parameter convergence in adaptive control," *Automatica*, vol. 22, pp. 629–639, 1986.
[9] S. Dasgupta, B. D. O. Anderson, and A. Tsoi, "Input conditions for continuous-time adaptive systems," in *Proc. CDC*, 1983, pp. 211–216.
[10] B. D. O. Anderson and R. Johnstone, "Global adaptive pole positioning," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 11–21, 1985.
[11] H. Elliott, R. Cristi, and M. Das, "Global stability of adaptive pole-placement algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 348-356, 1985.

The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients

HENK A. P. BLOM AND YAAKOV BAR-SHALOM

Abstract—**An important problem in filtering for linear systems with Markovian switching coefficients (dynamic multiple model systems) is the one of management of hypotheses, which is necessary to limit the computational requirements. A novel approach to hypotheses merging is presented for this problem. The novelty lies in the timing of hypotheses merging. When applied to the problem of filtering for a linear system with Markovian coefficients this yields an elegant way to derive the interacting multiple model (IMM) algorithm. Evaluation of the IMM algorithm makes it clear that it performs very well at a relatively low computational load. These results imply a significant change in the state of the art of approximate Bayesian filtering for systems with Markovian coefficients.**

Manuscript received June 24, 1987; revised October 21, 1987. This paper is based on a prior submission of October 20, 1986. The work of the second author was supported by the Air Force Office of Scientific Research under Grant 84-0112.
H. A. P. Blom is with the National Aerospace Laboratory, NLR, Amsterdam, The Netherlands.
Y. Bar-Shalom is with the University of Connecticut, Storrs, CT 06268.
IEEE Log Number 8821022.

I. INTRODUCTION

In this contribution we present a novel approach to the problem of filtering for a linear system with Markovian coefficients

$$x_t=a(\theta_t)x_{t-1}+b(\theta_t)w_t \tag{1}$$

with observations

$$y_t=h(\theta_t)x_t+g(\theta_t)v_t \tag{2}$$

θ_t is a finite state Markov chain taking values in $\{1,\cdots,N\}$ according to a transition probability matrix H, and w_t, v_t are mutually independent white Gaussian processes. The exact filter consists of a growing number of linear Gaussian hypotheses, with the growth being exponential with the time. Obviously, for filtering we need recursive algorithms whose complexity does not grow with time. With this, the main problem is to avoid the exponential growth of the number of Gaussian hypotheses in an efficient way.

This hypotheses management problem is also known for several other filtering situations [10], [5], [6], [9], and [4]. All these problems have stimulated during the last two decades the development of a large variety of approximation methods. For our problem the majority of these are techniques that reduce the number of Gaussian hypotheses, by pruning and/or merging of hypotheses. Well-known examples of this approach are the detection estimation (DE) algorithms and the generalized pseudo Bayes (GPB) algorithms. For overviews and comparisons see [14], [7], [12], and [17]. None of the algorithms discussed appeared to have good performance at modest computational load. Because of that, other approaches have been also developed, mainly by way of approximating the model (1), (2). Examples are the modified multiple model (MM) algorithms [20], [7], the modified gain extended Kalman (MGEK) filter of Song and Speyer [13], [7], and residual based methods [19], [2]. These algorithms, however, also lack good performance at modest computational load in too many situations. In view of this unsatisfactory situation and the practical importance of better solutions, the filtering problem for the class of systems (1), (2) needed further study.

One item that has not received much attention in the past is the timing of hypotheses reduction. It is common practice to reduce the number of Gaussian hypotheses immediately after a measurement update. Indeed, on first sight there does not seem to be a better moment. However, in two recent publications [3], [1], this point has been exploited to develop, respectively, the so-called IMM (interacting multiple model) and AFMM (adaptive forgetting through multiple models) algorithms. The latter exploits pruning to reduce the number of hypotheses, while the IMM exploits merging. The IMM algorithm was the reason for a further evaluation of the timing of hypotheses reduction. A novel approach to hypotheses merging is presented for a dynamic MM situation, which leads to an elegant derivation of the IMM algorithm. Next Monte Carlo simulations are presented to judge the state of the art in MM filtering after the introduction of the IMM algorithm.

II. TIMING OF HYPOTHESES REDUCTION

To show the possibilities of timing the hypothesis reduction, we start with a filter cycle from one measurement update up to and including the next measurement update. For this, we take a cycle of recursions for the evolution of the conditional probability measure of our hybrid state Markov process (x_t, θ_t). This cycle reads as follows:

$$P\{\theta_{t-1}|Y_{t-1}\}\xrightarrow{\text{Mixing}}P\{\theta_t|Y_{t-1}\} \tag{3}$$