1987-01-01

Simultaneous system identification and decision-directed detection and estimation of jump inputs to linear systems

Wynn C. Stirling
wynn_stirling@byu.edu

Original Publication Citation

Stirling, W. "Simultaneous System Identification and Decision-Directed Detection and Estimation of Jump Inputs to Linear Systems." Automatic Control, IEEE Transactions on 32.1 (1987): 86-9

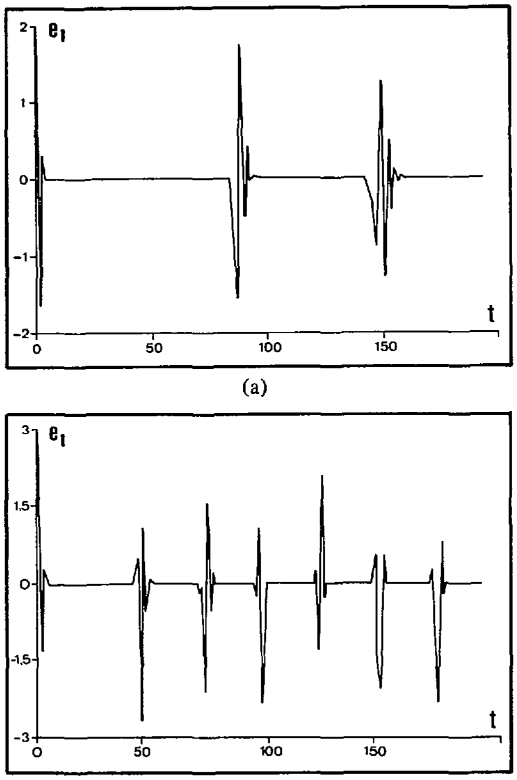

Fig. 2. (a) $G(s) = (s + 2)/(s^2 - s - 2)$, $\sigma = 0.03$. (b) $G(s) = (s + 1)/(s + 0.1)(s - 10)$, $\sigma = 0.03$.

remain virtually unaltered as $\sigma \to 0$, although the quiescent intervals ($e_1 \simeq 0$) become larger. This is in agreement with the bound of the "mean value" of $|e_1|$ given in [9], which is small if σ is small.

Hence, we note that, contrary to what can be predicted from the residual set established in [3], [4], in general, $\|e\|$ is not ultimately of order $\sigma^{1/2}$ as in the simple example of Section III (the two nontrivial equilibria were stable). A possible missing condition in the analysis of [3], [4] seems to be the "persistent excitation" requirement.

Indeed, bursts disappear when a rich input is introduced in the simulated examples. This is in agreement with the points raised in [5] about the effects of persistent (or rich) excitation. However, more is needed besides richness. The signal must also have a sufficient "level" [8] in terms of amplitude; otherwise bursts remain. We have observed that with a sufficiently rich and strong excitation the error does not tend to zero as in the case of $\sigma = 0$, but is ultimately bounded by a value that decreases to zero as $\sigma \to 0$. This indicates that, in order to guarantee a small residual error when a σ-modification is used, one should consider the input signal richness and "level."

If a persistent excitation is not available for a particular system (e.g, in regulation problems), it can be shown that a "switching" σ-factor [10] ($\sigma = \sigma_0$ if $\|\theta\| > M > \|\theta^*\|$; $\sigma = 0$ otherwise, where θ^* is the exact matching gain vector) leads to zero residual error in the ideal case. Yet, this property is not known to be robust in the sense that small perturbations would lead to an asymptotically small output error, except in the mean value sense [9]. Thus, although all signals in the system may remain uniformly bounded, bursting phenomena cannot be excluded in the presence of perturbations, however small.

VI. Conclusion

In this paper we have shown that a continuous-time adaptive system can exhibit bursting phenomena when a σ-modification is introduced in the adaptation law and the system operates in regulation mode. As a consequence, some results presented in recent papers should be revised.

Through the analysis of the system equilibria, bursting phenomena were inferred and confirmed by simulation. Persistent excitation with a sufficient level is pointed out as a means for eliminating bursts and for obtaining small residual errors.

Appendix

Stability of Nontrivial Equilibria

Let c_1, d_1, and d_0 denote the nonnull elements of (4.7) in proper sequence. The system $\Delta\dot{e} = [\bar{A} - b_c\theta_0^T W]\,\Delta e$ with θ_0 given by (4.7) corresponds to the following input–output equations: $L(p)u_1 = -c_1 u$; $L(p)u_2 = -d_1 y$; $u_3 = -d_0 y$; $u = u_1 + u_2 + u_3$; $N(p)u = D(p)y$, where $p = d/dt$; $L(s)$ is the monic characteristic polynomial of Λ, and $G(p) = N(p)/D(p)$ is the plant transfer operator. From these equations one has $NL(d_1 + d_0 L)u_1 = -c_1(d_1 + d_0 L)Dy$ and $D(c_1 + L)Lu_1 = c_1(d_1 + d_0 L)Dy$. The characteristic polynomial of $\bar{A} - b_c\theta_0^T W$ is thus given by

$$P(s) = L(s)\{N(s)[d_1 + d_0 L(s)] + D(s)[c_1 + L(s)]\}. \qquad \text{(A.1)}$$

Stability then follows from $P(s)$.

References

[1] K. J. Astrom, "Theory and applications of adaptive control: A survey," *Automatica,* vol. 19, no. 5, 1983.

[2] T. R. Fortescue, L. S. Kershenbaum, and B. E. Ydstie, "Implementation of self-tuning regulators with variable forgetting factors," *Automatica,* vol. 17, 1981.

[3] P. A. Ioannou and P. V. Kokotovic, "Robust redesign of adaptive control," *IEEE Trans. Automat. Contr.,* vol. AC-29, Mar. 1984.

[4] P. A. Ioannou and P. V. Kokotovic, "Instability analysis and improvement of robustness of adaptive control," *Automatica,* vol. 20, no. 5, 1984.

[5] B. D. O. Anderson, "Adaptive systems, lack of persistency of excitation and bursting phenomena," *Automatica,* vol. 21, no. 3, 1985.

[6] K. S. Narendra and L. S. Valavani, "Stable adaptive controller design. Direct control," *IEEE Trans. Automat. Contr.,* vol. AC-23, Aug. 1978.

[7] L. Hsu and R. R. Costa, "Improvement of adaptive control robustness in presence of unmodeled dynamics," presented at 24th IEEE Conf. Decision Contr., Fort Lauderdale, FL, Dec. 1985.

[8] M. Bodson and S. Sastry, "Exponential convergence and robustness margins in adaptive control," presented at 23rd IEEE Conf. Decision Contr., Las Vegas, NV, Dec. 1984.

[9] P. Ioannou and J. Sun, "Theory and design of robust direct and indirect adaptive control schemes," Univ. Southern California, Los Angeles, Rep. 86-06-1, 1986.

[10] P. Ioannou, "Robust direct adaptive control," presented at 23rd IEEE Conf. Decision Contr., Las Vegas, NV, Dec. 1984.

Simultaneous System Identification and Decision-Directed Detection and Estimation of Jump Inputs to Linear Systems

WYNN C. STIRLING

Abstract—**A decision-directed approach is presented for analyzing linear systems with unknown jump inputs. The system model parameters**

Manuscript received February 24, 1986. This paper is based on a prior submission of September 9, 1985. This work was supported by SPAWAR under Contract N00039-85-C-0223.

The author is with the Department of Electrical Engineering, Brigham Young University, Provo, UT 84602.

IEEE Log Number 8611409.

0018-9286/87/0100-0086$01.00 © 1987 IEEE

are estimated using a Kalman filter, and an empirical Bayes detection procedure is introduced to set the detector parameters, resulting in a decision-directed generalized *likelihood ratio test* coupled with recursive system parameter estimation. Monte Carlo results are presented to validate the performance of the algorithm.

I. Introduction

This correspondence concerns the problem of estimating the model parameters of a linear system subjected to jump inputs. Standard system identification procedures are adequate when the input consists of zero-mean independent variates, but may yield biased estimates when these assumptions do not hold. Also, it is desirable to estimate the jump component of the input as well as to identify the system parameters. We address the problem of simultaneously estimating the system model parameters and detecting the jump inputs when the rate and amplitude of the jump process is unknown and possibly time-varying.

This study will be limited to discrete-time systems, and attention will be restricted to signal models of the form

$$\theta_{t-1}=\theta_t$$

$$y_t=\phi_t^T\theta_t+j_t+v_t \tag{1}$$

where for $t = 0, 1, \cdots$. $\theta_t = \theta = [a_1, \cdots, a_n]^T$ is a vector of unknown signal model parameters, ϕ_t is vector function of past observations (we shall take $\phi_t = [-y_{t-1}, \cdots, -y_{t-n}]^T$, resulting in an autoregressive model for the process $\{y_t\}$), $\{v_t\}$ is a white noise process of variance $r(t)$, and the jump process $j_t = cn_t$ is a marked point process where c is a random variable and $\{n_t\}$ is a discrete-time point process (DTPP). The estimation/detection problem associated with the above signal model is to estimate the model coefficient vector θ, detect the jumps (i.e., when n_t = 1), and estimate the jump amplitude, c.

II. Problem Formulation

We take a Bayesian approach to the parameter estimation problem, and assume that θ is a random variable with *a priori* mean and covariance θ_0 and P_0, respectively. In the absence of jump inputs the Kalman filter provides an estimate for θ of the form

$$\hat{\theta}_t=\hat{\theta}_{t-1}+P_t\phi_t e_t \tag{2}$$

where $e_t = y_t - \phi_t^T\hat{\theta}_t$ is the innovations of y_t with

$$P_t=P_{t-1}-\frac{P_{t-1}\phi_t\phi_t^TP_{t-1}}{r(t)+\phi_t^TP_{t-1}\phi_t} \tag{3}$$

and initial conditions θ_0 and P_0. It is well known that this estimator is consistent if the driving process is white [1], in which case, $\hat{\theta}_t$ is the conditional expectation of θ given the data $y_0, \cdots, y_{t-1}$, and P_t is the conditional variance of the estimation error, $\hat{\theta}_t - \theta$.

For the present problem, however, the driving process contains a jump component which may be neither zero-mean nor independent, and the Kalman filter results will not, in general, yield consistent estimates of the state. In [2], a general nonlinear minimum-mean-square-error estimator is derived based on the Girsanov transformation, and a solution is presented for Gauss-Bernoulli jumps. This approach, however, requires a growing number of estimators to account for all possible hypotheses.

We propose an alternative solution incorporating an empirical Bayes decision rule to detect the occurrence of jumps. We assume that the point process n_t is governed by a random rate, λ_t. The empirical Bayes procedure [3] uses the observations to obtain an estimate the prior distribution which, in this case, corresponds to estimating λ_t. To formalize this problem, define

$$\lambda_t=P\{n_t=1|\sigma\{\theta,\ j_s,\ v_s,\ s\le t\}\},$$

where $\sigma\{\cdot\}$ denotes the σ-field generated by the arguments. We shall assume explicitly that λ_t may be time-varying. If λ_t and c were known, the Bayes decision rule would result in a likelihood ratio test at each time t of the form

$$\delta(e_t,\ \lambda_t,\ c)=\begin{cases}1 & \text{if } (1-\lambda_t)f_0(e_t)<\lambda_t f_1(e_t|c)\\ 0 & \text{otherwise}\end{cases},$$

where f_0 and f_1 are the probability density functions of e_t under the hypotheses H_0 (no jump occurred at time t) and H_1 (a jump occurred at time t). The density f_1 must be parameterized by the amplitude c of the jump. It should be noted, when jumps are possible, that e_t may no longer be treated as an innovations process since, even if λ_t and c are known exactly, there is a nonzero probability of detection error. Thus, consistency of the estimates cannot be guaranteed with a detection approach.

Unfortunately, neither λ_t nor c is known *a priori*. Consequently, we adopt a decision-directed empirical Bayes approach to estimate them, resulting in a generalized likelihood ratio test of the form $\delta(e_t, \hat{\lambda}_{t|t-1}, \hat{c}_{t-1})$ where $\hat{\lambda}_{t|t-1}$ and $\hat{c}_{t-1}$ are estimates of λ_t and c given observations of y_s, $s \le t-1$. Let

$$N_t=\delta(e_t,\ \hat{\lambda}_{t|t-1},\ \hat{c}_{t-1})$$

denote the *detected* discrete-time point-process, and modify (2) to become

$$\hat{\theta}_t=\hat{\theta}_{t-1}+P_t\phi_t\epsilon_t$$

where $\epsilon_t = e_t - \hat{c}_{t-1}N_t$.

Amplitude Estimation

The decision-directed rule for estimating the jump amplitude is very simple: compute the conditional mean of c given the detected jump history $\{N_0, \cdots, N_t\}$. Hence,

$$\hat{c}_t=\frac{\sum_{s=1}^{t} e_sN_s}{\sum_{s=1}^{t} N_s}. \tag{4}$$

When no detections are made (i.e., $\Sigma N_s = 0$)$\hat{c}_t$ must be set to some appropriate *a priori* value. It is evident that the estimate for c is unbiased if there are no false alarms. Analysis of this bias error is difficult, however, since the problem is embedded in the larger problem of system model parameter estimation and jump detection.

Rate Estimation

The detected DTPP, N_t, represents the detector output; it is the rate of this process that is to be estimated and used to approximate the rate of j_t. The relationship between π_t, the rate of N_t, and λ_t, the rate of n_t, is discussed in [4] for constant λ, where it is shown that the relative frequency estimate of π_t either undergoes a runaway (i.e., converges to 0 or 1) or converges to a steady-state value that is close to λ. Bounds are obtained for the probability of runaway as a function of the signal-to-noise ratio, and it is shown that, even for moderate-to-low signal-to-noise ratios, the probability of runaway is extremely low. In [5], a bias-removing transformation is introduced to ensure convergence of π_t to the true value of the prior.

A significant generalization from a constant (or even slowly varying) prior is to model the prior as a finite-state Markov chain [7], [6] with state vector $\rho = [\rho_1, \cdots, \rho_m]^T$, where $\rho_1 < \cdots < \rho_m$, with transition probabilities

$$q_{ij}(t)=Pr\ \{\pi_t=\rho_j|\pi_{t-1}=\rho_i\}$$

with initial distribution $\alpha = [\alpha_1, \cdots, \alpha_m]^T$, where $\alpha_i = Pr\{\pi_0 = \rho_i\}$.

With such a model, runaway can be eliminated by choice of ρ, and the dynamical behavior of the prior can be modeled by the transition matrix $Q = \{q_{ij}\}$. Following Segall [7], define the vector $x_t = [x_1(t), \cdots, x_m(t)]^T$ by

$$x_i(t)=\begin{cases}1, & \text{if } \pi_t=\rho_i \\ 0, & \text{otherwise}\end{cases} \quad i=1, 2, \cdots, m.$$

Thus, $\pi_t = \rho^T x_t$. The vector x_t can be viewed as the state vector of a system obeying dynamics and observation equations of the form

$$x_{t+1}=Q_t^T x_t+u_t \tag{5}$$

$$N_t=\rho^T x_t+w_t \tag{6}$$

The processes $\{u_t\}$ and $\{w_t\}$ are Martingale difference (MD) sequences with respect to the family of σ-fields $\{\mathcal{B}_t\}$ where $\mathcal{B}_t = \sigma\{N_0, \cdots, N_t, x_0, \cdots, x_{t+1}\}$. Equation (6) represents the Doob decomposition of $\{N_t\}$ with respect to $\{\mathcal{B}_t\}$, which family of σ-fields is unobservable. The estimation problem, consequently, consists of obtaining the Doob decomposition with respect to a family of σ-fields that is observable, namely $\{\mathfrak{F}_t\}$ where $\mathfrak{F}_t = \sigma\{N_0, \cdots, N_t\}$.

For the case where the events $x_i(t+1) = 1$ for $i = 1, \cdots, m$ and $N_t = 1$ are conditionally independent given x_t, the one-step prediction of x_{t+1} given the σ-field $\mathfrak{F}_t = \sigma\{N_0, \cdots, N_t\}$ is

$$\hat{x}_{t+1|t}=Q_t^T\hat{x}_{t|t-1}+Q_t^T\frac{\text{diag }(\hat{x}_{t|t-1})-\hat{x}_{t|t-1}\hat{x}_{t|t-1}^T}{\rho^T\hat{x}_{t|t-1}-(\rho^T\hat{x}_{t|t-1})^2}\,\nu_{t|t-1}$$

where $\nu_{t,t-1} = (N_t - \rho^T\hat{x}_{t|t-1})$ is the innovations process of N_t and diag (x) denotes a diagonal matrix composed of the elements of the vector x. Thus, $N_t = \rho^T\hat{x}_{t|t-1} + \nu_{t|t-1}$ is the Doob decomposition of $\{N_t\}$ with respect to $\{\mathfrak{F}_t\}$. The conditional estimation error covariance for $\tilde{x}_{t-1\,t} = x_{t+1} - \hat{x}_{t+1|t}$ given by

$$\Sigma_{t+1|t}=E^{\mathfrak{F}_t}\tilde{x}_{t+1|t}\tilde{x}_{t+1|t}^T=\text{diag }(\hat{x}_{t+1|t})-\hat{x}_{t+1|t}\hat{x}_{t+1|t}^T$$

which follows from the fact that $x_t x_t^T = \text{diag }(x_t)$. The conditional expectation of the rate is thus $\hat{\pi}_{t+1|t} = \rho^T\hat{x}_{t-1|t}$ and the conditional variance of the estimation error $\tilde{\pi}_{t+1|t} = \pi_{t+1} - \hat{\pi}_{t+1|t}$ is

$$\text{Var }(\tilde{\pi}_{t+1|t}|\mathfrak{F}_t)=E^{\mathfrak{F}_t}\rho^T\tilde{x}_{t+1|t}\tilde{x}_{t+1|t}^T\rho=\rho^T\text{ diag }(\rho)\hat{x}_{t+1|t}-(\hat{\pi}_{t+1|t})^2.$$

The philosophy of the decision-directed empirical Bayes approach is to use $\hat{\pi}_{t|t-1}$ as an approximation for λ_t, i.e., the rate estimate used for the generalized likelihood ratio test is $\hat{\lambda}_{t|t-1} = \hat{\pi}_{t\,t-1}$, which corresponds to the one-step prediction of the rate at time t, given data up to time $t-1$. Bias in the estimate may be removed by appropriate transformations [5], but is not considered a critical factor in the current analysis and is ignored.

III. Monte Carlo Simulations

The interaction between detection and estimation makes performance analysis of the above algorithm extremely complex. The difficulty is due not only to the dependencies present in the adaptive detector, but also to the non-Gaussian dependencies introduced by the two-way coupling between the detector and estimator which are virtually impossible to treat since the multivariate distributions are not available in analytic form. As an alternative to theoretical performance analysis, therefore, selected Monte Carlo simulation results are provided to indicate performance of the algorithm.

For the simulations described below, a Gauss–Bernoulli mixed process consisting of the sum of a unit-variance white Gaussian noise and a constant-rate constant-amplitude DTPP drives a fourth-order autoregressive model, with $\theta = [-1.3636, -1.4401, -1.0919, 0.8353]^T$. A nine-state Markov chain model was applied to this estimation problem, with states and transition matrix given as

$$\rho=[0.001,\ 0.002,\ 0.005,\ 0.01,\ 0.02,\ 0.05,\ 0.1,\ 0.2,\ 0.5]^T$$

$$Q=\begin{bmatrix} 1-\gamma & \gamma & 0 & \cdots & 0 & 0 \\ \frac{\gamma}{2} & 1-\gamma & \frac{\gamma}{2} & \cdots & \cdot & \cdot \\ 0 & \frac{\gamma}{2} & 1-\gamma & \cdots & \cdot & \cdot \\ \cdot & 0 & \frac{\gamma}{2} & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \frac{\gamma}{2} & 0 \\ \cdot & \cdot & 0 & \cdots & 1-\gamma & \frac{\gamma}{2} \\ 0 & 0 & 0 & \cdots & \frac{\gamma}{2} & 1-\gamma \end{bmatrix}$$

with $\gamma = 0.002$. Four different Monte Carlo series were performed with different jump parameters as indicated in Table I.

Results are provided in Tables II and III. Table II displays the sample mean, $\tilde{a}_i$, and sample standard deviation, $\hat{\sigma}_{a_i}$, of the model parameter estimation errors. This table also provides an approximation to the "ideal" standard deviation, denoted $\bar{\sigma}_{a_i}$, consisting of the conditional standard deviation derived from (3) for a representative trial adjusted for detection errors by invoking the assumption that the detection errors are zero-mean and independent. The resulting conditional covariance matrix is

$$\bar{P}_t=P_t[1+c^2(\lambda P_{MD}+(1-\lambda)P_{FA}] \tag{7}$$

where P_{MD} and P_{FA} denote the theoretical probabilities of missed detection and false alarm, respectively, assuming no model parameter estimation errors and known values of c and λ. It is evident from Table II that there is no appreciable bias on the parameter estimates; compared with the sample standard deviations, all sample mean errors lie within one standard deviation of zero. The sample standard deviations, however, are somewhat larger than the "ideal" standard deviations computed via (7), indicating that the ideal values may be optimistic, and the assumptions used to generate this covariance represent a significant oversimplification.

Table III displays the sample means and standard deviations of the amplitude and rate estimation errors, denoted by $\tilde{c}$ and $\tilde{\lambda}$, and $\hat{\sigma}_c$ and $\hat{\sigma}_\lambda$, respectively. This table also provides the empirical probabilities of missed detection ($\hat{P}_{MD}$) and false alarm ($\hat{P}_{FA}$), along with the corresponding theoretical values. It is evident that the estimation errors all lie within one sample standard deviation of zero, indicating consistent performance. Not surprisingly, the higher the rate, the more precise the estimates, since, for the low-rate cases, the missed detection rate is extremely high, resulting in few correct decisions and, hence, few values to process in the decision-directed estimator. For the high-rate cases, however, considerable improvement in performance is evident.

The joint estimation-detection procedure described above appears to work well, with increased precision obtained as the rate-amplitude product is increased. In no case do the estimates diverge, nor is there a significant bias in the estimates. In fact, a self-stabilizing effect is observed: when the rate-amplitude product is low, the detection errors are large, but the effect is small on the model parameter estimates, since the mean value of the process is very nearly zero, and when the rate-amplitude product is large, the detection errors are small, and the decision directed detector/estimator works well, resulting in unbiased estimates of the model parameters.

TABLE I
MONTE CARLO INPUTS

Case	1	2	3	4
Amplitude	2.0	4.0	4.0	6.0
Rate	0.02	0.02	0.2	0.2

TABLE II
MONTE CARLO STATISTICS FOR MODEL PARAMETER ESTIMATES

Case	1	2	3	4
$\tilde{a}_1$	−0.008	−0.01	−0.027	−0.019
$\tilde{a}_1$	0.012	0.013	0.040	0.027
$\tilde{a}_3$	0.005	−0.007	−0.024	−0.016
$\tilde{a}_4$	0.002	−0.002	0.004	0.001
$\hat{\sigma}_{a_1}$	0.019	0.020	0.031	0.021
$\hat{\sigma}_{a_2}$	0.036	0.032	0.046	0.031
$\hat{\sigma}_{a_3}$	0.025	0.024	0.029	0.020
$\hat{\sigma}_{a_4}$	0.014	0.014	0.029	0.007
$\bar{\sigma}_{a_1}$	0.018	0.017	0.011	0.007
$\bar{\sigma}_{a_2}$	0.030	0.028	0.018	0.012
$\bar{\sigma}_{a_3}$	0.030	0.029	0.018	0.008
$\bar{\sigma}_{a_4}$	0.018	0.017	0.011	0.007

TABLE III
MONTE CARLO STATISTICS FOR JUMP PARAMETER ESTIMATES

Case	1	2	3	4
$\tilde{c}$	−0.140	−0.328	−0.057	0.052
$\tilde{\lambda}$	0.016	0.003	0.006	0.006
$\hat{\sigma}_c$	0.936	0.390	0.163	0.087
$\hat{\sigma}_\lambda$	0.016	0.010	0.019	0.014
$\hat{P}_{FA}$	0.0015	0.001	0.008	0.0005
$\hat{P}_{MD}$	0.929	0.214	0.076	0.007
P_{FA}	0.0016	0.001	0.009	0.0006
P_{MD}	0.828	0.152	0.049	0.003

IV. CONCLUSIONS

A joint system model parameter estimation-jump input detection and estimation procedure has been presented. In order to apply this joint estimator-detector, the structure of the jump process must be known, but the parameters of the jump (i.e., the amplitude and rate) may be estimated. Analysis of performance of this algorithm is difficult, since there is two-way coupling between the decision-directed jump input detector/estimator and the system model parameter estimator. Consequently, Monte Carlo studies have been performed to assess empirically the performance. These simulation studies indicate that the algorithm performs well, yielding parameter estimates with negligible bias and satisfactory detection performance. Furthermore, a self-stabilizing effect is observed, indicating that the problem of simultaneously identifying system model parameters and detecting and estimating jump inputs is robust.

REFERENCES

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

[2] A. Makowski, W. S. Levine, and M. Asher, "The nonlinear mmse filter for partially observed systems driven by non-Gaussian white noise, with applications to failure estimation," in *23rd IEEE Conf. Decision and Control*, 1984, pp. 644–650.

[3] H. Robbins, "The empirical bayes approach to statistical decision problems," *Rev. Inst. Int. Statist.*, vol. 31, pp. 195–208, 1963.

[4] L. D. Davisson and S. C. Schwartz, "Analysis of a decision-directed receiver with unknown priors," *IEEE Trans. Inform. Theory*, vol. IT-16, no. 3, pp. 270–276, May 1970.

[5] A. Katopis and S. C. Schwartz, "Decision-directed learning using stochastic approximation," in *Proc. 3rd Ann. Pittsburgh Conf. Modeling and Simulation*, Univ. Pittsburgh, Pittsburgh, PA, Apr. 1972, pp. 473–481.

[6] W. C. Stirling and M. Morf, "A new decision-directed algorithm for nonstationary priors," *IEEE Trans. Automat. Contr.*, vol. AC-29, no. 10, pp. 928–930, Oct. 1984.

[7] A. Segall, "Recursive estimation from discrete-time point processes," *IEEE Trans. Inform. Theory*, vol. IT-22, no. 4, pp. 422–431, July 1976.

[8] S. C. Schwartz and A. Katopis, "Modified stochastic approximation to enhance unsupervised learning," in *Proc. 1977 Conf. Decision and Control*, 1977, pp. 1067–1079.

[9] D. Kazakos and L. D. Davisson, "An improved decision-directed detector," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 1, pp. 113–116, Jan. 1980.

Some Asymptotic Properties of Multivariable Models Identified by Equation Error Techniques

PAUL VAN DEN HOF AND PETER JANSSEN

Abstract—**Some interesting properties are derived for simple equation error identification techniques—least squares and basic instrumental variable methods—applied to a class of linear, time-invariant, time-discrete multivariable models. The system at hand is not supposed to be contained in the chosen model set. Assuming that the input is unit variance white noise, it is shown that the Markov parameters of the system are estimated asymptotically unbiased over a certain interval around $t = 0$.**

I. INTRODUCTION

In system identification literature, there is a growing interest in considering situations where the process at hand is not necessarily contained in the chosen model set. This interest is motivated by the fact that in many practical situations of system identification, a model will be required that is of restricted complexity, approximating the essential characteristics of the (possibly very complex) process, rather than a very sophisticated model that exactly models the process behavior. The way in which the original process is approximated by the model now is dictated by the applied identification method, and the chosen model set. Equation error techniques are rather popular, mainly due to their computational simplicity. Since, in many situations, the performance of an identified model is judged upon its ability to simulate the process under study, it is important to analyze the simulation behavior of an approximate model obtained by equation error techniques. By considering the Markov parameters of the identified model, we will focus on properties in the time domain. For an analysis in the frequency domain, see [1].

Manuscript received May 5, 1986; revised September 15, 1986.

P. Van den Hof was with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. He is now with the Laboratory for Measurement and Control, Department of Mechanical Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands.

P. Janssen is with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands.

IEEE Log Number 8611556.

0018-9286/87/0100-0089$01.00 © 1987 IEEE