where $T_n(a, b, k)$ is defined by (3.8) and

$$F_n(a, b)=\frac{\operatorname{tr} \tilde{\theta}^\tau(a, b)P_n(a, b)X_n^\tau(a, b)\bar{\tilde{X}}_n(K-a, K-b)D(K, a, b)\bar{\tilde{X}}_n^\tau(K-a, K-b)X_n(a, b)P_n(a, b)\tilde{\theta}(a, b)}{r_n(a, b)\,(\log r_n(a, b))^{1+\epsilon}}$$
$$\leqslant \operatorname{tr} \tilde{\theta}^\tau(a, b)P_n(a, b)X_n^\tau(a, b)X_n(a, b)P_n(a, b)\tilde{\theta}(a, b)\,(\log r_n(a, b))^{-1-\epsilon}$$
$$\leqslant \operatorname{tr} \tilde{\theta}^\tau(a, b)\tilde{\theta}(a, b)\,(\log r_n(a, b))^{-1-\epsilon}\rightarrow 0. \quad (3.27)$$

Together, (3.8), (3.26), and (3.27) confirm (3.25).

Lemma 4: If conditions A and B hold, then there exist constants $K_3 > 0$ and $M_3 > 0$ such that

$$S_n(a, b, K)>K_3r_n(a, b)\,(\log r_n(a, b))^{1+2\epsilon}, \quad \forall n>M_3 \quad (3.28)$$

where either $a < p$, $b \leq d_o$ or $a \leq d_o$, $b < q$; $K \geqslant d_o = \max(p, q)$.

Proof: From (2.10)–(2.13), (3.1) and (3.23)–(3.24), it can be shown that

$$\begin{aligned}S_n(a, b, k) \geqslant{}& 2^{-k}\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|^2-r_n(a, b)\\ &\cdot(\log r_n(a, b))^{1+\epsilon}-2\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|\\ &\cdot\|D(K, a, b)\bar{\tilde{X}}_n^\tau(K-a, K-b)Q_n(a, b)W_{n+1}\|\\ &-2\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|\cdot\|D(K, a, b)\\ &\cdot\bar{\tilde{X}}_n^\tau(K-a, K-b)X_n(a, b)P_n(a, b)\tilde{\theta}(a, b)\|. \end{aligned} \quad (3.29)$$

By Lemma 1 and (3.24)

$$\begin{aligned}&\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|^2\\ &=\|\tilde{\theta}^\tau(p_0, q_0)(M_n(p_0-a, q_0-b))^2\tilde{\theta}(p_0, q_0)\|\\ &\geqslant K_1^2\|\tilde{\theta}(p_0, q_0)\|^2r_n(d_0, d_0)\,(\log r_n(d_0, d_0))^{1+2\epsilon}.\end{aligned} \quad (3.30)$$

By Lemma 2 and (3.30)

$$\begin{aligned}L_n=&\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|^{-2}\\ &\cdot\|D(K, a, b)\bar{\tilde{X}}_n^\tau(k-a, k-b)Q_n(a, b)W_{n+1}\|^2\\ \leqslant{}& K_1^{-2}\|\tilde{\theta}(p_0, q_0)\|^{-2}T_n(a, b, K)\rightarrow 0.\end{aligned} \quad (3.31)$$

Obviously,

$$\begin{aligned}\tilde{L}_n=&\|D(K, a, b)\bar{\tilde{X}}_n^\tau(K-a, K-b)X_n(a, b)P_n(a, b)\tilde{\theta}(a, b)\|\\ &\cdot\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|^{-1}\rightarrow 0. \end{aligned} \quad (3.32)$$

From (3.30)–(3.32), it is clear that

$$R_n(a, b)=2^{-k}-\frac{r_n(a, b)\,(\log r_n(a, b))^{1+\epsilon}}{\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|^2}-2L_n^{0.5}-2\tilde{L}_n \geqslant 2^{-k-1}$$

for sufficiently large n; hence,

$$\begin{aligned}S_n(a, b, K) &\geqslant R_n(a, b)\|M_n(p_0-a, q_0-b)\tilde{\theta}(p_0, q_0)\|^2\\ &\geqslant K_3r_n(a, b)\,(\log r_n(a, b))^{1+2\epsilon}.\end{aligned}$$

Then (3.28) is established by taking

$$K_3=2^{-k-1}K_1^2\min(\|A_p\|^2, \|B_q\|^2).$$

Proof of Theorem 1: If $d_o = d$, then (2.14) holds by Lemma 3 and (2.15) is valid by Lemma 4, with $a = b = d - 1$ and $K_2 = \min(K_2', K_3)$. Conversely, if (2.14) is true, then it must be that $d \geqslant d_o$; otherwise, $d < d_o$. Then Lemma 4 and the fact that $r_n(d, d) < r_n(d_o, d_o)$ imply that

$$S_n(d, d, K)>K_3r_n(d, d)\,(\log r_n(d, d))^{1+2\epsilon}$$

this is a contradiction of (2.14). Similarly, the validity of (2.15) means that $d \leq d_o$ or we infer that $d - 1 \geqslant d_o$; then by Lemma 3,

$$S_n(d-1, d-1, K)<-K_2'r_n(d-1, d-1)\,(\log r_n(d-1, d-1))^{1+\epsilon}.$$

This contradicts (2.15). Thus assertion (1) of Theorem 1 has been verified.

In case of $d = d_o$, suppose that (2.16)–(2.18) hold; we prove that $p_1 = p$ and $q_1 = q$. When $p = q$, if either $p_1 < p$ or $q_1 < q$, then (2.16) must fail by Lemma 4. When $p \neq q$, say $q < p = d_o$, then p_1 must equal p; otherwise, (2.16) cannot hold. If $p_1 = p$ and $q_1 > q$, then (2.18) fails and if $p_1 = p$, $q_1 < q$, then (2.16) is not true by reasoning similar to that used in the proof of assertion (1). We see that $p_1 = p$ and $q_1 = q$.

On the other hand, if $p_1 = p$ and $q_1 = q$, the validity of (2.16)–(2.18) is quite easily checked by Lemmas 3 and 4. End of proof.

IV. Conclusion

Theorems for order-determination without *a priori* knowledge of upper bounds on the order in dynamic systems are developed. Deterministic procedures to determine orders and estimate parameters simultaneously are introduced. In order to be successful and practical, it is necessary that $P_n(a, b)$, $\theta_n(a, b)$, and $S_n(a, b, k)$ be generated from recursive forms. When the upper bounds of orders are known (which means that k is fixed), these can easily be found. The other case will be further discussed and developed in detail elsewhere. It is desirable to weaken Condition A and generalize the results to systems with correlated noise.

References

[1] P. Whittle, "Hypothesis testing in time series analysis," Ph.D. thesis, Almquist and Wiksell, Uppsala, 1951.

[2] H. Akaike, "Markovian representations of stochastic processes and its application to the analysis of autoregressive moving average processes," *Ann. Inst. Statist. Math.*, vol. 26, pp. 363–387, 1974.

[3] H. F. Chen and L. Guo, "Consistent estimation of the order of stochastic control systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 531–535, 1987.

[4] T. Z. Lai and C. Z. Wei, "Least-squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *Ann. Statist.*, vol. 10, pp. 154–166, 1982.

[5] E. J. Hannan and L. Kavalieris, "Multivariate linear time series models," *Adv. Appl. Prob.*, vol. 16, pp. 492–561, 1984.

[6] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 16, pp. 461–464, 1978.

Integral Action in Robust Adaptive Control

ARIE FEUER AND GRAHAM C. GOODWIN

***Abstract*—It is well known that integral action can be added to linear systems to achieve zero steady-state error for constant reference inputs and disturbances. The result has also been extended [3] to exponentially stable nonlinear systems. This note shows that a similar property holds for robust adaptive control systems.**

Manuscript received December 8, 1986; revised March 18, 1988. This work was supported by the Australian Technion Society and by the Technion—Israel Institute of Technology, Haifa, Israel.

A. Feuer is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.

C. G. Goodwin is with the Department of Electrical and Computer Engineering, University of Newcastle, New South Wales, Australia.

IEEE Log Number 8929178.

I. INTRODUCTION

There exists an extensive literature on the use of integral action in linear control systems; see, for example, [1], [2]. Recently, these results have also been extended to exponentially stable nonlinear systems [3].

In a separate line of development, considerable progress has recently been made on the problem of designing adaptive control laws which are robust to certain kinds of unmodeled dynamics; see, for example, [4]-[8]. In one of these approaches [8], integral action was included in the design and it was shown that the tracking error converged to zero in the presence of constant reference inputs, constant disturbances, and unmodeled dynamics. In [8], the integrator was included in the design from the beginning. However, an option available in classical control, as exemplified in [2] and [3], is to retrofit the integrator to an otherwise stable design. We have tested this idea in practice for adaptive control and found it to work well. The purpose of the current note is to establish the conditions under which this design option is soundly based.

Our method of analysis amounts to verifying that the key condition in [8], needed for convergence, is satisfied here as well. This illustrates a proof paradigm which could be similarly applied to other problems. The principal assumptions used will be as in [8], namely that the plant when augmented by an integrator has a known degree of controllability and that an overbounding function is known for the unmodeled system response.

We will treat the continuous-time case, but the corresponding discrete-time results follow mutatis mutandis by simply replacing $\rho = d/dt$ by $\delta = q - 1/\Delta$, L_2 by l_2, $\int$ by Σ, and so on.

II. THE SYSTEM MODEL

As in [8], let the plant be described by

$$y = H_o(1+H')u \tag{2.1}$$

where u, y denote the plant input and output, respectively, H_o denotes the "modeled" part of the plant, and H' denotes unmodeled dynamics.

The model H_o will be parameterized as a rational transfer function, i.e.,

$$H_o = \frac{B}{A} \tag{2.2}$$

where

$$A = \rho^n + a_{n-1}\rho^{n-1} + \cdots + a_o$$

$$B = b_{n-1}\rho^{n-1} + b_{n-2}\rho^{n-2} + \cdots + b_o; \qquad b_o \neq 0.$$

We require the following assumptions about the plant (2.1).

Assumptions A1:

i) The operator H' describing the unmodeled dynamics is analytic outside the stability region and has an impulse response which is bounded by a decaying exponential;

ii) the plant $H_o(1 + H')$ is strictly proper; and

iii) the plant, when augmented by $1/\rho$ has a coprime fractional representation in the ring of causal stable operators. More precisely, we require that for a given Hurwitz polynomial J of degree $(n + 1)$ there exist causal stable operators χ and Ω such that [9],

$$\chi\left[\frac{A\rho}{J}\right] + \Omega\left[\frac{B(1+H')}{J}\right] = 1. \tag{2.3}$$

III. PARAMETER ESTIMATION

We plan to design an adaptive controller including integral action which should: a) stabilize the system; and b) cause y to asymptotically track a constant reference input y^*. Towards this end, we express the model (2.1) in filtered (or fractional [9]) form as

$$\bar{y} \triangleq Ey_f = (E-A)y_f + Bu_f + \eta_f \tag{3.1}$$

where

$$E = \rho^n + e_{n-1}\rho^{n-1} + \cdots + e_o$$

is Hurwitz and

$$y_f = \frac{\rho}{EQ} y \tag{3.2}$$

$$u_f = \frac{\rho}{EQ} u \tag{3.3}$$

$$\eta_f = \frac{\rho B}{EQ} H'u \tag{3.4}$$

where Q is Hurwitz of degree 1 and where η_f denotes the filtered unmodeled response. Note that the filter $1/E$ is a low-pass filter, whereas ρ/Q is a high-pass filter. Thus, (3.1) is simply obtained from (2.1) by band-pass filtering.

Equation (3.1) can be written in the standard regression form [10] as

$$\bar{y}(t) = \phi(t)^T\theta_* + \eta_f(t) \tag{3.5}$$

where ϕ and θ_* are the regression vector and tuned parameters, respectively, given by

$$\phi(t)^T = [y_f, \cdots, \rho^{n-1}y_f, u_f, \cdots, \rho^{n-1}u_f] \tag{3.6}$$

$$\theta_*^T = [e_o - a_o, \cdots, e_{n-1} - a_{n-1}, b_o, \cdots, b_{n-1}]. \tag{3.7}$$

Because H' is exponentially stable it is readily seen [5], [8] that the filtered unmodeled error η_f can be bounded by an exponential function of past inputs, i.e.,

$$|\eta_f| < \beta \qquad \forall t \tag{3.8}$$

where

$$\beta(k) = \epsilon\sigma_o^{k-t}|v(\rho)u_f(t)| \tag{3.9}$$

where $\sigma_o < 1$ and $\nu(\rho)$ is an arbitrary Hurwitz polynomial of degree $n - 1$.

As in [4]-[8], we assume sufficient knowledge of the unmodeled dynamics to find ϵ, σ_o such that (3.8) is satisfied. In practice, this does not represent a major difficulty since one can start with conservative values and then "tune-up" the bound.

The parameter estimator can now be constructed using ordinary least squares (for example) incorporating a relative dead zone [5]-[8] to "protect" the algorithm against the unmodeled errors. An appropriate estimator is

$$\rho\hat{\theta} = \frac{aP\phi e}{1+\phi^TP\phi+\bar{C}\phi^T\phi}, \qquad \bar{C}>0 \tag{3.10}$$

$$\rho P = -\frac{aP\phi\phi^TP}{1+\phi^TP\phi+\bar{C}\phi^T\phi} \tag{3.11}$$

$$e = \bar{y} - \phi^T\hat{\theta} \tag{3.12}$$

and a implements a relative dead zone as follows.

Choose $\gamma > 0$, and let

$$\xi \triangleq \sqrt{\gamma+1}. \tag{3.13}$$

Then, with β as in (3.9)

$$a = \begin{cases} 0 & \text{if } |e| \leq \xi\beta \quad (3.14) \\ f[\xi\beta, e]/e & \text{otherwise} \quad (3.15) \end{cases}$$

where

$$f(g, e) \triangleq \begin{cases} e-g & \text{if } e>g \\ 0 & \text{if } |e| \leq g \\ e+g & \text{if } e<-g. \end{cases} \tag{3.16}$$

The properties of the above parameter estimator are, by now, standard and given in [8]. In summary, its key properties are as follows:

i) $\dfrac{f[\xi\beta, e]}{\sqrt{1+\phi^T\phi}} \in L_2$

ii) $\rho\hat{\theta} \in L_2$

iii) $\|\hat{\theta}-\theta_*\| \le \|\hat{\theta}_o-\theta_*\|$.

IV. Control Law Synthesis

We will use a pole placement design. However, unlike the result in [8] will not include the integrator in the design equation. Instead, we will add it afterwards as in [2] and [3]. We use the usual certainty equivalence principle [11]. Thus, let

$$\hat{\theta}^T \triangleq [\hat{\theta}_1, \cdots, \hat{\theta}_{2n}] \triangleq [e_o-\hat{a}_o, \cdots, e_{n-1}-\hat{a}_{n-1}, \hat{b}_o, \cdots, \hat{b}_{n-1}] \tag{4.1}$$

and define

$$\hat{A}=\rho^n+\hat{a}_{n-1}\rho^{n-1}+\cdots+\hat{a}_o \tag{4.2}$$

$$\hat{B}=\hat{b}_{n-1}\rho^{n-1}+\cdots+\hat{b}_o. \tag{4.3}$$

We then define $\hat{L}$ and $\hat{P}$ of degree n and $(n-1)$, respectively, by the following identity:

$$\hat{A}\hat{L}+\hat{B}\hat{P}=A^* \tag{4.4}$$

where A^* is an arbitrary Hurwitz polynomial of degree $2n$. (The issue of relative primeness of $\hat{A}$, $\hat{B}$ will be addressed in Section V.)

We next implement the feedback law conceptually as shown in Fig. 1 where the dotted line represents the retrofit integrator and ϵ^I is a small constant.

Using filtered signals, the control law of Fig. 1 can be implemented as

$$\hat{L}\rho\left(\frac{1}{EQ}u\right)=-(\hat{P}\rho+\epsilon^I)\left(\frac{1}{EQ}z\right); \qquad z=y-y^* \tag{4.5}$$

or

$$\hat{L}u_f=-(\hat{P}\rho+\epsilon^I)z_f'; \; u_f=\frac{\rho}{EQ}u; \; z_f'=\frac{1}{EQ}z. \tag{4.6}$$

Actually, there are several other ways that the integrator might be added to the feedback system, e.g., as in Fig. 1 but with the summation point after $1/\hat{L}$. We have found the setup of Fig. 1 to be most suitable in terms of the stability analysis.

One final key point is that we must choose the sign of ϵ^I correctly. The analysis in [2], [3] for the nonadaptive case, suggests the following choice:

$$\text{sign } \epsilon^I=\text{sign } \hat{b}_o. \tag{4.7}$$

V. Stabilizability of the Estimated Model

In common with all analyses of adaptive control laws to date we require stabilizability of the estimated models. Also, since our control law involves integral action, we need to ensure that the estimated model does not have a zero at the origin. Thus, we require $|\hat{b}_o|$ to be bounded away from zero.

These two requirements can be achieved by running parallel parameter estimator as in [8]. Each parameter estimator ensures that the corresponding estimates lie in a convex region, say D_i, for the ith estimator. It is then simply required that:

i) $\theta_* \in \bigcup_{i=1}^{p} D_i$

ii) for all $\theta \in \bigcup_{i=1}^{p} D_i$, A and B are uniformly relatively prime and b_o is bounded away from zero.

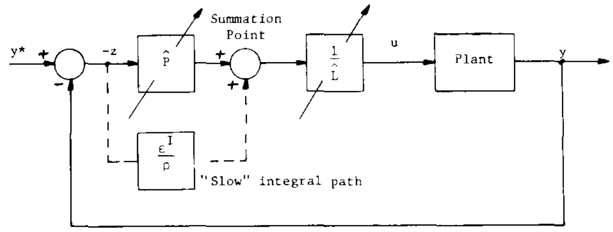

Fig. 1. Conceptual implementation of feedback law.

We then employ a projection as in [11] to ensure that $\hat{\theta}_i \in D_i$. Also, for each estimator we monitor

$$M_i(t)=\int_o^t \frac{f[\xi\beta, e_i]^2}{1+\phi^T P\phi}\, d\tau \tag{5.1}$$

and basically choose that estimator which has the lowest value of $M_i(t)$. (Actually, some hysteresis is needed to avoid switching back and forth between estimators infinitely often—see [8] for details.)

It is readily shown [8] that this selection procedure:

a) retains the standard properties of the estimators;

b) ensures that there exists some time t_o after which only one estimator is selected; and

c) ensures that each estimated model is uniformly stabilizable with $\hat{b}_o$ bounded away from zero; thus

d) in view of b), c), and (4.7), ϵ^I is constant for all $t \ge t_o$.

VI. Stability Analysis of the Adaptive Control Law

The key equations describing the adaptive feedback law are as follows:

i) the prediction error equation (3.14) which can be rewritten as

$$\hat{A}y_f=\hat{B}u_f+e \tag{6.1}$$

ii) the feedback control law (4.6)

$$\hat{L}u_f=-(\hat{P}\rho+\epsilon^I)z_f' \tag{6.2}$$

iii) the design identity (4.4)

$$\hat{A}\hat{L}+\hat{B}\hat{P}=A^* \tag{6.3}$$

(where the coefficients of $\hat{L}$, $\hat{P}$ are guaranteed bounded by the multiple estimator strategy). Noting that $\rho y^*=0$, (6.1), (6.2) can be rewritten in state-space form as

$$\rho x=A(t)x+B_1 e \tag{6.4}$$

where

$$A(t)=\begin{bmatrix} -\hat{a}_{n-1}, & \cdots, & -\hat{a}_o, & 0, & \hat{b}_{n-1}, & \cdots, & \hat{b}_o \\ 1 & & \vdots & & & 0 & \\ & 1 & 0 & & & & \\ -\hat{P}_{n-1}, & \cdots, & -\hat{P}_o, & \epsilon^I, & -\hat{l}_{n-1}, & \cdots, & -\hat{l}_o \\ & & & & 1 & & \\ & 0 & & & & 1 & 0 \end{bmatrix}; \quad B_1=\begin{bmatrix}1\\0\\0\end{bmatrix}$$

$$x(t)^T \triangleq [\rho^n z_f', \cdots, z_f', \rho^{n-1}u_f, \cdots, u_f]. \tag{6.6}$$

The time-varying matrix $A(t)$ can be readily seen to have eigenvalues which are the $(2n+1)$ zeros of the polynomial $(A^*\rho+\epsilon^I\hat{B})$. In order that we can apply the proof paradigm of [8], we require that the matrix has all its eigenvalues in the strict left-half plane for each fixed time instant. We therefore establish the following lemma.

Lemma 6.1: For some fixed ϵ^I sufficiently small, the zeros of the polynomial $(A^*\rho+\epsilon^I\hat{B})$ are in the strict left-half plane.

Proof: This result is actually implicit in [2] but can be easily established in the SISO case as follows. Property iii) of the parameter estimator implies that the coefficients of $\hat{B}$ are bounded. Also, by design, A^* has bounded coefficients and zeros in the strict left-half plane. If we now form the Routh-Hurwitz array for $A^*\rho + \epsilon^I\hat{B}$ we find that the left column of the array is (G_i'; $i = 1, 2, \cdots, 2n + 2$) where

$$G_i' = \begin{cases} G_i; & i=1, 2, \cdots, n+2 \\ G_i + 0(\epsilon^I); & i=n+3, \cdots, (2n+1) \\ \epsilon^I\hat{b}_o; & i=2n+2 \end{cases}$$

where (G_i; $i = 1, 2, \cdots, 2n + 1$) is the left column of the Routh-Hurwitz array of A^* and where $\lim_{\epsilon^I \to 0} 0(\epsilon^I) = 0$ (follows from the boundedness of $\hat{B}$'s coefficients).

Since the polynomial A^* has its zeros in the strict left-half plane, we know that each G_i is positive. Thus, since sign ϵ^I = sign $\hat{b}_o$ and $|\hat{b}_o|$ is bounded away from zero, we can choose $\epsilon^I > 0$ (independent of time) such that G_i' is positive for all time. However, this is a necessary and sufficient condition for $(A^*\rho + \epsilon^I\hat{B})$ to have its zeros in the strict left-half plane. □

Thus, we see that for each frozen time instant the matrix $A(t)$ in (6.4) has eigenvalues in the strict left-half plane. Inspection of the proof in [8] shows that this is the key requirement for convergence of the adaptive control algorithm. Thus, using the proof paradigm of [8] we have

Theorem 6.1: Subject to Assumption A.1 and provided

i) (3.8) is satisfied for some ϵ, ν, E, Q, σ_o

ii) ϵ^I and ϵ are sufficiently small (both depending on A, ν, E, Q, σ_o, A^*), then:

a) all signals remain bounded; and

b) $\lim_{t\to\infty} y - y^* = 0$.

VII. Conclusions

It is known that in classical control, it is permissible in certain circumstances to retrofit an integrator to an otherwise stable design. This note has shown that a similar procedure is valid for adaptive control. A simple strategy has been described for choosing the integral constant and global convergence has been established for the resultant algorithm.

Acknowledgment

The research reported here was made possible by a visit, of the second author, to Haifa in November 1986.

References

[1] N. Sandell and M. Athans, "On type L multivariable linear systems," *Automatica*, vol. 9, pp. 131-136, 1973.

[2] M. Morari, "Robust stability of systems with integral action," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 574-577, June 1985.

[3] C. A. Desoer and C. A. Lin, "Tracking and disturbance rejection of MIMO nonlinear systems with PI controller," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 861-867, 1985.

[4] P. A. Ioannou and K. Tsakalis, "A robust direct adaptive controller," *IEEE Trans. Automat. Contr.*, to be published.

[5] G. Kreisselmeier and B. D. O. Anderson, "Robust model reference adaptive control," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 127-133, 1986.

[6] G. Kreisselmeir, "A robust indirect adaptive control approach," *Int. J. Contr.*, vol. 43, no. 1, pp. 161-175, 1986.

[7] K. S. Narendra and A. M. Annaswamy, "Robust adaptive control using reduced order models," in *Proc. 2nd IFAC Workshop on Adaptive Syst.*, 1986, pp. 49-54.

[8] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne, "Design issues in adaptive control," *IEEE Trans. Automat. Contr.*, vol. 33, Jan. 1988.

[9] C. A. Desoer, R.-W. Lin, and R. Sacks, "Feedback system design—The fractional representation approach to analysis and synthesis," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 3, pp. 399-412, 1980.

[10] G. C. Goodwin and D. Q. Mayne, "A parameter estimation perspective of continuous time model reference adaptive control," *Automatica*, Jan. 1987.

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

Recovering the Poles from Third-Order Cumulants of System Output

JITENDRA K. TUGNAIT

Abstract—**The problem of identifying the poles of single-input/single-output (SISO) linear stochastic systems from the higher order statistics of noisy observations is considered. It is assumed that the system is driven by an independent and identically distributed non-Gaussian process with nonzero third-order cumulant function at zero lag. There is no other restriction on the probability distribution of the driving noise. The system is assumed to be of known order, causal, and exponentially stable, but is not required to be minimum phase. The system output is observed in additive, possibly non-Gaussian, noise. We show that if there are no pole-zero cancellations in the transfer function of the given system, then it is necessary and sufficient for a block Hankel matrix to have rank equal to the system order where the matrix is constructed from a partial set of third-order cumulants of the noisy output sequence. This fundamental result then leads to a linear solution to the problem of estimating the coefficients of the system characteristic polynomial from which the system poles can be found via root-finding.**

I. Introduction

A vast majority of the literature on system identification and parameter estimation *using only the output data* is restricted to minimum-phase system models [5]. However, there are several cases of practical interest where the underlying signal/system model is either causal nonminimum-phase or noncausal; see, e.g., [1], [2], [6]-[14]. System identification for such models may be accomplished by the use of the higher order statistics [1], [2], [6]-[14]. The area of parametric modeling via cumulant statistics has attracted considerable attention in recent years; for a tutorial and a perspective, see [7] and [8], respectively, where further references may be found.

This note is concerned with the problem of recovering the poles of a causal stable ARMA (autoregressive moving average) model of known order from the third-order statistics of the system output. We show that if there are no pole-zero cancellations in the transfer function of the given system, then it is necessary and sufficient for a block Hankel matrix to have rank equal to the system order where the matrix is constructed from a partial set of third-order cumulants of the noisy output sequence. This fundamental result then leads to a linear solution to the problem of estimating the AR coefficients from which the system poles can then be obtained via root-finding.

The problem of linear estimation of the AR coefficients of an ARMA model from the higher order statistics of the system output has not been satisfactorily addressed so far; previous attempts include [1, proof of Lemma 6] (also repeated in [10, proof of Lemma 3] and [13, proof of Lemma 3]), [11], and [12]. The approach of [11] is flawed as evidenced by our (counter-) example in Section III (see also Remark 1). In [1] it has been claimed that there exists a subset of output cumulants that yield the AR coefficients; the proof is by contradiction and it invokes some results of Lii and Rosenblatt [14] which require that the system transfer function be nonzero at zero frequency. As discussed in Remark 2 in Section III, the arguments of [1] are flawed; the results of [1] remain true, however. In [12] the same flaw recurs since the arguments used are the same as in [1]. Our results proved in Section III do not require that the system transfer function be nonzero at zero frequency.

The problem of pole recovery from fourth-order statistics has been addressed in [3] and [9] following the approach of this note.

The fundamental results pertaining to the recovery of the poles of a causal stable ARMA model from the third-order statistics of the process

Manuscript received June 28, 1988; revised September 21, 1988. This paper is based on a prior submission of August 19, 1987.

The author is with Exxon Production Research Company, Houston, TX 77252-2189.

IEEE Log Number 8929490.

