# Convergence Properties of an Extended Least Squares (ELS) Variant*

R. HENRIKSEN†

By factorizing the $A$- and $B$-polynomials in an ARMAX model and filtering the input/output data with the appropriate factors thereof, the parameters of the model can be estimated in a decentralized fashion. This may improve the robustness of some estimators significantly, e.g., when applied to very stiff systems. Earlier work on these techniques has established both local and global properties for some LS, IV, and PE variants. An ELS variant has, however, never been considered, and a variant of this type is introduced. Local and global convergence properties of this variant are analyzed.

## 1. Introduction

The extended least squares (ELS) method, also known as pseudolinear regression (PLR) or the approximate maximum likelihood (ML) method is a well-known and popular identification method which often is employed for ARMAX models. Consistency and convergence properties of this method were derived quite some years ago, and in addition to the usual consistency conditions which apply to all LS methods, some additional conditions have to be satisfied in order to ensure convergence of the parameters to the true values. These conditions put restrictions on the $C$-polynomial in the ARMAX model, and for *local* convergence a sufficient condition is that $C$ be strictly positive real (in the appropriate frequency domain), whereas for *global* convergence a sufficient condition is that $[1/C - 1/2]$ be strictly positive real, the latter condition naturally being more restrictive.

Parameter estimators based upon standard estimation techniques, viz. least squares (LS) methods, instrumental variable (IV) methods, and prediction error (PE) methods, or adaptive controllers based upon them, do occasionally have difficulties with systems that have a somewhat ill-conditioned nature, e.g. stiff systems. By prefiltering the input/output data in a specific way, utilizing factors of the $A$- and $B$-polynomials, and estimating the parameters in a decoupled fashion it is, however, possible to improve significantly the robustness of the estimators and of their ultimate use in control, i.e. in controller design or as part of an adaptive controller.

Another application of these estimation techniques could be the following. Assume the underlying process, described by an ARMAX model, contains a very slowly varying part and or more rapidly changing one, the latter needing continuous updating. By factorizing the polynomials in accordance with these two parts, the slowly varying one, described by factors of the $A$- and $B$-polynomials, can be 'filtered out' in order to get a lower order model which often is more robust both for estimation purposes and for controller design, see, e.g., Clary and Franklin (1984) where this concept is utilized in connection with self-tuning control. Whenever the slowly changing model

needs updating, the full model estimator can be employed using the schemes described herein.

Previous papers on these methods have treated robustness and local convergence properties, see Henriksen (1988, 1989) and Young *et al.* (1987), global convergence properties, see Henriksen and Weyer (1990) and Weyer (1991), and accuracy aspects, see Henriksen (1991, 1992). In those references it is shown that local convergence can be guaranteed for all the variants (LS, IV, and PE) provided the $A$- and $B$-polynomials were factorized into coprime factors, whereas global convergence could be guaranteed for the LS variants provided the same coprimeness conditions were satisfied.

In this paper an ELS variant based upon these techniques is introduced, and its convergence properties are analyzed. Since the LS variants turned out to have the same convergence properties as the ordinary LS method, provided the coprimeness conditions were satisfied, it was expected that the ELS variant would possess the same convergence properties as the ordinary ELS method. It turned out, however, that this is not quite true.

The paper is organized as follows. In Section 2 a brief outline of the system and the estimator is presented, including previous results concerning local and global convergence properties, robustness and accuracy. Section 3 is devoted to analysis of local and global convergence of the particular ELS variant presented herein. Concluding remarks are given in Section 4.

## 2. System description and previous results

We consider a system described by the ARMAX model

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t \tag{1}$$

where $y_t$ is the output at time $t$, $u_t$ is the input, whereas $e_t$ is (zero-mean) white noise. The processes $\{u_t\}$ and $\{e_t\}$ are assumed to be independent, and the system assumed to be asymptotically stable (the latter being assumed only for the purpose of simplifying the analysis). The polynomials $A$ and $B$ are written in the following factorized forms, viz.

$$A(q^{-1}) = A_1(q^{-1})A_2(q^{-1}) \tag{2}$$

$$B(q^{-1}) = B_1(q^{-1})B_2(q^{-1}) \tag{3}$$

where

$$A_i(q^{-1}) = 1 + a_1^i q^{-1} + a_2^i q^{-2} + \ldots + a_{n_i}^i q^{-n_i} \qquad i = 1, 2 \tag{4}$$

$$B_i(q^{-1}) = 1 + b_1^i q^{-1} + b_2^i q^{-2} + \ldots + b_{m_i}^i q^{-m_i} \qquad i = 1, 2 \tag{5}$$

and where $b_0^1 = 1$, $b_0^2 = 0$. Moreover, $n_1 + n_2 = n_a$, the degree of $A(q^{-1})$, whereas $m_1 + m_2 = n_b$, the degree of $B(q^{-1})$. The $C$-polynomial is assumed to be of the form

$$C(q^{-1}) = 1 + c_1 q^{-1} + c_2 q^{-2} + \ldots + c_{n_c} q^{-n_c} \tag{6}$$

Assuming for a moment the two polynomials $A_2(q^{-1})$ and $B_2(q^{-1})$ to be known, we can define two new variables which are moving averages of respectively the output and the input of the system, viz.

$$w_t = A_2(q^{-1})y_t \qquad r_t = B_2(q^{-1})u_t \tag{7}$$

and (1) thus takes the form

$$A_1(q^{-1})w_t = B_1(q^{-1})r_t + C(q^{-1})e_t \tag{8}$$

This model will in the sequel be referred to as model $\mathcal{M}_1$.

Similarly, assuming for a moment the polynomials $A_1(q^{-1})$ and $B_1(q^{-1})$ to be known, we can define two other variables which also are moving averages of respectively the output and the input, viz.

$$z_t = A_1(q^{-1})y_t \qquad s_t = B_1(q^{-1})u_t \tag{9}$$

and (1) in this case take the form

$$A_2(q^{-1})z_t = B_2(q^{-1})s_t + C(q^{-1})e_t \tag{10}$$

This model will in the sequel be referred to as model $\mathcal{M}_2$.

Equations (9)–(12) form the basis for the estimators which have been treated in the aforementioned papers and for the estimator considered herein. Note, however, that estimation of the $C$-polynomial was never considered in the previous papers. This implies that $C(q^{-1})) = 1$ had to be assumed for the LS variants, whereas $C(q^{-1}) \neq 1$ could be allowed for the IV variants.

We rewrite model $\mathcal{M}_1$, i.e. (10) in the form

$$w_t = \psi_t^T \beta + r_t + e_t \tag{11}$$

where

$$\psi_t = [-w_{t-1} - w_{t-2} \ldots - w_{t-n_1} r_{t-1} r_{t-2} \ldots r_{t-m_1} e_{t-1} e_{t-2} \ldots e_{t-n_c}]^T \tag{12}$$

$$\beta = [a_1^1 a_2^1 \ldots a_{n_1}^1 b_1^1 b_2^1 \ldots b_{m_1}^1 c_1 c_2 \ldots c_{n_c}]^T \tag{13}$$

An ELS estimator for model $\mathcal{M}_1$ will take the (batch) form

$$\hat{\beta} = \left[ \sum_{t=1}^N \psi_t \psi_t^T \right]^{-1} \left[ \sum_{t=1}^N \psi_t (w_t - r_t) \right] = \left[ \frac{1}{N} \sum_{t=1}^N \psi_t \psi_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \psi_t (w_t - r_t) \right] \tag{14}$$

where we have to utilize

$$\psi_t = [-w_{t-1} \ldots - w_{t-n_1} r_{t-1} \ldots r_{t-m_1} \varepsilon_{t-1} \ldots \varepsilon_{t-n_c}]^T \tag{15}$$

because the $e$-terms are unknown, but can be estimated from

$$\varepsilon_t = w_t - \psi_t^T \hat{\beta} - r_t \tag{16}$$

Similarly, we can write model $\mathcal{M}_2$, i.e. (12) in the form, assuming the $C$-polynomial and the $e$-terms to be known,

$$z_t = \phi_t^T \theta + [C(q^{-1}) - 1]e_t + e_t \tag{17}$$

where

$$\phi_t = [-z_{t-1} - z_{t-2} \ldots - z_{t-n_2} s_{t-1} s_{t-2} \ldots s_{t-m_2}]^T \tag{18}$$

$$\theta = [a_1^2 a_2^2 \ldots a_{n_2}^2 b_1^2 b_2^2 \ldots b_{m_2}^2]^T \tag{19}$$

An ordinary LS estimator for model $\mathcal{M}_2$ now takes the (batch) form

$$\hat{\theta} = \left[ \sum_{t=1}^N \phi_t \phi_t^T \right]^{-1} \left[ \sum_{t=1}^N \phi_t (z_t - [C(q^{-1}) - 1]e_t) \right]$$

$$= \left[ \frac{1}{N} \sum_{t=1}^N \phi_t \phi_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \phi_t (z_t - [C(q^{-1}) - 1]e_t) \right] \tag{20}$$

We could of course have chosen to estimate the $C$-polynomial with model $\mathcal{M}_2$, but the whole matter is in this respect symmetric.

Now, the above estimators are both based on the assumption that factors in the $A$- and $B$-polynomials are known. This is normally not true, i.e. generally we will have to estimate both factors of the $A$- and $B$-polynomials. In order to utilize the above, we therefore have to calculate the moving averages of the input and the output by using *estimates* of the appropriate factors, i.e., we calculate the moving averages as

$$w_t(=w_t(\hat{\theta})) = \hat{A}_2(q^{-1})y_t \qquad r_t(=r_t(\hat{\theta})) = \hat{B}_2(q^{-1})u_t \tag{21}$$

$$z_t(=z_t(\hat{\beta})) = \hat{A}_1(q^{-1})y_t \qquad s_t(=s_t(\hat{\beta})) = \hat{B}_1(q^{-1})u_t \tag{22}$$

where

$$\hat{A}_1(q^{-1}) = A_1(q^{-1};\hat{\beta}) \qquad \hat{B}_1(q^{-1}) = \hat{B}_1(q^{-1};\hat{\beta}) \tag{23}$$

$$\hat{A}_2(q^{-1}) = A_2(q^{-1};\hat{\theta}) \qquad \hat{B}_2(q^{-1}) = B_2(q^{-1};\hat{\theta}) \tag{24}$$

Furthermore, the $\varepsilon$-terms can be calculated from

$$\varepsilon_t(=\varepsilon_t(\hat{\beta},\hat{\theta})) = w_t(\hat{\theta}) - r_t(\hat{\theta}) - \psi_t^T\hat{\beta} \tag{25}$$

where

$$\psi_t(=\psi_t(\hat{\beta},\hat{\theta})) = [-w_{t-1}(\hat{\theta})\ldots r_{t-1}(\hat{\theta})\ldots \varepsilon_{t-1}(\hat{\beta},\hat{\theta})\ldots \varepsilon_{t-n_c}(\hat{\beta},\hat{\theta})]^T \tag{26}$$

Likewise, we will make use of

$$\phi_t(=\phi_t(\hat{\beta})) = [-z_{t-1}(\hat{\beta})\ldots s_{t-1}(\hat{\beta})\ldots s_{t-m_2}(\hat{\beta})]^T \qquad \hat{C}(q^{-1}) = C(q^{-1};\hat{\beta}) \tag{27}$$

We now write down the expressions for the two estimators with the modifications introduced above:

$$\hat{\beta} = \left[\frac{1}{N}\sum_{t=1}^{N}\psi_t(\hat{\beta},\hat{\theta})\psi_t^T(\hat{\beta},\hat{\theta})\right]^{-1}\left[\frac{1}{N}\sum_{t=1}^{N}\psi_t(\hat{\beta},\hat{\theta})\{w_t(\hat{\theta}) - r(\hat{\theta})\}\right] \tag{28}$$

$$\hat{\theta} = \left[\frac{1}{N}\sum_{t=1}^{N}\phi_t(\hat{\beta})\phi_t^T(\hat{\beta})\right]^{-1}\left[\frac{1}{N}\sum_{t=1}^{N}\phi_t(\hat{\beta})\{z_t - [C(q^{-1};\hat{\beta}) - 1]\varepsilon_t(\hat{\beta},\hat{\theta})\}\right] \tag{29}$$

Of course, in the form depicted above, constituting a set of implicit equations in the unknowns $\hat{\beta}$ and $\hat{\theta}$, the common way to solve the system is to do it iteratively, i.e., to implement the above as a bootstrap estimator. However, by doing a few approximations we can write down the above equations in a recursive fashion which greatly simplifies the computational effort.

In the above scheme, filtering of the input/output data and computation of the $\varepsilon$-terms have to be carried out using the *final* values of the appropriate polynomials. In a recursive scheme this would for all practical purposes be an impossible task to do. We therefore introduce the following recursive approximation to the above.

*Step 1*

At time instant $t$ compute

$$w_t = A_2(q^{-1};\hat{\theta}_{t-1})y_t \qquad r_t = B_2(q^{-1};\hat{\theta}_{t-1})u_t \tag{30}$$

$$z_t = A_1(q^{-1};\hat{\beta}_{t-1})y_t \qquad s_t = B_1(q^{-1};\hat{\beta}_{t-1}) \qquad (31)$$

$$\varepsilon_t = w_t - r_t - \psi_t^T \hat{\beta}_{t-1} \qquad (32)$$

### Step 2

Compute

$$P_t = P_{t-1} + \frac{1}{t}\{\psi_t \psi_t^T - P_{t-1}\} \qquad (33)$$

$$\hat{\beta}_t = \hat{\beta}_{t-1} + \frac{1}{t}P_t^{-1}\psi_t\{w_t - r_t - \psi_t^T\hat{\beta}_{t-1}\} \qquad (34)$$

$$Q_t = Q_{t-1} + \frac{1}{t}\{\phi_t\phi_t^T - Q_{t-1}\} \qquad (35)$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{t}Q_t\phi_t\{z_t - [C(q^{-1};\hat{\beta}_{t-1})-1]\varepsilon_t - \phi_t^T\hat{\theta}_{t-1}\} \qquad (36)$$

### Step 3

Replace $(\psi_t,\phi_t)$ by $(\psi_{t+1},\phi_{t+1})$ using the values computed in (30) and (31), increase $t$ by 1, and return to Step 1.

In the above recursive scheme, note that we instead of (32) can compute $\varepsilon_t$ from two other alternative expressions, viz.

$$\varepsilon_t = z_t - [C(q^{-1};\hat{\beta}_{t-1})-1]\varepsilon_t - \phi_t^T\hat{\theta}_{t-1} = y_t - [\psi_t^T \phi_t^T]\begin{bmatrix}\hat{\beta}_{t-1}\\\hat{\theta}_{t-1}\end{bmatrix} \qquad (37)$$

(34) and (36) can be written in the following simple forms, respectively,

$$\hat{\beta}_t = \hat{\beta}_{t-1} + \frac{1}{t}P_t^{-1}\psi_t\varepsilon_t \qquad \hat{\theta}_t = \hat{\theta}_{t-1} + \frac{1}{t}Q_t^{-1}\phi_t\varepsilon_t \qquad (38)$$

Moreover, convergence of the above scheme can significantly be speeded up in the following ways, (i) by moving (31) down immediately after (34) and changing the expressions into

$$z_t = A_1(q^{-1};\hat{\beta}_t)y_t \qquad s_t = B_1(q^{-1};\hat{\beta}_t) \qquad (39)$$

and (ii) by computing a new, and presumably improved, value of $\varepsilon_t$ immediately after (34) for use in the sequel, i.e. by recomputing it according to

$$\varepsilon_t = w_t - \psi_t\hat{\beta}_t - r_t \qquad (40)$$

In order to accommodate (36) to these modifications we have to replace $C(q^{-1};\hat{\beta}_{t-1})$ by $C(q^{-1};\hat{\beta}_t)$.

We can write the scheme given by (30)–(36) in a more compact form as follows. Define

$$\kappa_t = \begin{bmatrix}\beta_t\\\theta_t\end{bmatrix} \qquad R_t = \begin{bmatrix}P_t & 0\\0 & Q_t\end{bmatrix} \qquad (41)$$

$$\Pi_t = \begin{bmatrix} \psi_t & 0 \\ 0 & \phi_t \end{bmatrix} \qquad p_t = \begin{bmatrix} w_t - r_t \\ z_t - [\hat{C}_{t-1} - 1]\varepsilon_t \end{bmatrix} \tag{42}$$

where we have introduced the simplified notation $\hat{C}_{t-1} = C(q^{-1}; \hat{\beta}_{t-1})$. The scheme can then be written in the form given below, omitting obvious details which are given before.

*Step 1*

At time instant $t$ compute

$$\varepsilon_t = w_t - r_t - \psi_t^T \hat{\beta}_{t-1} = y_t - [\psi_t^T \, \phi_t^T] \begin{bmatrix} \hat{\beta}_{t-1} \\ \hat{\theta}_{t-1} \end{bmatrix} = y_t - [\psi_t^T \, \phi_t^T] \hat{\kappa}_{t-1} \tag{43}$$

*Step 2*

$$R_t = R_{t-1} + \frac{1}{t}[\Pi_t \Pi_t^T - R_{t-1}] \tag{44}$$

$$\hat{\kappa}_t = \hat{\kappa}_{t-1} + \frac{1}{t} R_t^{-1} \Pi_t [p_t - \Pi_t^T \hat{\kappa}_{t-1}] \tag{45}$$

*Step 3*

Replace $\Pi_t$ by $\Pi_{t+1}$, increase $t$ by 1, and return to Step 1.

The above scheme will be used in the convergence analysis which will be the topic for the next section, but before we go into that some previous results about LS ('ordinary', i.e. not extended), IV, and PE estimators based upon the polynomial factorization described herein are summarized in what follows.

*Fact*

*The bootstrap LS, IV, and PE variants converge <u>locally</u>, and the LS variant converges <u>globally</u>, if and only if the polynomials $A_1(q^{-1})$, $A_2(q^{-1})$, $B_1(q^{-1})$, and $B_2(q^{-1})$ are all mutually coprime.*

*As the stiffness of the system tends to infinity, local convergence of the bootstrap LS, IV, and PE variants tends to becoming instantaneous.*

*The accuracy of the bootstrap LS and IV variants, with input $u_t$ is the same as the accuracy of the corresponding ordinary LS and IV methods, respectively, for the reduced-order models $\mathcal{M}_1$ and $\mathcal{M}_2$ with inputs $B_2(q^{-1})u_t$ and $B_1(q^{-1})u_t$, respectively.*

## 3.  Convergence analysis of the ELS variant

The convergence analysis, both locally and globally, will be based upon the ODE approach introduced by Ljung (1977), see also Ljung (1987) or Ljung and Söderström (1983). The associated differential equations to (44)–(45) have the form

$$\frac{d}{d\tau} \kappa(\tau) = R^{-1}(\tau) f(\kappa(\tau)) \tag{46}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau} R(\tau) = G(\kappa(\tau)) - R(\tau) \tag{47}$$

where

$$f(\kappa) = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E\left[ \Pi_t(\kappa) \left[ \begin{array}{c} w_t(\theta) - r_t(\theta) - \psi_t^{\mathrm{T}}(\kappa)\beta \\ z_t(\beta) - [C(q^{-1};\beta) - 1]\varepsilon_t(\kappa) - \phi_t^{\mathrm{T}}(\beta)\theta \end{array} \right] \right]$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E\left[ \begin{array}{c} \psi_t(\kappa)\{w_t(\theta) - r_t(\theta) - \psi_t^{\mathrm{T}}(\kappa)\beta\} \\ \phi_t(\beta)\{z_t(\beta) - [C(q^{-1};\beta) - 1]\varepsilon_t(\kappa) - \phi_t^{\mathrm{T}}(\beta)\theta\} \end{array} \right] \tag{48}$$

$$G(\kappa) = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E[\Pi_t(\kappa)\Pi_t^{\mathrm{T}}(\kappa)]$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E\left[ \begin{array}{cc} \psi_t(\kappa)\psi_t^{\mathrm{T}}(\kappa) & 0 \\ 0 & \phi_t(\beta)\phi_t^{\mathrm{T}}(\beta) \end{array} \right] \tag{49}$$

For the sake of simplicity, assuming the processes involved to be ergodic (which, of course, requires the assumption about the system being stable), the above expressions simply become

$$f(\kappa) = E\left[ \left[ \begin{array}{c} \psi_t(\kappa)\{w_t(\theta) - r_t(\theta) - \psi_t^{\mathrm{T}}(\kappa)\beta\} \\ \phi_t(\beta)\{z_t(\beta) - [C(q^{-1};\beta) - 1]\varepsilon_t(\kappa) - \phi_t^{\mathrm{T}}(\beta)\theta\} \end{array} \right] \right] \tag{50}$$

$$G(\kappa) = E[\Pi_t(\kappa)\Pi_t^{\mathrm{T}}(\kappa)] = E\left[ \begin{array}{cc} \psi_t(\kappa)\psi_t^{\mathrm{T}}(\kappa) & 0 \\ 0 & \phi_t(\beta)\phi_t^{\mathrm{T}}(\beta) \end{array} \right] \tag{51}$$

Now, assume sufficient conditions for the matrix $R(\tau)$ to be invertible are given, e.g., that the input process is pe. of a sufficient order and that the coprimeness conditions stated in the Fact are satisfied, see Henriksen and Weyer (1990). Let $(\kappa^*, R^*)$ denote a point of convergence for (46)–(47). It follows that

$$0 = (R^*)^{-1} f(\kappa^*) \Rightarrow f(\kappa^*) = 0 \tag{52}$$

$$0 = G(\kappa^*) - R^* \Rightarrow G(\kappa^*) = R^* \tag{53}$$

which shows what has to be satisfied in a possible convergence point.

*Local Convergence Analysis*

Linearizing (46)–(47) about $(\kappa^*, R^*)$ yields, where $\mathbf{r} = \mathrm{vec}\, R$,

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \left[ \begin{array}{c} \kappa(\tau) - \kappa^* \\ \mathbf{r}(\tau) - \mathbf{r}^* \end{array} \right] = \left[ \begin{array}{cc} G^{-1}(\kappa^*)\dfrac{\partial f}{\partial \kappa^{\mathrm{T}}}(\kappa^*) & 0 \\ \mathbf{N}_t & -I \end{array} \right] \left[ \begin{array}{c} \kappa(\tau) - \kappa^* \\ \mathbf{r}(\tau) - \mathbf{r}^* \end{array} \right] \tag{54}$$

where the matrix $\mathbf{N}_t$ is of no interest because it has no influence on the eigenvalues of the system matrix. Local convergence, i.e., local stability about the convergence point,

can therefore be determined from the stability of

$$\frac{d}{d\tau}[\kappa(\tau) - \kappa^*] = G^{-1}(\kappa^*)\tilde{G}(\kappa^*)[\kappa(\tau) - \kappa^*] = L(\kappa^*)[\kappa(\tau) - \kappa^*] \quad (55)$$

where we have defined

$$\tilde{G}(\kappa^*) = \frac{\partial f}{\partial \kappa^T}(\kappa^*) \qquad L(\kappa^*) = G^{-1}(\kappa^*)\tilde{G}(\kappa^*) \quad (56)$$

We want to investigate local convergence about the true parameter vector, so let

$$\kappa^0 = \begin{bmatrix} \beta^0 \\ \theta^0 \end{bmatrix} \quad (57)$$

denote this vector. Local convergence about the true parameter vector $\kappa^0$ is thus determined from the stability of the system

$$\frac{d}{d\tau}[\kappa(\tau) - \kappa^0] = L(\kappa^0)[\kappa(\tau) - \kappa^0] \quad (58)$$

where

$$L(\kappa^0) = G^{-1}(\kappa^0)\tilde{G}(\kappa^0) \qquad \tilde{G}(\kappa^0) = \frac{\partial f}{\partial \kappa^T}(\kappa^0) \qquad G(\kappa^0) = E[\Pi(\kappa^0)\Pi^T(\kappa^0)] \quad (59)$$

We are now ready to present what follows.

**Theorem** (Local convergence of the ELS variant)
*The ELS variant converges* <u>*locally*</u> *if the polynomial* $C(q^{-1}; \beta^0)$, *i.e., the true C-polynomial, is strictly positive real for* $\omega \in [-\pi, \pi]$.

*Proof.* A quite detailed proof is found in the Appendix.

□

Readers familiar with the ELS method will recognize this as exactly the same *sufficient* condition for local convergence of the ordinary ELS method, see Söderström and Stoica (1989), Wellstead and Zarrop (1991).


*Global Convergence Analysis*
Now, having established that a sufficient condition for local stability of the ELS variant presented herein is exactly the same as for local stability of the ordinary ELS estimator (provided, of course, that the coprimeness conditions given before are satisfied),we will now to try to establish sufficient conditions for global stability.
First, consider the Lyapunov function

$$V(\kappa(\tau), R(\tau)) = [\kappa(\tau) - \kappa^0]^T R(\tau)[\kappa(\tau) - \kappa^0] \quad (60)$$

which is the one used in Söderström and Stoica (1989) for global analysis of the ordinary ELS method. Straightforward differentiation of $V$ with respect to $\tau$ yields

$$\frac{d}{d\tau} V(\kappa(\tau), R(\tau)) = f^T(\kappa(\tau))[\kappa(\tau) - \kappa^0] + [\kappa(\tau) - \kappa^0]^T f(\kappa(\tau))$$

$$+ [\kappa(\tau) - \kappa^0]^T [G(\kappa(\tau)) - R(\tau)][\kappa(\tau) - \kappa^0] \quad (61)$$

where

$$f(\kappa) = E\left[ \Pi_t(\kappa)\left[ \begin{array}{c} w_t(\theta) - r_t(\theta) - \psi_t^T(\kappa)\beta \\ z_t(\beta) - [C(q^{-1};\beta) - 1]\varepsilon_t(\kappa) - \phi_t^T(\beta)\theta \end{array} \right] \right]$$

$$= E\left[ \begin{array}{c} \psi_t(\kappa)[w_t(\theta) - r_t(\theta) - \psi_t^T(\kappa)\beta] \\ \phi_t(\beta)[z_t(\beta) - [C(q^{-1};\kappa) - 1]\varepsilon_t(\kappa) - \phi_t^T(\beta)\theta] \end{array} \right] = E\left[ \begin{array}{c} \psi_t(\kappa)\varepsilon_t(\kappa) \\ \phi_t(\beta)\varepsilon_t(\kappa) \end{array} \right] \quad (62)$$

Since none of the components of $\psi_t(\kappa)$ and $\phi_t(\beta)$ are correlated with $e_t$ we can write $f(\kappa)$ as

$$f(\kappa) = E\left[ \begin{array}{c} \psi_t(\kappa)[\varepsilon_t(\kappa) - e_t] \\ \phi_t(\beta)[\varepsilon_t(\kappa) - e_t] \end{array} \right] \quad (63)$$

Now, trying the approach used in Söderström and Stoica (1989) we obtain

$$\varepsilon_t(\kappa) - e_t = \{y_t - [\psi_t^T(\kappa)\phi_t^T(\kappa)]\kappa\} - \{y_t - [\psi_t^T(\kappa^0)\phi_t^T(\beta^0)]\kappa^0\}$$

$$= -[\psi_t^T(\kappa)\phi_t^T(\beta)][\kappa - \kappa^0] - \{[\psi_t^T(\kappa)\phi_t^T(\beta)] - [\psi_t^T(\kappa^0)\phi_t^T(\beta^0)]\}\kappa^0 \quad (64)$$

For the ordinary ELS method the last term on the right-hand side is simply

$$[C(q^{-1};\beta^0) - 1][\varepsilon_t(\kappa) - e_t] \quad (65)$$

because all components of the regressors except the $\varepsilon$- and the $e$-components are identical. This is certainly not true for our ELS variant because

$$\psi_t^T(\kappa) = [-A_2(q^{-1};\theta)y_{t-1}\ldots B_2(q^{-1};\theta)u_{t-1}\ldots\varepsilon_{t-1}(\kappa)\ldots\varepsilon_{t-n_c}(\kappa)] \quad (66)$$

$$\phi_t^T(\beta) = [-A_1(q^{-1};\beta)y_{t-1}\ldots - A_1(q^{-1};\beta)y_{t-n_2}B_1(q^{-1};\beta)u_{t-1}\ldots B_1(q^{-1};\beta)u_{t-m_2}] \quad (67)$$

whereas

$$\psi_t^T(\kappa^0) = [-A_2(q^{-1};\theta^0)y_{t-1}\ldots B_2(q^{-1};\theta^0)u_{t-1}\ldots e_{t-1}\ldots e_{t-n_c}] \quad (68)$$

$$\phi_t(\beta^0) = [-A_1(q^{-1};\beta^0)y_{t-1}\ldots - A_1(q^{-1};\beta^0)y_{t-n_2}B_1(q^{-1};\beta^0)u_{t-1} \quad (69)$$

$$\ldots B_1(q^{-1};\beta^0)u_{t-m_2}]$$

Thus, we do not from the above obtain a simple expression for $\varepsilon_t(\kappa) - e_t$ (it is this simple expression for $\varepsilon_t - e_t$, viz. $\varepsilon_t - e_t = -[1/C^0(q^{-1})]\varphi_t^T[\kappa - \kappa^0]$, which eventually leads to the sufficient condition for global convergence of the ordinary ELS method), and we have to conclude that the above approach does not lead to any sufficient condition for global convergence of the particular ELS variant described herein.

Another Lyapunov function is now attempted, viz.

$$V(\kappa, R) = E[\varepsilon_t^2(\kappa)] \quad (70)$$

which is the one that was used in Henriksen and Weyer (1990). We obtain

$$\frac{d}{d\tau} V(\kappa, R) = E\left[ \varepsilon_t(\kappa)\frac{\partial\varepsilon_t}{\partial\kappa^T}(\kappa) \right]\frac{d\kappa}{d\tau} = -E\left[ \varepsilon_t(\kappa)\frac{\partial\varepsilon_t}{\partial\kappa^T}(\kappa) \right]R^{-1}f(\kappa)$$

$$= -E\left[ \varepsilon_t(\kappa)\frac{1}{C(q^{-1};\beta)}\left[ \psi_t^T(\kappa)\phi_t^T(\beta) \right]R^{-1}f(\kappa) \right] \quad (71)$$

where

$$f(\kappa) = E\left[\begin{bmatrix}\psi_t(\kappa)\\\phi_t(\beta)\end{bmatrix}\varepsilon_t(\kappa)\right] \qquad (72)$$

The right-hand side of (71) is not a quadratic form, so it cannot be concluded that it is strictly negative (unless $C(q^{-1};\beta) \equiv 1$).

From the above it follows that we have not succeeded in finding sufficient conditions for global convergence of the ELS variant, and whether meaningful, although restricted, conditions can be established for $C(q^{-1};\beta) \neq 1$ is an open question.

Now, finally, why did we not succeed in finding similar global convergence conditions as for the ordinary ELS estimator, whereas the local convergence conditions turned out to be the same? The answer to the first part follows from the form of the two regressors, see (66)–(69). The components of these are *not independent* of the parameters even if we disregard the $\varepsilon$-components, whereas in the ordinary ELS method the pertinent components actually are. The latter fact allows a fairly simple expression for $\varepsilon_t(\kappa) - e_t$ to be established (something that failed in (64)), and this subsequently leads to a meaningful sufficient condition for global convergence. The answer to the second part of the question follows also from the form of the two regressors, (66)–(69). Although the components depend upon the parameters it is a fact that aside from the $\varepsilon$-components all the other ones are *linear* in the parameters. Since local analysis involves differentiating the regressors with respect to the parameters, we thus end up with something which aside from the entries resulting from the $\varepsilon$-components is *independent* of the parameters, see (87) and (89). This latter fact allows a fairly simple expression for $\partial\varepsilon_t/\partial\kappa^T$ to be established which then subsequently leads to the sufficient condition for local convergence derived herein.

## 4. Conclusion

The convergence properties of an ELS variant for ARMAX models based upon factorizing the $A$- and $B$-polynomials and filtering the input/output data with appropriate factors of these polynomials have been analyzed. The ELS variant is similar to other estimators (of MK, IV and PE types) which have been analyzed in earlier papers by Henriksen (1988, 1989), Henriksen and Weyer (1990), Weyer (1991). For *local* convergence it turned out that the convergence properties were exactly the same if we compare it with the ordinary ELS method, i.e. (1) the $C$-polynomial has to be strictly positive real (which also is an additional requirement for the ordinary ELS method compared with the ordinary LS method), and (2) the factors of the $A$- and $B$-polynomials have to be mutually coprime.

Having initially expected that similar properties for *global* convergence, when comparing the ELS variant herein with the ordinary ELS method, could be established, i.e., that the only additional requirement would be the coprimeness condition, the analysis revealed that this is not true. Strictly positive realness of $[1/C - 1/2]$ is not sufficient to ensure global convergence, although this is sufficient for global convergence of the ordinary ELS method.

In connection with the above it is worth mentioning that *global* convergence of the variants analyzed in the papers by Henriksen and Weyer (1990), Weyer (1991) could only be established for the LS variant, not for any of the IV or PE variants. The reason for the latter was partially similar to the one pointed out above, e.g., the regressors in the IV variants depended upon the parameters.

**Appendix**

*Proof of the Local Convergence Theorem*

The proof partially resembles proofs that can be found in Söderström and Stoica (1989), Wellstead and Zarrop (1991), but the calculations involved are quite more tedious. Stability of the system given by (58) will be determined by introducing the following Lyapunov function

$$V(\kappa(\tau)) = [\kappa(\tau) - \kappa^0]^T G(\kappa^0)[\kappa(\tau) - \kappa^0] \tag{73}$$

We find

$$\frac{d}{d\tau} V(\kappa(\tau)) = \left(\frac{d}{d\tau}[\kappa(\tau) - \kappa^0]\right)^T G(\kappa^0)[\kappa(\tau) - \kappa^0] + [\kappa(\tau) - \kappa^0]^T G(\kappa^0)\left(\frac{d}{d\tau}[\kappa(\tau) - \kappa^0]\right)$$

$$= [\kappa(\tau) - \kappa^0]^T \{\tilde{G}^T(\kappa^0) + \tilde{G}(\kappa^0)\}[\kappa(\tau) - \kappa^0] \tag{74}$$

where we first have to find an expression for

$$\tilde{G}(\kappa^0) = \frac{\partial f}{\partial \kappa^T}(\kappa^0) \tag{75}$$

The form of $f(\kappa)$ is given in (50), and omitting all arguments we write it simplified as

$$f = E\left[\begin{bmatrix} \psi_t\{w_t - r_t - \psi_t^T\beta\} \\ \phi_t\{z_t - [C-1]\varepsilon_t - \phi_t^T\theta\} \end{bmatrix}\right] \tag{76}$$

We have

$$\frac{\partial f}{\partial \kappa^T} = E\begin{bmatrix} \frac{\partial \psi_t}{\partial \kappa^T}\{w_t - r_t - \psi_t^T\beta\} \\ \frac{\partial \phi_t}{\partial \kappa^T}\{z_t - [C-1]\varepsilon_t - \phi_t^T\theta\} \end{bmatrix} + E\begin{bmatrix} \psi_t\frac{\partial}{\partial \kappa^T}\{w_t - r_t - \psi_t^T\beta\} \\ \phi_t\frac{\partial}{\partial \kappa^T}\{z_t - [C-1]\varepsilon_t - \phi_t^T\theta\} \end{bmatrix} \tag{77}$$

Now,

$$w_t(\theta) - r_t(\theta) - \psi_t^T(\kappa)\beta = \varepsilon_t(\kappa) \tag{78}$$

$$z_t(\beta) - [C(q^{-1};\beta) - 1]\varepsilon_t(\kappa) - \phi_t^T(\beta)\theta = \varepsilon_t(\kappa) \tag{79}$$

and for $\kappa = \kappa^0$ we have

$$w_t(\theta^0) - r_t(\theta^0) - \psi_t^T(\kappa^0)\beta^0 = \varepsilon_t(\kappa^0) \equiv e_t \tag{80}$$

$$z_t(\beta^0) - [C(q^{-1};\beta^0) - 1]\varepsilon_t(\kappa^0) - \phi_t^T(\beta^0)\theta^0$$

$$= z_t(\beta^0) - [C(q^{-1};\beta^0) - 1]e_t - \phi_t^T(\beta^0)\theta^0 = \varepsilon_t(\kappa^0) \equiv e_t \tag{81}$$

The first term on the right-hand side of (77) therefore vanishes at the point $\kappa^0$ because none of the components of $\psi_t$ and $\phi_t$ are correlated with $e_t$, the latter being white.

In the second right-hand term of (77) we have

$$\frac{\partial}{\partial \kappa^T}[w_t - r_t] = \frac{\partial}{\partial \kappa^T}[A_2 y_t - B_2 u_t]$$

$$= [0_{n_1}\, 0_{m_1}\, 0_{n_c}\, y_{t-1}\, y_{t-2}\cdots y_{t-n_2}\, -u_{t-1}\, -u_{t-2}\cdots -u_{t-m_2}] \tag{82}$$

where $0_k$ denotes a row $k$-vector containing exactly $k$ zeros. Likewise,

$$\frac{\partial}{\partial \kappa^{\mathrm{T}}}[z_t - [C-1]\varepsilon_t] = \frac{\partial}{\partial \kappa^{\mathrm{T}}}[A_1 y_t - [C-1]\varepsilon_t]$$

$$= [y_{t-1} y_{t-2} \ldots y_{t-n_1} 0_{m_1} - \varepsilon_{t-1} - \varepsilon_{t-2} \ldots - \varepsilon_{t-n_c} 0_{n_2} 0_{m_2}] - [C-1]\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} \quad (83)$$

Furthermore,

$$\frac{\partial}{\partial \kappa^{\mathrm{T}}}[\psi_t^{\mathrm{T}}\beta] = \beta^{\mathrm{T}}\frac{\partial \psi_t}{\partial \kappa^{\mathrm{T}}} + \psi_t^{\mathrm{T}}\frac{\partial \beta}{\partial \kappa^{\mathrm{T}}} = \beta^{\mathrm{T}}\frac{\partial \psi_t}{\partial \kappa^{\mathrm{T}}} + \psi_t^{\mathrm{T}}[I_{n_1+m_1+n_c} 0_{n_2+m_2}^{n_1+m_1+n_c}] \quad (84)$$

where $I_k$ denotes the $k \times k$ identity matrix whereas $0_k^l$ denotes the $l \times k$ matrix containing solely zero elements. Likewise,

$$\frac{\partial}{\partial \kappa^{\mathrm{T}}}[\phi_t^{\mathrm{T}} 0] = 0^{\mathrm{T}}\frac{\partial \phi_t}{\partial \kappa^{\mathrm{T}}} + \phi_t^{\mathrm{T}}\frac{\partial 0}{\partial \kappa^{\mathrm{T}}} = 0^{\mathrm{T}}\frac{\partial \phi_t}{\partial \kappa^{\mathrm{T}}} + \phi_t^{\mathrm{T}}[0_{n_1+m_1+n_c}^{n_2+m_2} I_{n_2+m_2}] \quad (85)$$

Introducing the following notations, in the form of two Hankel matrices,

$$Y_q^p(t) = \begin{bmatrix} y_t & y_{t-1} & \cdots & y_{t-q+1} \\ y_{t-1} & y_{t-2} & \cdots & y_{t-q} \\ \vdots & \vdots & \vdots & \vdots \\ y_{t-p+1} & y_{t-p} & \cdots & y_{t-p-q+2} \end{bmatrix} \quad U_q^p(t) = \begin{bmatrix} u_t & u_{t-1} & \cdots & u_{t-q+1} \\ u_{t-1} & u_{t-2} & \cdots & u_{t-q} \\ \vdots & \vdots & \vdots & \vdots \\ u_{t-p+1} & u_{t-p} & \cdots & u_{t-p-q+2} \end{bmatrix}$$

$$(86)$$

we find

$$\frac{\partial \psi_t}{\partial \kappa^{\mathrm{T}}} = \begin{bmatrix} 0_{n_1+m_1+n_c}^{n_1} & -Y_{n_2}^{n_1}(t-2) & 0_{m_2}^{n_1} \\ 0_{n_1+m_1+n_c}^{m_1} & 0_{n_2}^{m_1} & U_{m_2}^{m_1}(t-2) \\ & \dfrac{\partial \varepsilon_{t-1}}{\partial \kappa^{\mathrm{T}}} & \\ & \dfrac{\partial \varepsilon_{t-2}}{\partial \kappa^{\mathrm{T}}} & \\ & \vdots & \\ & \dfrac{\partial \varepsilon_{t-n_c}}{\partial \kappa^{\mathrm{T}}} & \end{bmatrix},$$

$$(n_1 + m_1 + n_c) \times (n_1 + m_1 + n_c + n_2 + m_2) \quad (87)$$

It follows that

$$\beta^{\mathrm{T}}\frac{\partial \psi_t}{\partial \kappa^{\mathrm{T}}} = [C-1]\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} + [A_1-1][0_{n_1+m_1+n_c} - y_{t-1} - y_{t-2} \ldots - y_{t-n_2} 0_{n_2}]$$

$$+ [B_1-1][0_{n_1+m_1+n_c} 0_{n_2} u_{t-1} u_{t-2} \ldots u_{t-n_2}]$$

and we obtain

$$\frac{\partial}{\partial \kappa^{\mathrm{T}}}[w_t - r_t - \psi_t^{\mathrm{T}}\beta] = -[C-1]\frac{\partial \varepsilon_1}{\partial \kappa^{\mathrm{T}}} - [0_{n_1+m_1+n_c} \phi_t^{\mathrm{T}}] - [\psi_t^{\mathrm{T}} 0_{n_2+m_2}]$$

$$= -[C-1]\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} - [\psi_t^{\mathrm{T}} \phi_t^{\mathrm{T}}] \quad (88)$$

Furthermore,

$$\frac{\partial \phi_t}{\partial \kappa^{\mathrm{T}}} = \begin{bmatrix} -Y_{n_1}^{n_2}(t-2) & 0_{m_1}^{n_2} & 0_{n_c+n_2+m_2}^{n_2} \\ 0_{n_1}^{m_2} & U_{m_1}^{m_2}(t-2) & 0_{n_c+n_2+m_2} \end{bmatrix}, (n_2+m_2) \times (n_1+m_1+n_c+n_2+m_2) \tag{89}$$

which yields

$$0^{\mathrm{T}}\frac{\partial \phi_t}{\partial \kappa^{\mathrm{T}}} = [A_2 - 1][-y_{t-1} \ -y_{t-2}\cdots \ -y_{t-n_1}\ 0_{m_1}\ 0_{n_c+n_2+m_2}]$$

$$+ B_2[0_{n_1}\ u_{t-1}\ u_{t-2}\cdots \ u_{t-m_1}\ 0_{n_c+n_2+m_2}]$$

and

$$\frac{\partial}{\partial \kappa^{\mathrm{T}}}[z_t - [C-1]\varepsilon_t - \phi_t^{\mathrm{T}}0] = [\psi_t^{\mathrm{T}}\ 0_{n_2+m_2}] - [C-1]\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} - [0_{n_1+m_1+n_c}\ \phi_t^{\mathrm{T}}]$$

$$= -[C-1]\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} - [\psi_t^{\mathrm{T}}\ \phi_t^{\mathrm{T}}] \tag{90}$$

Since we have

$$\varepsilon_t = w_t - r_t - \psi_t^{\mathrm{T}}\beta = z_t - [C-1]\varepsilon_t - \phi_t^{\mathrm{T}}0 \tag{91}$$

it follows from either of the above equations, i.e., (88) or (90),

$$\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} = -[C-1]\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} - [\psi_t^{\mathrm{T}}\ \phi_t^{\mathrm{T}}] \tag{92}$$

which yields

$$\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}} = -\frac{1}{C(q^{-1})}[\psi_t^{\mathrm{T}}\ \phi_t^{\mathrm{T}}] \tag{93}$$

and, particularly,

$$\frac{\partial \varepsilon_t}{\partial \kappa^{\mathrm{T}}}(\kappa^0) = -\frac{1}{C(q^{-1};\beta^0)}[\psi_t^{\mathrm{T}}(\kappa^0)\ \phi_t^{\mathrm{T}}(\beta^0)] \tag{94}$$

From (77) and what follows thereafter we now have that

$$\frac{\partial f}{\partial \kappa^{\mathrm{T}}}(\kappa^0) = E\begin{bmatrix} \psi_t(\kappa^0)\cdot\left[-\dfrac{1}{C(q^{-1};\beta^0)}[\psi_t^{\mathrm{T}}(\kappa^0)\ \phi_t^{\mathrm{T}}(\beta^0)]\right] \\[2ex] \phi_t(\beta^0)\cdot\left[-\dfrac{1}{C(q^{-1};\beta^0)}[\psi_t^{\mathrm{T}}(\kappa^0)\ \phi_t^{\mathrm{T}}(\beta^0)]\right] \end{bmatrix}$$

$$= -E\left[\begin{bmatrix} \psi_t(\kappa^0) \\ \phi_t(\beta^0) \end{bmatrix}\cdot\frac{1}{C(q^{-1};\beta^0)}[\psi_t^{\mathrm{T}}(\kappa^0)\ \phi_t^{\mathrm{T}}(\beta^0)]\right] = \tilde{G}(\kappa^0) \tag{95}$$

We define the scalar process

$$v_t = \frac{1}{C(q^{-1};\beta^0)}[\psi_t^{\mathrm{T}}(\kappa^0)\ \phi_t^{\mathrm{T}}(\beta^0)]\cdot[\kappa(\tau)-\kappa^0] \tag{96}$$

and obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tau} V(\kappa(\tau)) = - E[v_t \cdot C(q^{-1}; \beta^0)v_t + C(q^{-1}; \beta^0)v_t \cdot v_t] = - 2E(v_t \cdot C(q^{-1}; \beta^0)v_t]$$

$$(97)$$

since

$$C(q^{-1}; \beta^0)v_t = [\psi_t^{\mathrm{T}}(\kappa^0)\, \phi_t^{\mathrm{T}}(\beta^0)] \cdot [\kappa(\tau) - \kappa^0] = [\kappa(\tau) - \kappa^0]^{\mathrm{T}} \cdot \begin{bmatrix} \psi_t(\kappa^0) \\ \phi_t(\beta^0) \end{bmatrix}$$

Let $\Phi_{vv}(e^{j\omega})$ denote the power spectral density of $v_t$ (which is real and nonnegative since $v_t$ is scalar). We now obtain what follows

$$\frac{\mathrm{d}}{\mathrm{d}\tau} V(\kappa(\tau)) = - 2E[v_t \cdot C(q^{-1}; \beta^0)] = - 2\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{vv}(e^{j\omega})C(e^{-j\omega}; \beta^0)d\omega$$

$$= -\frac{1}{\pi}\int_{-\pi}^{\pi} \Phi_{vv}(e^{j\omega})\mathrm{Re}[C(e^{-j\omega}; \beta^0)]d\omega \quad (98)$$

From this we conclude that a *sufficient* condition for $\mathrm{d}/\mathrm{d}\tau\, V(\kappa(\tau)) < 0$ simply is that

$$\mathrm{Re}[C(e^{-j\omega}; \beta^0)] > 0, \qquad -\pi \leqslant \omega \leqslant \pi \quad (99)$$

i.e., $C(q^{-1}; \beta_0)$ *strictly positive real* (in the appropriate frequency domain). As will be well-known to many of the readers, this is exactly the same condition which ensures local convergence of the ordinary ELS method.                                                $\square$

## REFERENCES

CLARY, J. P. and FRANKLIN, G. F. (1984). Self-tuning control with *a priori* plant knowledge. *Proc. 23rd IEEE Conference on Decision and Control*, Las Vegas, Nevada, December 1984, pp. 369–374.

HENRIKSEN, R. (1988). Convergence analysis of some decentralized estimators. *Proc. 8th IFAC/ IFORS Symposium on Identification and System Parameter Estimation* (selected papers), Beijing, P. R. China, August 27–31, pp. 361–366 (*IFAC Proceedings Series*, 1989, Number 8, Volume 1, Pergamon Press).

HENRIKSEN, R. (1989). Convergence analysis of some decentralized estimators. *Modeling, Identification and Control*, **10**, No. 1, pp. 13–34.

HENRIKSEN, R. (1991). Accuracy of some robust estimators based upon prefiltering of the input/ output data. *Proc. 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, August 8–12.

HENRIKSEN, R. (1992). Asymptotic properties of some estimators based upon prefiltering of the input/output data. *Proc. 11th IASTED International Conference Modeling, Identification and Control*, Innsbruck, Austria, February 10–12, pp. 231–235.

HENRIKSEN, R. and WEYER, E. (1990). Convergence aspects of some robust estimators based upon prefiltering of the input/output data. *Proc. 11th IFAC World Congress*, Tallinn, Estonia, August 13–17, Vol. 3, pp. 215–220.

LJUNG, L. (1977). Analysis of recursive stochastic algorithms. *IEEE Trans. Automatic Control*, AC-22, 551–575.

LJUNG, L. (1987). *System Identification: Theory for the User* (Prentice Hall).

LJUNG, L. and SÖDERSTRÖM, T. (1983). *Theory and Practice of Recursive Identification*. The MIT Press, Cambridge, Massachusetts.

SÖDERSTRÖM, T. and STOICA, P. (1989). *System Identification* (Prentice Hall).

WELLSTEAD, P. E. and ZARROP, M. B. (1991). *Self-tuning Systems—Control and Signal Processing* (John Wiley and Sons).

WEYER, E. (1991). A recursive decentralized parameter estimator for a general linear SISO system. *Proc. 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, August 8–12.

YOUNG, R. E., HENRIKSEN, R. and MELLICHAMP, D. A. (1987). A multi-rate decentralized parameter estimation method for stiff systems. *Proc. 26th IEEE Conference on Decesion and Control*, Los Angeles, California, December 1987, pp. 1902–1907.