

## Accuracy of some robust estimators based upon prefiltering of the input/output data†

ROLF HENRIKSEN‡

**Keywords:** *Parameter estimation, estimation accuracy, stiff systems.*

By prefiltering the input/output data and employing certain decentralized estimation techniques, it is possible to improve the robustness of some estimators significantly. Earlier papers on these techniques have been focused on local convergence properties of four bootstrap estimators, a LS variant and three IV variants, global convergence properties of the LS variant, and convergence rates of the LS variant and two of the IV variants when the underlying system is stiff. This paper is devoted to the accuracy properties of the LS variant and one of the IV variants.

### 1. Introduction

Parameter estimators based upon standard estimation techniques, viz. least squares (LS) methods, instrumental variable (IV) methods, etc. do occasionally have difficulties with systems that have a somewhat ill-conditioned nature, e.g., stiff systems. By prefiltering the input/output data and employing certain decentralized estimation techniques, it is, however, possible to improve the robustness significantly. Previous papers on the methods presented herein have been focused on robustness properties, local and global convergence properties, and convergence rates when the underlying system is stiff, see Young *et al.* (1987), Henriksen (1988, 1989), and Henriksen and Weyer (1990).

The paper is organized as follows. In Section 2 we present a brief outline of the system and a resume of previous results concerning local convergence properties, global convergence properties, convergence rates, and robustness. Section 3 is devoted to analysis of accuracy properties of two of the estimators, the LS variant and one of the IV variants.

### 2. System description and previous results

We consider a system described by the linear discrete-time model

$$A(q^{-1})y_t = B(q^{-1})u_t + v_t \quad (1)$$

where  $y_t$  is the output at time  $t$ ,  $u_t$  is the input, whereas  $v_t$  is the disturbance or residual.  $\{v_t\}$  is assumed to be a zero-mean stochastic process with a rational nonsingular spectral density matrix. The processes  $\{u_t\}$  and  $\{v_t\}$  are assumed to be mutually independent, and the system is assumed to be asymptotically stable.

---

Received 4 February 1992.

† This paper was presented at the 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation, 8-12 July 1991, Budapest, Hungary. Reprinted with permission.

‡ The Norwegian Institute of Technology, Division of Engineering Cybernetics, N-7034, Trondheim, Norway.

In the previous papers by Henriksen (1988, 1989) it was originally assumed that the model (1) could be MIMO (multi-input, multi-output). The most neat and interesting results were obtained, however, by assuming the system to be SISO (single-input, single-output). We will therefore, for the sake of simplicity, also assume that in this paper. It should be noted, however, that the analysis can be brought quite far without making this additional assumption.

The polynomials  $A(q^{-1})$  and  $B(q^{-1})$  are factored as, respectively,

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

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

where

$$A_1(q^{-1}) = 1 + a_1^1 q^{-1} + a_2^1 q^{-2} + \dots + a_{n_1}^1 q^{-n_1} \quad (4)$$

$$A_2(q^{-1}) = 1 + a_1^2 q^{-1} + a_2^2 q^{-2} + \dots + a_{n_2}^2 q^{-n_2} \quad (5)$$

$$B_1(q^{-1}) = 1 + b_1^1 q^{-1} + b_2^1 q^{-2} + \dots + b_{m_1}^1 q^{-m_1} \quad (6)$$

$$B_2(q^{-1}) = b_1^2 q^{-1} + b_2^2 q^{-2} + \dots + b_{m_2}^2 q^{-m_2} \quad (7)$$

and where  $n_1 + n_2 = n$ , the degree of  $A(q^{-1})$ , whereas  $m_1 + m_2 = m$ , the degree of  $B(q^{-1})$ .

Assuming 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 input of the system, viz.

$$w_t = A_2(q^{-1})y_t; \quad r_t = B_2(q^{-1})u_t \quad (8)$$

(1) thus takes the form

$$A_1(q^{-1})w_t = B_1(q^{-1})r_t + v_t \quad (9)$$

which is a reduced model of the system. We shall refer to it as model  $M_1$ .

Similarly, assuming  $A_1(q^{-1})$  and  $B_1(q^{-1})$  to be known, we can define

$$z_t = A_1(q^{-1})y_t; \quad s_t = B_1(q^{-1})u_t \quad (10)$$

which leads to the reduced model

$$A_2(q^{-1})z_t = B_2(q^{-1})s_t + v_t \quad (11)$$

This model will be referred to as model  $M_2$ .

Equations (8)–(11) form the basis for the estimators considered in this paper. We can rewrite (9) and (11) as, respectively,

$$w_t = \psi_t^T \beta + r_t + v_t \quad (12)$$

$$z_t = \varphi_t^T \theta + v_t \quad (13)$$

where

$$\psi_t = [-w_{t-1}, \dots, -w_{t-n_1}, r_{t-1}, \dots, r_{t-m_1}]^T \quad (14)$$

$$\beta = [a_1^1, \dots, a_{n_1}^1, b_1^1, \dots, b_{m_1}^1]^T \quad (15)$$

$$\varphi_t = [-z_{t-1}, \dots, -z_{t-n_2}, s_{t-1}, \dots, s_{t-m_2}]^T \quad (16)$$

$$\theta = [a_1^2, \dots, a_{n_2}^2, b_1^2, \dots, b_{m_2}^2]^T \quad (17)$$

From (12) and (13) we can derive the following LS estimator:

$$\hat{\beta}_N^{LS} = \left[ \frac{1}{N} \sum_{t=1}^N \psi_t \psi_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \psi_t (w_t - r_t) \right] \quad (18)$$

$$\hat{\theta}_N^{LS} = \left[ \frac{1}{N} \sum_{t=1}^N \varphi_t \varphi_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \varphi_t z_t \right] \quad (19)$$

Since neither of the true values  $\beta^*$  and  $\theta^*$  generally are known beforehand, the two estimators (18) and (19) will have to be employed in a bootstrap fashion by replacing  $w_t$ ,  $r_t$ ,  $z_t$ , and  $s_t$  with, respectively,

$$\hat{w}_t = A_1(q^{-1}, \hat{\beta}) y_t; \quad \hat{r}_t = B_1(q^{-1}, \hat{\beta}) u_t \quad (20)$$

$$\hat{z}_t = A_2(q^{-1}, \hat{\theta}) y_t; \quad \hat{s}_t = B_2(q^{-1}, \hat{\theta}) u_t \quad (21)$$

where  $\hat{\beta}$  is short for  $\hat{\beta}_N^{LS}$  etc. For more details, see Henriksen (1988, 1989), and Henriksen and Weyer (1990).

From (18) and (19) we can also immediately derive an IV estimator of the form

$$\hat{\beta}_N^{IV} = \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\psi}_t \tilde{\psi}_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\psi}_t (w_t - r_t) \right] \quad (22)$$

$$\hat{\theta} = \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\varphi}_t \tilde{\varphi}_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\varphi}_t z_t \right] \quad (23)$$

where

$$\tilde{\psi}_t = [-\tilde{w}_{t-1}, \dots, -\tilde{w}_{t-n_1}, r_{t-1}, \dots, r_{t-m_1}]^T \quad (24)$$

$$\tilde{w}_t = A_1^{-1}(q^{-1}) B(q^{-1}) u_t \quad (25)$$

$$\tilde{\varphi}_t = [-\tilde{z}_{t-1}, \dots, -\tilde{z}_{t-n_2}, s_{t-1}, \dots, s_{t-m_2}]^T \quad (26)$$

$$\tilde{z}_t = A_2^{-1}(q^{-1}) B(q^{-1}) u_t \quad (27)$$

A couple of other IV estimators have been suggested in the earlier papers by Henriksen (1988, 1989), and Henriksen and Weyer (1990), e.g., a symmetric variant of the form

$$\hat{\beta}_N^{SIV} = \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\psi}_t \tilde{\psi}_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\psi}_t (w_t - r_t) \right] \quad (28)$$

$$\hat{\theta}_N^{SIV} = \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\varphi}_t \tilde{\varphi}_t^T \right]^{-1} \left[ \frac{1}{N} \sum_{t=1}^N \tilde{\varphi}_t z_t \right] \quad (29)$$

but the consistency and accuracy properties of these variants are doubtful, see Henriksen (1988, 1989) or Söderström and Stoica (1983). We will therefore not pursue any of these other variants any further.

The convergence properties of the above estimators have been thoroughly investigated in a couple of papers by Henriksen (1988, 1989), and in a paper by Henriksen and Weyer (1990). Provided certain standard consistency conditions are satisfied, e.g., (1) the polynomials  $A(q^{-1})$  and  $B(q^{-1})$  are coprime, (2) the input process  $\{u_t\}$  is stationary, ergodic with respect to second-order moments, and persistently exciting of order  $n+m$ , (3) the input  $u_t$  and the disturbance  $v_s$  are independent for all  $t$  and  $s$ , etc., see Henriksen (1989) or Söderström and Stoica (1983), some neat results about the convergence properties of the above estimators can be given. These results will be summarized in what follows.

After some elaborate computations it can be derived that local convergence (about the point  $(\beta^*, \theta^*)$ ) of the LS variant can be determined from the eigenvalues of the matrix

$$F = [E\psi_t\psi_t^T]^{-1}E\psi_t\varphi_t^T[E\varphi_t\varphi_t^T]^{-1}E\varphi_t\psi_t^T \quad (30)$$

whereas local convergence of the IV variant can be determined from the eigenvalues of the matrix

$$\tilde{F} = [E\tilde{\psi}_t\tilde{\psi}_t^T]^{-1}E\tilde{\psi}_t\tilde{\varphi}_t^T[E\tilde{\varphi}_t\tilde{\varphi}_t^T]^{-1}E\tilde{\varphi}_t\tilde{\psi}_t^T \quad (31)$$

where  $E$  denotes the expectation operator.

#### Fact 1

If  $\lambda$  is an eigenvalue of  $F$  (or of  $\tilde{F}$ ), then  $\lambda$  is real and  $0 \leq \lambda \leq 1$ . Moreover,  $\lambda = 1$  is an eigenvalue of  $F$  (or of  $\tilde{F}$ ) of multiplicity  $k = i + j$  if and only if  $A_1(q^{-1})$  and  $A_2(q^{-1})$  have exactly  $i$  common zeros, and  $B_1(q^{-1})$  and  $B_2(q^{-1})$  have exactly  $j$  common zeros.

#### Fact 2

The above bootstrap estimators converge *locally* if both the polynomials  $A_1(q^{-1})$  and  $A_2(q^{-1})$  and the polynomials  $B_1(q^{-1})$  and  $B_2(q^{-1})$  are coprime.

These two facts were essentially shown in Henriksen (1988, 1989). Some further results were obtained in a recent paper by Henriksen and Weyer (1990).

#### Fact 3

The LS variant converges *globally* if it converges *locally*.

Furthermore, as the stiffness of the system tends to infinity, all eigenvalues of  $F$  (or of  $\tilde{F}$ ) tend to zero, i.e., local convergence of the two estimators tends to becoming instantaneous.

### 3. Asymptotic distributions of the estimators

In this section we will evaluate the asymptotic distributions of the two estimators (considered as bootstrap estimators). We start with the accuracy result for the LS variant.

#### Theorem 1 (Accuracy of the LS variant)

Assume sufficient conditions for consistency of the LS variant are satisfied and that the processes  $\{u_t\}$  and  $\{v_t\}$  are stationary (and mutually independent). Then the parameter estimates  $\beta_N^{LS}$  and  $\theta_N^{LS}$  are asymptotically Gaussian distributed in the sense that

$$(N)^{1/2}[\beta_N^{LS} - \beta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_\beta^{LS} \quad (32)$$

$$(N)^{1/2}[\theta_N^{LS} - \theta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_\theta^{LS} \quad (33)$$

where  $\gamma_\beta^{LS} \sim N(0, P_\beta^{LS})$  and  $\gamma_\theta^{LS} \sim N(0, P_\theta^{LS})$  and where the covariance matrices  $P_\beta^{LS}$  and  $P_\theta^{LS}$  are given by, respectively,

$$P_\beta^{LS} = R_v R_\psi^{-1}(\theta^*); \quad P_\theta^{LS} = R_v R_\varphi^{-1}(\beta^*) \quad (34)$$

where

$$R_\psi(\theta^*) = E\psi_t(\theta^*)\psi_t^T(\theta^*); \quad R_\varphi(\beta^*) = E\varphi_t(\beta^*)\varphi_t^T(\beta^*) \quad (35)$$

whereas  $R_v = E(v_t)^2$ .

*Proof.* See the Appendix.

By assuming the input process  $\{u_t\}$  to be white we will make some further refinements of the above result (the noise process  $\{v_t\}$  has to be white in order to satisfy the consistency conditions). Define the series

$$\sum_{i=0}^{\infty} h_i^1 z^{-i} = \frac{1}{A_1^*(z^{-1})}; \quad \sum_{i=0}^{\infty} h_i^2 z^{-i} = \frac{1}{A_2^*(z^{-1})} \quad (36)$$

$$\sum_{i=0}^{\infty} g_i^1 z^{-i} = \frac{B^*(z^{-1})}{A_1^*(z^{-1})}; \quad \sum_{i=0}^{\infty} g_i^2 z^{-i} = \frac{B^*(z^{-1})}{A_2^*(z^{-1})} \quad (37)$$

where

$$A_1^*(z^{-1}) = A_1(z^{-1}, \beta^*), \quad A_2^*(z^{-1}) = A_2(z^{-1}, \theta^*),$$

$$B^*(z^{-1}) = B(z^{-1}, \beta^*, \theta^*) = B_1(z^{-1}, \beta^*)B_2(z^{-1}, \theta^*).$$

Furthermore, define the matrix

$$H_\beta = \begin{bmatrix} H_\beta^{11} & 0 \\ 0 & 0 \end{bmatrix} \quad (38)$$

where the submatrix  $H_\beta^{11}$  is given by

$$H_\beta^{11} = [h_{\beta,ij}^{11}], \quad i = 1, \dots, n_1, j = 1, \dots, n_1 \quad (39)$$

with entry  $h_{\beta,ij}^{11}$  given by

$$h_{\beta,ij}^{11} = \sum_{k=0}^{\infty} h_k^1 h_{k+|i-j|}^1 \quad (40)$$

We also define the matrix

$$G_\beta = \begin{bmatrix} G_\beta^{11} & G_\beta^{12} \\ G_\beta^{21} & G_\beta^{22} \end{bmatrix} \quad (41)$$

where the submatrices  $G_\beta^{pq}$  are given by

$$G_\beta^{11} = [g_{\beta,ij}^{11}], \quad i = 1, \dots, n_1, j = 1, \dots, n_1 \quad (42)$$

$$G_\beta^{12} = [g_{\beta,ij}^{12}], \quad i = 1, \dots, n_1, j = 1, \dots, m_1 \quad (43)$$

$$G_\beta^{21} = [g_{\beta,ij}^{21}], \quad i = 1, \dots, m_1, j = 1, \dots, n_1 \quad (44)$$

$$G_\beta^{22} = [g_{\beta,ij}^{22}], \quad i = 1, \dots, m_1, j = 1, \dots, m_1 \quad (45)$$

and where the entries  $g_{\beta,ij}^{pq}$  of these submatrices are given by, respectively,

$$g_{\beta,ij}^{11} = \sum_{k=0}^{\infty} g_k^1 g_{k+|i-j|}^1; \quad g_{\beta,ij}^{12} = - \sum_{k=0}^{m_2} g_{k-i+j}^1 b_k^{2*} \quad (46)$$

$$g_{\beta,ij}^{21} = - \sum_{k=0}^{m_2} g_{k+i-j}^1 b_k^{2*}; \quad g_{\beta,ij}^{22} = \sum_{k=0}^{m_2} b_k^{2*} b_{k+|i-j|}^{2*} \quad (47)$$

Note that  $g_k^1 = 0$  for  $k < 0$  and that  $b_k^{2*} = 0$  for  $k > m_2$ . Furthermore,  $b_k^{2*}$  means the true value of  $b_k^2$ .

In a similar fashion we define the matrices

$$H_{\theta} = \begin{bmatrix} H_{\theta}^{11} & 0 \\ 0 & 0 \end{bmatrix} \quad (48)$$

$$G_{\theta} = \begin{bmatrix} G_{\theta}^{11} & G_{\theta}^{12} \\ G_{\theta}^{21} & G_{\theta}^{22} \end{bmatrix} \quad (49)$$

where the submatrices are given by

$$H_{\theta}^{11} = [h_{\theta,ij}^{11}], \quad i = 1, \dots, n_2, j = 1, \dots, n_2 \quad (50)$$

$$G_{\theta}^{11} = [g_{\theta,ij}^{11}], \quad i = 1, \dots, n_2, j = 1, \dots, n_2 \quad (51)$$

$$G_{\theta}^{12} = [g_{\theta,ij}^{12}], \quad i = 1, \dots, n_2, j = 1, \dots, m_2 \quad (52)$$

$$G_{\theta}^{21} = [g_{\theta,ij}^{21}], \quad i = 1, \dots, m_2, j = 1, \dots, n_2 \quad (53)$$

$$G_{\theta}^{22} = [g_{\theta,ij}^{22}], \quad i = 1, \dots, m_2, j = 1, \dots, m_2 \quad (54)$$

with entries given by, respectively,

$$h_{\theta,ij}^{11} = \sum_{k=0}^{\infty} h_k^2 h_{k+|i-j|}^2; \quad (55)$$

$$g_{\theta,ij}^{11} = \sum_{k=0}^{\infty} g_k^2 g_{k+|i-j|}^2; \quad g_{\theta,ij}^{12} = - \sum_{k=0}^{m_1} g_{k-i+j}^2 b_k^{1*} \quad (56)$$

$$g_{\theta,ij}^{21} = - \sum_{k=0}^{m_1} g_{k+i-j}^2 b_k^{1*}; \quad g_{\theta,ij}^{22} = \sum_{k=0}^{m_1} b_k^{1*} b_{k+|i-j|}^{1*} \quad (57)$$

Note that  $g_k^2 = 0$  for  $k < 0$  and that  $b_k^{1*} = 0$  for  $k > m_1$ . Furthermore  $b_k^{1*}$  means the true value of  $b_k^1$ .

With the above definitions we are now able to present what follows.

#### Corollary 1

Assume the input process  $\{u_t\}$  to be white with  $R_u = E(u_t)^2$ . Then the covariance matrices  $P_{\beta}^{LS}$  and  $P_{\theta}^{LS}$  of the LS variant are given by, respectively,

$$P_{\beta}^{LS} = R_v(R_u G_{\beta} + R_v H_{\beta})^{-1}; \quad P_{\theta}^{LS} = R_v(R_u G_{\theta} + R_v H_{\theta})^{-1} \quad (58)$$

*Proof.* See the Appendix.

We now proceed with the IV variant. In this case we do not have to assume the disturbance to be white; it is instead assumed to be of the form

$$v_t = C(q^{-1})\varepsilon_t \quad (59)$$

where the process  $\{\varepsilon_t\}$  is white and stationary with  $R_{\varepsilon} = E(\varepsilon_t)^2$  and where  $C(q^{-1})$  is a filter such that both  $C(q^{-1})$  and  $C^{-1}(q^{-1})$  are asymptotically stable whereas  $C(0) = 1$ .

**Theorem 2** (Accuracy of the IV variant)

Assume sufficient conditions for consistency for the IV variant are satisfied and that the processes  $\{u_t\}$  and  $\{v_t\}$  are stationary (and mutually independent). Then the parameter estimates  $\hat{\beta}_N^{IV}$  and  $\hat{\theta}_N^{IV}$  are asymptotically Gaussian distributed in the sense that

$$(N)^{1/2}[\hat{\beta}_N^{IV} - \beta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_\beta^{IV} \quad (60)$$

$$(N)^{1/2}[\hat{\theta}_N^{IV} - \theta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_\theta^{IV} \quad (61)$$

where  $\gamma_\beta^{IV} \sim N(0, P_\beta^{IV})$  and  $\gamma_\theta^{IV} \sim N(0, P_\theta^{IV})$  and where the covariance matrices  $P_\beta^{IV}$  and  $P_\theta^{IV}$  are given by, respectively,

$$P_\beta^{IV} = R_\psi^{-1}(\beta^*, \theta^*) Q_\psi(\beta^*, \theta^*) R_\psi^{-1}(\beta^*, \theta^*) \quad (62)$$

$$P_\theta^{IV} = R_\phi^{-1}(\beta^*, \theta^*) Q_\phi(\beta^*, \theta^*) R_\phi^{-1}(\beta^*, \theta^*) \quad (63)$$

where

$$R_\psi(\beta^*, \theta^*) = E\tilde{\psi}_t(\beta^*, \theta^*)\tilde{\psi}_t^T(\beta^*, \theta^*) = E\tilde{\psi}_t(\beta^*, \theta^*)\tilde{\psi}_t^T(\beta^*, \theta^*) \quad (64)$$

$$R_\phi(\beta^*, \theta^*) = E\tilde{\phi}_t(\beta^*, \theta^*)\tilde{\phi}_t^T(\beta^*, \theta^*) = E\tilde{\phi}_t(\beta^*, \theta^*)\tilde{\phi}_t^T(\beta^*, \theta^*) \quad (65)$$

$$Q_\psi(\beta^*, \theta^*) = R_\epsilon E C(q^{-1})\tilde{\psi}_t(\beta^*, \theta^*) \cdot C(q^{-1})\tilde{\psi}_t^T(\beta^*, \theta^*) \quad (66)$$

$$Q_\phi(\beta^*, \theta^*) = R_\epsilon E C(q^{-1})\tilde{\phi}_t(\beta^*, \theta^*) \cdot C(q^{-1})\tilde{\phi}_t^T(\beta^*, \theta^*) \quad (67)$$

*Proof.* See the Appendix.

By assuming the input process  $\{u_t\}$  to be white we can also in this case make some further refinements. Of course, the noise process  $\{v_t\}$  does not have to be white in this case, but we will for the case of simplicity and for the purpose of making a direct comparison also make this assumption.

**Corollary 2**

Assume both the input process  $\{u_t\}$  and the noise process  $\{v_t\}$  to be white with  $R_u = E(u_t)^2$  and  $R_v = E(v_t)^2$ . Then the covariance matrices  $P_\beta^{IV}$  and  $P_\theta^{IV}$  of the IV variant are given by, respectively,

$$P_\beta^{IV} = R_v R_u^{-1} G_\beta^{-1}; \quad P_\theta^{IV} = R_v R_u^{-1} G_\theta^{-1} \quad (68)$$

*Proof.* See the Appendix.

The two matrices  $G_\beta$  and  $G_\theta$  are the same as defined before, i.e., as defined by Eqns. (41)–(47) and by (49), (51)–(54), and (56)–(57), respectively.

We will now make a comparison of the above asymptotic properties with the asymptotic properties of more ordinary LS and IV variants. For this purpose we consider a system of the form

$$A_1(q^{-1})y_t = B_1(q^{-1})u_t + v_t \quad (69)$$

where the polynomials  $A_1(q^{-1})$  and  $B_1(q^{-1})$  are the same as before. Assume the input  $u_t$  to be an MA process of the form

$$u_t = D(q^{-1})\eta_t \quad (70)$$

where  $\{\eta_t\}$  is a white process with  $E(\eta_t)^2 = R_\eta$  (it is, furthermore, assumed to be independent of the white process  $\{v_t\}$ ). The asymptotic covariance matrix of this estimator is given by

$$P_\beta^{\text{OLS}} = R_v(R_\eta F_\beta + R_v H_\beta)^{-1} \quad (71)$$

where the matrix  $H_\beta$  is the same as before, i.e., as defined by Eqns. (38)–(40), whereas  $F_\beta$  is defined by

$$F_\beta = \begin{bmatrix} F_\beta^{11} & F_\beta^{12} \\ F_\beta^{21} & F_\beta^{22} \end{bmatrix} \quad (72)$$

where the submatrices  $F_\beta^{pq}$  are given by

$$F_\beta^{11} = [f_{\beta,ij}^{11}], \quad i=1, \dots, n_1, \quad j=1, \dots, n_1 \quad (73)$$

$$F_\beta^{12} = [f_{\beta,ij}^{12}], \quad i=1, \dots, n_1, \quad j=1, \dots, m_1 \quad (74)$$

$$F_\beta^{21} = [f_{\beta,ij}^{21}], \quad i=1, \dots, m_1, \quad j=1, \dots, n_1 \quad (75)$$

$$F_\beta^{22} = [f_{\beta,ij}^{22}], \quad i=1, \dots, m_1, \quad j=1, \dots, m_1 \quad (76)$$

and where the entries  $f_{\beta,ij}^{pq}$  are given by, respectively,

$$f_{\beta,ij}^{11} = \sum_{k=0}^{\infty} f_k^1 f_{k+|i-j|}^1, \quad f_{\beta,ij}^{12} = - \sum_{k=0}^{\delta} f_{k-i+j}^1 d_k \quad (77)$$

$$f_{\beta,ij}^{21} = - \sum_{k=0}^{\delta} f_{k+i-j}^1 d_k, \quad f_{\beta,ij}^{22} = \sum_{k=0}^{\delta} d_k d_{k+|i-j|} \quad (78)$$

Furthermore, the series  $f_0^1, f_1^1, f_2^1, \dots$  is defined by

$$\sum_{k=0}^{\infty} f_k^1 z^{-k} = \frac{B_1^*(z^{-1})D(z^{-1})}{A_1^*(z^{-1})} \quad (79)$$

and we have assumed that

$$D(z^{-1}) = 1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_\delta z^{-\delta} \quad (80)$$

From (71) and (58) we see that our LS variant has the same accuracy as an ordinary LS estimator for the system (69) with input  $u_t = B_2(q^{-1})\eta_t$  and with  $E(\eta_t)^2 = R_\eta$ . A similar thing applies to the estimator for the parameter  $\theta$ .

With the same assumptions as above we find the asymptotic covariance matrix of the corresponding ordinary IV variant for the system described by (69) to be

$$P_\beta^{\text{OIV}} = R_v R_\eta^{-1} F_\beta^{-1} \quad (81)$$

When we compare this with (68) we see that the above conclusion also applies to the IV variant, i.e., the accuracy is the same as for the corresponding ordinary IV estimator for the system (69) when the input is  $u_t = B_2(q^{-1})\eta_t$  with  $E(\eta_t)^2 = R_\eta$ .

What has been shown above is not very surprising. Our prefiltering of the input/output data simply amounts to, as  $N$  tends to infinity, removing known dynamics, and we are left with the problem of estimating a reduced-order model with input  $B_2(q^{-1})u_t$  or  $B_1(q^{-1})u_t$  respectively.

It is easily verified that the above conclusion also holds when the input process  $\{u_t\}$  is not white. We summarize this in what follows.



## Fact 4

The accuracy of the LS and IV variants with input  $u_t$  is the same as the accuracy of the corresponding ordinary LS and IV variants, respectively, for the reduced-order models  $M_1$  and  $M_2$  with inputs given by  $B_2(q^{-1})u_t$  and  $B_1(q^{-1})u_t$  respectively.

## 4. Conclusion

We have considered the problem of evaluating the accuracy of certain robust estimators based upon filtering of the input/output data and employing certain decentralized estimation techniques. These estimators have in some earlier papers, e.g., Young *et al.* (1987) or Henriksen (1988, 1989), been shown to have superior robustness properties in situations where ordinary parameter estimators often fail, in particular when dealing with stiff systems.

The accuracy analysis presented herein reveals that the loss in accuracy may be small or even nothing compared with the corresponding ordinary LS or IV variants for the whole system. This is due to the fact that the variants herein are faced with a lower order model which should be expected to yield higher accuracy in the estimated parameters for a fixed sample size  $N$ . On the other hand, the estimators have to deal with inputs which are moving averages of the original inputs and this should lead to lower accuracy. The net outcome of these two effects could depend upon the actual system. However, the estimators have proved to be useful when the underlying system has an ill-conditions dynamics, e.g., when the underlying system is stiff. For stiff systems it has been shown that the convergence rate of the estimators will be very high, see Henriksen and Weyer (1990). This surely makes the estimators more attractive from a user's point of view.

## REFERENCES

- HENRIKSEN, R. (1988). Convergence analysis of some decentralized parameter 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 parameter estimators. *Modeling, Identification and Control*, **10**, 13–34.
- 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, USSR, August 13–17, Volume 3, pp. 215–220.
- LJUNG, L. (1977). Some limit results for functionals of stochastic processes. Department of Electrical Engineering, Linköping University, Linköping, Sweden, Report No. LiTH-ISY-I-0167.
- SÖDERSTRÖM, T. and STOICA, P. G. (1983). *Instrumental Variable Methods for System Identification* (Springer-Verlag).
- VETTER, W. J. (1970). Derivative operations on matrices. *IEEE Trans. on Automatic Control*, **15**, 241–244.
- VETTER, W. J. (1973). Matrix calculus operations and Taylor expansions. *SIAM Review*, **15**, 352–369.
- 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 Decision and Control*, Los Angeles, California, pp. 1902–1907.

## Appendix

## Proof of Theorem 1.

For  $N$  large enough we can clearly write the LS variant in the form

$$(N)^{1/2}[\hat{\beta}_N^{LS} - \beta^*] = R_{\psi, N}^{-1}(\hat{\theta}_N^{LS}) \left[ \frac{1}{(N)^{1/2}} \sum_{t=1}^N \psi_t(\hat{\theta}_N^{LS}) v_t \right] \{1 + o(1)\} \quad (\text{A.1})$$

$$(N)^{1/2}[\hat{\theta}_N^{LS} - \theta^*] = R_{\varphi, N}^{-1}(\hat{\beta}_N^{LS}) \left[ \frac{1}{(N)^{1/2}} \sum_{t=1}^N \varphi_t(\hat{\beta}_N^{LS}) v_t \right] \{1 + o(1)\} \quad (\text{A.2})$$

where

$$R_{\psi, N}(\hat{\theta}_N^{LS}) = \frac{1}{N} \sum_{t=1}^N \psi_t(\hat{\theta}_N^{LS}) \psi_t^T(\hat{\theta}_N^{LS}) \quad (\text{A.3})$$

$$R_{\varphi, N}(\hat{\beta}_N^{LS}) = \frac{1}{N} \sum_{t=1}^N \varphi_t(\hat{\beta}_N^{LS}) \varphi_t^T(\hat{\beta}_N^{LS}) \quad (\text{A.4})$$

Expanding the right-hand sides of (A.1) and (A.2) in Taylor series about  $\theta^*$  and  $\beta^*$ , respectively, we obtain

$$(N)^{1/2}[\hat{\beta}_N^{LS} - \beta^*] = R_{\psi, N}^{-1}(\theta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \psi_t(\theta^*) v_t + \Psi_N(\theta^*) [\hat{\theta}_N^{LS} - \theta^*] + \dots \quad (\text{A.5})$$

$$(N)^{1/2}[\hat{\theta}_N^{LS} - \theta^*] = R_{\varphi, N}^{-1}(\beta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \varphi_t(\beta^*) v_t + \Phi_N(\beta^*) [\hat{\beta}_N^{LS} - \beta^*] + \dots \quad (\text{A.6})$$

where

$$\begin{aligned} \Psi_N(\theta^*) = R_{\psi, N}^{-1}(\theta^*) \left\{ -\frac{\partial R_{\psi, N}}{\partial \hat{\theta}_N^{LS}}(\theta^*) \left[ I \otimes \left[ R_{\psi, N}^{-1}(\theta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \psi_t(\theta^*) v_t \right] \right. \right. \\ \left. \left. + \frac{1}{(N)^{1/2}} \sum_{t=1}^N \frac{\partial \psi_t}{\partial \hat{\theta}_N^{LS}}(\theta^*) v_t \right] \right\} \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \Phi_N(\beta^*) = R_{\varphi, N}^{-1}(\beta^*) \left\{ -\frac{\partial R_{\varphi, N}}{\partial \hat{\beta}_N^{LS}}(\beta^*) \left[ I \otimes \left[ R_{\varphi, N}^{-1}(\beta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \varphi_t(\beta^*) v_t \right] \right. \right. \\ \left. \left. + \frac{1}{(N)^{1/2}} \sum_{t=1}^N \frac{\partial \varphi_t}{\partial \hat{\beta}_N^{LS}}(\beta^*) v_t \right] \right\} \end{aligned} \quad (\text{A.8})$$

$$\frac{\partial R_{\psi, N}}{\partial \hat{\theta}_N^{LS}}(\theta^*) = \frac{1}{N} \sum_{t=1}^N \left[ \frac{\partial \psi_t}{\partial \hat{\theta}_N^{LS}}(\theta^*) [I \otimes \psi_t^T(\theta^*)] + \psi_t(\theta^*) \frac{\partial \psi_t^T}{\partial \hat{\theta}_N^{LS}}(\theta^*) \right] \quad (\text{A.9})$$

$$\frac{\partial R_{\varphi, N}}{\partial \hat{\beta}_N^{LS}}(\beta^*) = \frac{1}{N} \sum_{t=1}^N \left[ \frac{\partial \varphi_t}{\partial \hat{\beta}_N^{LS}}(\beta^*) [I \otimes \varphi_t^T(\beta^*)] + \varphi_t(\beta^*) \frac{\partial \varphi_t^T}{\partial \hat{\beta}_N^{LS}}(\beta^*) \right] \quad (\text{A.10})$$

In the above, where partial derivatives with respect to  $\hat{\beta}_N^{LS}$  and  $\hat{\theta}_N^{LS}$  appear, these vectors have to be interpreted as row vectors. Furthermore,  $\otimes$  denotes the Kronecker matrix product, see Vetter (1970, 1973).

As  $N$  tends to infinity,  $\hat{\beta}_N^{LS}$  and  $\hat{\theta}_N^{LS}$  tend to  $\beta^*$  and  $\theta^*$  respectively, and we obtain

$$(N)^{1/2}[\hat{\beta}_N^{LS} - \beta^*] \xrightarrow{N \rightarrow \infty} R_{\psi, N}^{-1}(\theta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \psi_t(\theta^*) v_t \quad (\text{A.11})$$

$$(N)^{1/2}[\hat{\theta}_N^{LS} - \theta^*] \xrightarrow{N \rightarrow \infty} R_{\varphi, N}^{-1}(\beta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \varphi_t(\beta^*) v_t \quad (\text{A.12})$$

By Lemma A4.1 in Söderström and Stoica (1983) (which is a variant of the central limit theorem due to Ljung (1977))

$$\frac{1}{(N)^{1/2}} \sum_{t=1}^N \psi_t(\theta^*) v_t$$

and

$$\frac{1}{(N)^{1/2}} \sum_{t=1}^N \varphi_t(\beta^*) v_t$$

are asymptotically Gaussian distributed with zero means and with covariance matrices given by, respectively,

$$Q_{\psi}(\theta^*) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=1}^N E \psi_t(\theta^*) v_t v_s \psi_s^T(\theta^*) \quad (\text{A.13})$$

$$Q_{\varphi}(\beta^*) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \sum_{s=1}^N E \varphi_t(\beta^*) v_t v_s \varphi_s^T(\beta^*) \quad (\text{A.14})$$

It follows that

$$(N)^{1/2}[\hat{\beta}_N^{LS} - \beta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_{\beta}^{LS} \quad (\text{A.15})$$

$$(N)^{1/2}[\hat{\theta}_N^{LS} - \theta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_{\theta}^{LS} \quad (\text{A.16})$$

where  $\gamma_{\beta}^{LS} \sim N(0, P_{\beta}^{LS})$  and  $\gamma_{\theta}^{LS} \sim N(0, P_{\theta}^{LS})$  and where the covariance matrices  $P_{\beta}^{LS}$  and  $P_{\theta}^{LS}$  are given by

$$P_{\beta}^{LS} = R_{\psi}^{-1}(\theta^*) Q_{\psi}(\theta^*) R_{\psi}^{-1}(\theta^*); \quad P_{\theta}^{LS} = R_{\varphi}^{-1}(\beta^*) Q_{\varphi}(\beta^*) R_{\varphi}^{-1}(\beta^*) \quad (\text{A.17})$$

where

$$R_{\psi}(\theta^*) = \lim_{N \rightarrow \infty} R_{\psi, N}(\theta^*) = E \psi_t(\theta^*) \psi_t^T(\theta^*) \quad (\text{A.18})$$

$$R_{\varphi}(\beta^*) = \lim_{N \rightarrow \infty} R_{\varphi, N}(\beta^*) = E \varphi_t(\beta^*) \varphi_t^T(\beta^*) \quad (\text{A.19})$$

The matrices  $Q_{\psi}(\theta^*)$  and  $Q_{\varphi}(\beta^*)$  can be evaluated somewhat further. Since the process  $\{v_t\}$  is white, only terms where  $t=s$  will be nonzero in (A.13) and (A.14). We thus obtain

$$Q_{\psi}(\theta^*) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_{\psi} \psi_t(\theta^*) \psi_t^T(\theta^*) = R_{\psi} E \psi_t(\theta^*) \psi_t^T(\theta^*) = R_{\psi} R_{\psi}(\theta^*) \quad (\text{A.20})$$

$$Q_{\varphi}(\beta^*) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_{\varphi} \varphi_t(\beta^*) \varphi_t^T(\beta^*) = R_{\varphi} E \varphi_t(\beta^*) \varphi_t^T(\beta^*) = R_{\varphi} R_{\varphi}(\beta^*) \quad (\text{A.21})$$

which completes the proof.

*Proof of Corollary 1.*

We have

$$\begin{aligned} \psi_t^T(\theta^*) = & \left[ -\frac{B^*(q^{-1})}{A_1^*(q^{-1})} u_{t-1}, \dots, -\frac{B^*(q^{-1})}{A_1^*(q^{-1})} u_{t-n_1}, B_2^*(q^{-1}) u_{t-1}, \dots, B_2^*(q^{-1}) u_{t-m_1} \right] \\ & + \left[ \frac{1}{A_1^*(q^{-1})} v_{t-1}, \dots, \frac{1}{A_1^*(q^{-1})} v_{t-n_1}, 0, 0, \dots, 0 \right] \end{aligned} \quad (\text{A.22})$$

$$\begin{aligned} \varphi_t(\theta^*) = & \left[ -\frac{B^*(q^{-1})}{A_2^*(q^{-1})} u_{t-1}, \dots, -\frac{B^*(q^{-1})}{A_2^*(q^{-1})} u_{t-n_2}, B_1^*(q^{-1}) u_{t-1}, \dots, B_1^*(q^{-1}) u_{t-m_2} \right] \\ & + \left[ \frac{1}{A_2^*(q^{-1})} v_{t-1}, \dots, \frac{1}{A_2^*(q^{-1})} v_{t-n_2}, 0, 0, \dots, 0 \right] \end{aligned} \quad (\text{A.23})$$

From this and the definitions preceding Corollary 1 we find that

$$R_\psi(\theta^*) = R_u G_\beta + R_v H_\beta; \quad R_\varphi(\beta^*) = R_u G_\theta + R_v H_\theta \quad (\text{A.24})$$

which completes the proof.

*Proof of Theorem 2.*

For  $N$  large enough we can write the IV variant in the form

$$(N)^{1/2} [\hat{\beta}_N^{\text{IV}} - \beta^*] = R_{\tilde{\psi}, N}^{-1}(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) \left[ \frac{1}{(N)^{1/2}} \sum_{t=1}^N \tilde{\psi}_t(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) v_t \right] \{1 + o(1)\} \quad (\text{A.25})$$

$$(N)^{1/2} [\hat{\theta}_N^{\text{IV}} - \theta^*] = R_{\tilde{\varphi}, N}^{-1}(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) \left[ \frac{1}{(N)^{1/2}} \sum_{t=1}^N \tilde{\varphi}_t(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) v_t \right] \{1 + o(1)\} \quad (\text{A.26})$$

where

$$R_{\tilde{\psi}, N}(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) = \frac{1}{N} \sum_{t=1}^N \tilde{\psi}_t(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) \psi_t^T(\hat{\theta}_N^{\text{IV}}) \quad (\text{A.27})$$

$$R_{\tilde{\varphi}, N}(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) = \frac{1}{N} \sum_{t=1}^N \tilde{\varphi}_t(\hat{\beta}_N^{\text{IV}}, \hat{\theta}_N^{\text{IV}}) \varphi_t^T(\hat{\beta}_N^{\text{IV}}) \quad (\text{A.28})$$

The proof is now carried out in detail for the estimator of  $\beta$ .

Expanding (A.25) in a Taylor series about  $(\beta^*, \theta^*)$  yields

$$\begin{aligned} (N)^{1/2} [\hat{\beta}_N^{\text{IV}} - \beta^*] = & R_{\tilde{\psi}, N}^{-1}(\beta^*, \theta^*) \frac{1}{(N)^{1/2}} \sum_{t=1}^N \tilde{\psi}_t(\beta^*, \theta^*) v_t \\ & + \Psi_{\beta, N}(\beta^*, \theta^*) [\hat{\beta}_N^{\text{IV}} - \beta^*] + \Psi_{\theta, N}(\beta^*, \theta^*) [\hat{\theta}_N^{\text{IV}} - \theta^*] + \dots \end{aligned} \quad (\text{A.29})$$

where

$$\Psi_{\beta, N}(\beta^*, \theta) = R_{\tilde{\psi}, N}^{-1}(\beta^*, \theta^*) \left\{ -\frac{\partial R_{\tilde{\psi}, N}}{\partial \beta_N^{IV}}(\beta^*, \theta^*) \left[ I \otimes \left[ R_{\tilde{\psi}, N}^{-1}(\beta^*, \theta^*) \right. \right. \right. \\ \left. \left. \left. \times \frac{1}{(N)^{1/2}} \sum_{i=1}^N \tilde{\psi}_i(\beta^*, \theta^*) v_i \right] \right] + \frac{1}{(N)^{1/2}} \sum_{i=1}^N \frac{\partial \tilde{\psi}_i}{\partial \beta_N^{IV}}(\beta^*, \theta^*) v_i \right\} \quad (\text{A.30})$$

$$\Psi_{\theta, N}(\beta^*, \theta^*) = R_{\tilde{\psi}, N}^{-1}(\beta^*, \theta^*) \left\{ -\frac{\partial R_{\tilde{\psi}, N}}{\partial \theta_N^{IV}}(\beta^*, \theta^*) \left[ I \otimes \left[ R_{\tilde{\psi}, N}^{-1}(\beta^*, \theta^*) \right. \right. \right. \\ \left. \left. \left. \times \frac{1}{(N)^{1/2}} \sum_{i=1}^N \tilde{\psi}_i(\beta^*, \theta^*) v_i \right] \right] + \frac{1}{(N)^{1/2}} \sum_{i=1}^N \frac{\partial \tilde{\psi}_i}{\partial \theta_N^{IV}}(\beta^*, \theta^*) v_i \right\} \quad (\text{A.31})$$

and where

$$\frac{\partial R_{\tilde{\psi}, N}}{\partial \theta_N^{IV}}(\beta^*, \theta^*) = \frac{1}{N} \sum_{i=1}^N \left[ \frac{\partial \tilde{\psi}_i}{\partial \theta_N^{IV}}(\beta^*, \theta^*) [I \otimes \psi_i^T(\theta^*)] \right] \quad (\text{A.32})$$

$$\frac{\partial R_{\tilde{\psi}, N}}{\partial \theta_N^{IV}}(\beta^*, \theta^*) = \frac{1}{N} \sum_{i=1}^N \left[ \frac{\partial \tilde{\psi}_i}{\partial \theta_N^{IV}}(\beta^*, \theta^*) [I \otimes \psi_i^T(\theta^*)] + \tilde{\psi}_i(\beta^*, \theta^*) \frac{\partial \psi_i^T}{\partial \theta_N^{IV}}(\theta^*) \right] \quad (\text{A.33})$$

As  $N$  tends to infinity  $\beta_N^{IV}$  and  $\theta_N^{IV}$  tend to  $\beta^*$  and  $\theta^*$  (by assumption), respectively, and we obtain

$$(N)^{1/2} [\beta_N^{IV} - \beta^*] \xrightarrow{N \rightarrow \infty} R_{\tilde{\psi}, N}^{-1}(\beta^*, \theta^*) \frac{1}{(N)^{1/2}} \sum_{i=1}^N \tilde{\psi}_i(\beta^*, \theta^*) v_i \quad (\text{A.34})$$

Again, by Lemma A4.1 in Söderström and Stoica (1983)

$$\frac{1}{(N)^{1/2}} \sum_{i=1}^N \tilde{\psi}_i(\beta^*, \theta^*) v_i$$

is asymptotically Gaussian distributed with zero mean and covariance matrix

$$Q_{\tilde{\psi}, N}(\beta^*, \theta^*) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sum_{s=1}^N E \tilde{\psi}_i(\beta^*, \theta^*) v_i v_s^T \tilde{\psi}_s^T(\beta^*, \theta^*) \quad (\text{A.35})$$

It follows that

$$(N)^{1/2} [\beta_N^{IV} - \beta^*] \xrightarrow[N \rightarrow \infty]{\text{distribution}} \gamma_{\beta}^{IV} \quad (\text{A.36})$$

where  $\gamma_{\beta}^{IV} \sim N(0, P_{\beta}^{IV})$  and where the covariance matrix  $P_{\beta}^{IV}$  is given by

$$P_{\beta}^{IV} = R_{\tilde{\psi}}^{-1}(\beta^*, \theta^*) Q_{\tilde{\psi}}(\beta^*, \theta^*) R_{\tilde{\psi}}^{-1}(\beta^*, \theta^*) \quad (\text{A.37})$$

where

$$R_{\tilde{\psi}}(\beta^*, \theta^*) = \lim_{N \rightarrow \infty} R_{\tilde{\psi}, N}(\beta^*, \theta^*) = E \tilde{\psi}_i(\beta^*, \theta^*) \psi_i^T(\theta^*) \\ = E \tilde{\psi}_i(\beta^*, \theta^*) \tilde{\psi}_i^T(\beta^*, \theta^*) \quad (\text{A.38})$$

We will now evaluate the matrix  $Q_{\psi}(\beta^*, \theta^*)$  further. Let  $\varphi_{vv}(\tau)$  denote the autocorrelation function of  $v$ , i.e.,

$$\varphi_{vv}(\tau) = E v_t v_{t+\tau} \quad (\text{A.39})$$

We can then write (A.35) in the form (see the proof of Theorem 5.1. in Söderström and Stoica (1983))

$$\begin{aligned} Q_{\psi}(\beta^*, \theta^*) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\tau=-N}^N (N-|\tau|) E \tilde{\psi}_t(\beta^*, \theta^*) \varphi_{vv}(\tau) \tilde{\psi}_{t+\tau}^T(\beta^*, \theta^*) \\ &= \sum_{\tau=-\infty}^{\infty} E \tilde{\psi}_t(\beta^*, \theta^*) \varphi_{vv}(\tau) \tilde{\psi}_{t+\tau}^T(\beta^*, \theta^*) \\ &\quad - \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\tau=-\infty}^{\infty} |\tau| E \tilde{\psi}_t(\beta^*, \theta^*) \varphi_{vv}(\tau) \tilde{\psi}_{t+\tau}^T(\beta^*, \theta^*) \end{aligned} \quad (\text{A.40})$$

where the first term on the right hand side converges due to the assumption of stationarity whereas the second term vanishes (see the above reference). Now, define the series

$$\sum_{i=-\infty}^{\infty} c_i z^{-i} = C(z^{-1}) \quad (\text{A.41})$$

where  $c_i = 0$  for  $i < 0$ . We obtain

$$\begin{aligned} Q_{\psi_t}(\beta^*, \theta^*) &= \sum_{\tau=-\infty}^{\infty} E \tilde{\psi}_t(\beta^*, \theta^*) \cdot E C(q^{-1}) \varepsilon_t C(q^{-1}) \varepsilon_{t+\tau} \cdot \tilde{\psi}_{t+\tau}^T(\beta^*, \theta^*) \\ &= \sum_{\tau=-\infty}^{\infty} E \tilde{\psi}_t(\beta^*, \theta^*) \left[ \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} c_i [E \varepsilon_{t-i} \varepsilon_{t+\tau-j}] c_j \right] \tilde{\psi}_{t+\tau}^T(\beta^*, \theta^*) \\ &= \sum_{\tau=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} E \tilde{\psi}_t(\beta^*, \theta^*) c_i R_{\varepsilon} c_{\tau+i} \tilde{\psi}_{t+\tau}^T(\beta^*, \theta^*) \\ &= R_{\varepsilon} \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} c_j E \tilde{\psi}_t(\beta^*, \theta^*) \tilde{\psi}_{t+j-i}^T(\beta^*, \theta^*) c_i \\ &= R_{\varepsilon} E C(q^{-1}) \tilde{\psi}_t(\beta^*, \theta^*) \cdot C(q^{-1}) \tilde{\psi}_t^T(\beta^*, \theta^*) \end{aligned} \quad (\text{A.42})$$

*Proof of Corollary 2.*

We simply observe that

$$\tilde{\psi}_t^T(\beta^*, \theta^*) = \left[ -\frac{B^*(q^{-1})}{A_1^*(q^{-1})} u_{t-1}, \dots, -\frac{B^*(q^{-1})}{A_1^*(q^{-1})} u_{t-n_1}, B_2^*(q^{-1}) u_{t-1}, \dots, B_2^*(q^{-1}) u_{t-m_1} \right] \quad (\text{A.43})$$

$$\tilde{\varphi}_t^T(\beta^*, \theta^*) = \left[ -\frac{B^*(q^{-1})}{A_2^*(q^{-1})} u_{t-1}, \dots, -\frac{B^*(q^{-1})}{A_2^*(q^{-1})} u_{t-n_2}, B_1^*(q^{-1}) u_{t-1}, \dots, u_{t-m_2} \right] \quad (\text{A.44})$$