A RECURSIVE FRISCH SCHEME ALGORITHM FOR COLOURED OUTPUT NOISE

J. G. Linden and K. J. Burnham
Control Theory and Applications Centre, Coventry University, Priory Street, Coventry, U.K.
j.linden@coventry.ac.uk

Keywords: Adaptive Estimation, Errors-in-varibles, Recursive Identification, System Identification.

Abstract: A recursive (adaptive) algorithm for the identification of dynamical linear errors-in-variables systems in the case of coloured output noise is developed. The input measurement noise variance as well as the auto-covariance elements of the coloured output noise sequence are determined via two separate Newton algorithms. The model parameter estimates are obtained by a recursive bias-compensating instrumental variables algorithm with past noisy inputs as instruments, thus allowing the compensation for the explicitly computed bias at each discrete-time instance. The performance of the developed algorithm is demonstrated via simulation.

1 INTRODUCTION

Linear time-invariant (LTI) errors-in-variables (EIV) models are characterised by an exact linear relationship between input and output signals where both quantities are assumed to be corrupted by additive measurement noise (Söderström, 2007b). An EIV model representation can be advantageous, if the aim is to gain a better understanding of the underlying process rather than prediction. One interesting approach for the identification of dynamical systems within this framework is the so-called Frisch scheme (Beghelli et al., 1990), which yields estimates of the model parameters as well as the measurement noise variances. The dynamic Frisch scheme presented in (Beghelli et al., 1990; Söderström, 2007a) assumes that the additive disturbances on the system input and output are white. Such an assumption, however, can be rather restrictive since the output noise often not solely consists of measurement uncertainties, but also aims to account for process disturbances, which are usually correlated in time. In order to overcome this shortcoming, the Frisch scheme has recently been extended to the coloured output noise case (Söderström, 2008). This paper develops a recursive (adaptive) formulation of the algorithm developed in (Söderström, 2008), which allows the estimates to be calculated online as new data arrives. Recursive algorithms for the white noise case have been considered in (Linden et al., 2008; Linden et al., 2007).

The paper is organised as follows. Section 2 presents the EIV identification problem and introduces some notational conventions. Section 3 reviews the offline Frisch scheme procedure for the white noise as well as the coloured noise case. Section 4 develops the recursive algorithm and Section 5 provides a numerical example. Conclusions are given in Section 6.

2 PROBLEM STATEMENT AND NOTATION

In this paper, a discrete-time, LTI single-input single-output (SISO) EIV system is considered, which is described by

$$A(q^{-1})y_{0_i} = B(q^{-1})u_{0_i}, \tag{1}$$

where i is an integer valued time index and

$$A(q^{-1}) \triangleq 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a}, \tag{2a}$$
$$B(q^{-1}) \triangleq b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \tag{2b}$$

are polynomials in the backward shift operator q^{-1}, which is defined such that $x_i q^{-1} = x_{i-1}$. The noise-free input u_{0_i} and output y_{0_i} are unknown and only the measurements

$$u_i = u_{0_i} + \tilde{u}_i, \tag{3a}$$
$$y_i = y_{0_i} + \tilde{y}_i \tag{3b}$$

are available, where $\tilde{u}_i$ and $\tilde{y}_i$ denote input and output measurement noise, respectively. Such a setup is depicted in Figure 1. The following assumptions are introduced:

A1. The dynamic system (1) is asymptotically stable, i.e. $A(q^{-1})$ has all zeros inside the unit circle.

A2. All system modes are observable and controllable, i.e. $A(q^{-1})$ and $B(q^{-1})$ have no common factors.

A3. The polynomial degrees n_a and n_b are known *a priori* with $n_b \leq n_a$.

A4. The true input u_{0_i} is a zero-mean ergodic process and is persistently exciting of sufficiently high order.

A5a. The sequence $\tilde{u}_i$ is a zero-mean, ergodic, white noise process with unknown variance $\sigma_{\tilde{u}}$.

A5b. The sequence $\tilde{y}_i$ is a zero-mean, ergodic noise process with unknown auto-covariance sequence $\{r_{\tilde{y}}(0), r_{\tilde{y}}(1), \cdots\}$.

A6. The noise sequences $\tilde{u}_i$ and $\tilde{y}_i$ are mutually uncorrelated and uncorrelated with u_{0_i}.

The auto-covariance elements in A5b are defined by

$$r_{\tilde{y}}(\tau) \triangleq E\left[\tilde{y}_k \tilde{y}_{k-\tau}\right], \tag{4}$$

where $E[\cdot]$ denotes the expected value operator. Within this paper covariance matrices of two column vectors v_k and w_k are denoted

$$\Sigma_{vw} \triangleq E\left[v_k w_k^T\right], \qquad \Sigma_v \triangleq E\left[v_k v_k^T\right], \tag{5}$$

whilst vectors consisting of covariance elements are denoted

$$\xi_{vc} \triangleq E\left[v_k c_k\right] \tag{6}$$

with c_k being a scalar. The corresponding estimated sample covariance elements are denoted in a similar manner

$$\hat{\Sigma}_{vw}^k \triangleq \frac{1}{k}\sum_{i=1}^{k} v_k w_k^T, \quad \hat{\Sigma}_{v}^k \triangleq \frac{1}{k}\sum_{i=1}^{k} v_k v_k^T, \quad \hat{\xi}_{vc}^k \triangleq \frac{1}{k}\sum_{i=1}^{k} v_k c_k. \tag{7}$$

In addition, the parameter vectors are formed by

$$\theta \triangleq \begin{bmatrix} a^T & b^T \end{bmatrix}^T = \begin{bmatrix} a_1 & \dots & a_{n_a} & b_1 & \dots & b_{n_b} \end{bmatrix}^T, \tag{8a}$$

$$\bar{\theta} \triangleq \begin{bmatrix} \bar{a}^T & b^T \end{bmatrix}^T = \begin{bmatrix} 1 & \theta^T \end{bmatrix}^T, \tag{8b}$$

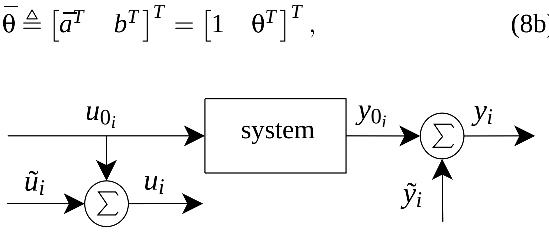

Figure 1: Errors-in-variables setup.

which gives an alternative description of (1)-(3) by

$$\bar{\varphi}_{0_i}^T \bar{\theta} = 0, \tag{9a}$$

$$\bar{\varphi}_i = \bar{\varphi}_{0_i} + \bar{\tilde{\varphi}}_i, \tag{9b}$$

where the regression vector is given by

$$\varphi_i \triangleq \begin{bmatrix} \varphi_{y_i}^T & \varphi_{u_i}^T \end{bmatrix}^T \tag{10}$$
$$\triangleq \begin{bmatrix} -y_{i-1} & \dots & -y_{i-n_a} & u_{i-1} & \dots & u_{i-n_b} \end{bmatrix}^T,$$

$$\bar{\varphi}_i \triangleq \begin{bmatrix} \bar{\varphi}_{y_i}^T & \varphi_{u_i}^T \end{bmatrix}^T \triangleq \begin{bmatrix} -y_i & \varphi_i^T \end{bmatrix}^T. \tag{11}$$

The noise-free regression vectors φ_{0_i}, $\bar{\varphi}_{0_i}$ and the vectors containing the noise contributions $\tilde{\varphi}_i$, $\bar{\tilde{\varphi}}_i$ are defined in a similar manner. The identification problem is now given by:

Problem 1. *Given k samples of noisy input-output data* $\{u_1, y_1, ..., u_k, y_k\}$, *determine an estimate of the augmented parameter vector*

$$\vartheta \triangleq \begin{bmatrix} a_1 & \dots & a_{n_a} & b_1 & \dots & b_{n_b} & \sigma_{\tilde{u}} \\ & r_{\tilde{y}}(0) & \cdots & r_{\tilde{y}}(n_a) \end{bmatrix}^T. \tag{12}$$

Throughout this paper, the convention is made that estimated quantities are marked by a ˆ whilst time dependent quantities have a sub- or superscript k, e.g. $\hat{\Sigma}_{\varphi}^k$ for a sample covariance matrix corresponding to Σ_{φ}.

3 FRISCH SCHEME

3.1 White Noise Case

If the least squares (LS) estimator is directly applied to estimate the parameters of the EIV system (1)-(3), the estimates will generally be biased (Söderström, 2007a). However, if the statistical nature of the noise sequences is known, it is possible to compensate for the bias. The Frisch scheme belongs to the class of such bias-compensating LS techniques. The compensated normal equations are given by

$$\left(\hat{\Sigma}_{\varphi}^k - \hat{\Sigma}_{\tilde{\varphi}}^k\right)\hat{\theta}_k = \hat{\xi}_{\varphi y}^k, \tag{13}$$

where $\hat{\Sigma}_{\varphi}^k$ and $\hat{\xi}_{\varphi y}^k$ are defined by (7). In the case of white noise sequences the compensating matrix $\hat{\Sigma}_{\tilde{\varphi}}^k$ is diagonal and given by

$$\begin{bmatrix} \hat{r}_{\tilde{y}}^k(0) I_{n_a} & 0 \\ 0 & \hat{\sigma}_{\tilde{u}}^k I_{n_b} \end{bmatrix}, \tag{14}$$

where I_n denotes the identity matrix of dimension n. Within the Frisch scheme, the variances $\hat{\sigma}_{\tilde{u}}^k$, $\hat{r}_{\tilde{y}}^k(0)$ of input and output measurement noise, respectively, are determined such that the extended compensated normal equations equate to zero

$$0 = \left(\hat{\Sigma}^k_{\bar{\varphi}} - \hat{\Sigma}^k_{\tilde{\varphi}}\right)\hat{\bar{\theta}}_k \tag{15}$$
$$= \left(\begin{bmatrix} \hat{\Sigma}^k_{\bar{\varphi}_y} & \hat{\Sigma}^k_{\bar{\varphi}_y\varphi_u} \\ \hat{\Sigma}^k_{\varphi_u\bar{\varphi}_y} & \hat{\Sigma}^k_{\varphi_u} \end{bmatrix} - \begin{bmatrix} \hat{r}^k_{\tilde{y}}(0)I_{n_a+1} & 0 \\ 0 & \hat{\sigma}^k_{\tilde{u}}I_{n_b} \end{bmatrix}\right)\hat{\bar{\theta}}_k,$$

i.e. such that $\hat{\Sigma}^k_{\bar{\varphi}} - \hat{\Sigma}^k_{\tilde{\varphi}}$ is singular. By utilising the Schur complement, the input noise variance can be expressed as a nonlinear function of the output noise variance and vice versa (Beghelli et al., 1990)

$$\hat{r}^k_{\tilde{y}}(0) = \lambda_{\min}\left(\hat{\Sigma}^k_{\bar{\varphi}_y} - \hat{\Sigma}^k_{\bar{\varphi}_y\varphi_u}\left[\hat{\Sigma}^k_{\varphi_u} - \sigma_{\tilde{u}}I_{n_b}\right]^{-1}\hat{\Sigma}^k_{\varphi_u\bar{\varphi}_y}\right), \tag{16a}$$

$$\hat{\sigma}^k_{\tilde{u}} = \lambda_{\min}\left(\hat{\Sigma}^k_{\varphi_u} - \hat{\Sigma}^k_{\varphi_u\bar{\varphi}_y}\left[\hat{\Sigma}^k_{\bar{\varphi}_y} - r_{\tilde{y}}(0)I_{n_a+1}\right]^{-1}\hat{\Sigma}^k_{\bar{\varphi}_y\varphi_u}\right), \tag{16b}$$

where $\lambda_{\min}$ denotes the minimum eigenvalue operator. Equation (16) together with (15) defines a whole set of possible solutions depending on the choice of $\sigma_{\tilde{u}}$ or $r_{\tilde{y}}(0)$, respectively. In order to uniquely solve the identification problem, another equation is required. Several choices are discussed in (Hong et al., 2007).

3.2 Coloured Noise Case

Now assume that $\tilde{y}_k$ is no longer white, i.e. it is correlated or coloured. For this case, the matrices $\hat{\Sigma}^k_{\bar{\varphi}}$ and $\hat{\Sigma}^k_{\tilde{\varphi}}$ in (15) can be expressed in block form as

$$\hat{\Sigma}^k_{\bar{\varphi}} = \begin{bmatrix} \times & \times & \times \\ -\hat{\xi}^k_{\varphi_y y} & \hat{\Sigma}^k_{\varphi_y} & \hat{\Sigma}^k_{\varphi_y\varphi_u} \\ -\hat{\xi}^k_{\varphi_u y} & \hat{\Sigma}^k_{\varphi_u\varphi_y} & \hat{\Sigma}^k_{\varphi_u} \end{bmatrix}, \tag{17a}$$

$$\hat{\Sigma}^k_{\tilde{\varphi}} = \begin{bmatrix} \times & \times & \times \\ -\hat{\xi}^k_{\tilde{\varphi}_y\tilde{y}} & \hat{\Sigma}^k_{\tilde{\varphi}_y} & 0 \\ 0 & 0 & \hat{\sigma}_{\tilde{u}}I^k_{n_b} \end{bmatrix}, \tag{17b}$$

where the first row consists of arbitrary entries $\times$ and

$$\hat{\Sigma}^k_{\tilde{\varphi}_y} = \begin{bmatrix} \hat{r}^k_{\tilde{y}}(0) & \cdots & \hat{r}^k_{\tilde{y}}(n_a-1) \\ \vdots & \ddots & \vdots \\ \hat{r}^k_{\tilde{y}}(n_a-1) & \cdots & \hat{r}^k_{\tilde{y}}(0) \end{bmatrix} \tag{18}$$

is a dense matrix, whilst the remaining entries in (17) are in accordance with (7). Consequently, the n_a+n_b compensated normal equations in the case of correlated output noise are given by

$$\left(\begin{bmatrix} \hat{\Sigma}^k_{\varphi_y} & \hat{\Sigma}^k_{\varphi_y\varphi_u} \\ \hat{\Sigma}^k_{\varphi_u\varphi_y} & \hat{\Sigma}^k_{\varphi_u} \end{bmatrix} - \begin{bmatrix} \hat{\Sigma}^k_{\tilde{\varphi}_y} & 0 \\ 0 & \hat{\sigma}^k_{\tilde{u}}I_{n_b} \end{bmatrix}\right)\hat{\theta}_k = \begin{bmatrix} \hat{\xi}^k_{\varphi_y y} - \hat{\xi}^k_{\tilde{\varphi}_y\tilde{y}} \\ \hat{\xi}^k_{\varphi_u y} \end{bmatrix}. \tag{19}$$

Now consider the Frisch equation (16b) which becomes

$$\hat{\sigma}^k_{\tilde{u}} = \lambda_{\min}\left(B_k\right), \tag{20}$$

with

$$B_k \triangleq \hat{\Sigma}^k_{\varphi_u} - \hat{\Sigma}^k_{\varphi_u\bar{\varphi}_y}\left[\hat{\Sigma}^k_{\bar{\varphi}_y} - \hat{\Sigma}^k_{\tilde{\bar{\varphi}}_y}\right]^{-1}\hat{\Sigma}^k_{\bar{\varphi}_y\varphi_u} \tag{21}$$

and it remains to specify n_a+1 equations for the determination the auto-covariance elements

$$\hat{\rho}^k_y \triangleq \left[\hat{r}^k_{\tilde{y}}(0) \quad \hat{r}^k_{\tilde{y}}(1) \quad \cdots \quad \hat{r}^k_{\tilde{y}}(n_a)\right]^T. \tag{22}$$

In (Söderström, 2008) several possibilities to obtain the remaining equations are discussed. It is shown that a covariance-matching criterion, as used in (Diversi et al., 2003), as well as correlating the residuals with past outputs, which corresponds to an instrumental variable (IV) -like approach with outputs as instruments, cannot be successful since it always leads to more unknowns than equations. However, by correlating the residuals, denoted ε_k, with past inputs, the remaining equations are obtained for the asymptotic case via

$$E\left[\bar{\zeta}_k\varepsilon_k\right] = 0, \tag{23}$$

where the instruments are given by

$$\bar{\zeta}_k = \left[u_{k-n_b-1} \quad \cdots \quad u_{k-n_b-l}\right]^T \tag{24}$$

and the residuals are obtained via

$$\varepsilon_k = A(q^{-1})y_k - B(q^{-1})u_k = y_k - \varphi_k^T\theta. \tag{25}$$

This yields

$$\xi_{\bar{\zeta}y} - \Sigma_{\bar{\zeta}\varphi}\theta = 0, \tag{26}$$

which can be expressed in block form, and using sample statistics, as

$$\begin{bmatrix} \hat{\Sigma}^k_{\bar{\zeta}\varphi_y} & \hat{\Sigma}^k_{\bar{\zeta}\varphi_u} \end{bmatrix}\hat{\theta}_k = \hat{\xi}^k_{\bar{\zeta}y}, \tag{27}$$

where the length l of the instrument vector $\bar{\zeta}_k$ must satisfy $l \geq n_a+1$ in order to obtain at least n_a+1 additional equations for the determination of $\hat{\rho}^k_y$.

In (Söderström, 2008), two algorithms have been proposed to solve the resulting (nonlinear) estimation problem. Here, the two step algorithm, which makes use of the separable LS technique is considered. Whilst in the white noise case the estimate of θ is obtained from the compensated normal equations after the noise variances have been determined, this approach is conceptually different as outlined in the remainder of this Section.

3.2.1 Step 1

Note that $\hat{\rho}_y^k$ only appears in the first n_a equations of (19) and by combining the last n_b equations of the LS normal equations (19) with the n_a+1 IV equations (27), one can express

$$\begin{bmatrix} \hat{\Sigma}^k_{\varphi_u\varphi_y} & \hat{\Sigma}^k_{\varphi_u} - \hat{\sigma}^k_{\tilde{u}} I_{n_b} \\ \hat{\Sigma}^k_{\underline{\zeta}\varphi_y} & \hat{\Sigma}^k_{\underline{\zeta}\varphi_u} \end{bmatrix} \hat{\theta}_k = \begin{bmatrix} \hat{\xi}^k_{\varphi_u y} \\ \hat{\xi}^k_{\underline{\zeta} y} \end{bmatrix}. \quad (28)$$

which constitute n_a+n_b+1 equations in n_a+n_b+1 unknowns ($\hat{\theta}$ and $\hat{\sigma}^k_{\tilde{u}}$). Equation (28) is an overdetermined system of normal equations with its first part obtained from the bias compensated LS and the second part given by the IV estimator, which uses delayed inputs. Moreover, it is nonlinear due to the multiplication of $\hat{\theta}_k$ with $\hat{\sigma}^k_{\tilde{u}}$.

In order to estimate θ and $\sigma_{\tilde{u}}$, (28) can be re-expressed as

$$\left(\hat{\Sigma}^k_{\underline{\delta}\varphi} - \hat{\sigma}^k_{\tilde{u}} \bar{J}\right) \hat{\theta}_k = \hat{\xi}^k_{\underline{\delta} y}, \quad (29)$$

where $\hat{\Sigma}^k_{\underline{\delta}\varphi}$ and $\hat{\xi}^k_{\underline{\delta} y}$ are defined by (7) with

$$\begin{aligned} \bar{\delta}_i &\triangleq \begin{bmatrix} \varphi_{u_i}^T & \bar{\zeta}_i^T \end{bmatrix}^T \\ &= \begin{bmatrix} u_{i-1} & \dots & u_{i-n_b} & u_{i-n_b-1} & \dots & u_{i-n_b-l} \end{bmatrix}^T, \end{aligned} \quad (30)$$

whilst $\bar{J}$ is given by

$$\bar{J} \triangleq \begin{bmatrix} 0 & I_{n_b} \\ 0 & 0 \end{bmatrix}. \quad (31)$$

Note that (29) can be interpreted as a bias-compensated IV approach, where the instrument vector $\bar{\delta}_i$ is constructed from past measured inputs. Introducing for convenience

$$\hat{G}_k \triangleq \hat{\Sigma}^k_{\underline{\delta}\varphi} - \hat{\sigma}^k_{\tilde{u}} \bar{J}, \quad (32)$$

the estimates for $\sigma^k_{\tilde{u}}$ and θ_k are obtained by minimising the (nonlinear) LS costfunction

$$\min_{\hat{\theta}_k, \hat{\sigma}^k_{\tilde{u}}} \left|\left| \hat{G}_k \hat{\theta}_k - \hat{\xi}^k_{\underline{\delta} y} \right|\right|^2 \quad (33)$$

which is minimised w.r.t. $\sigma^k_{\tilde{u}}$ and θ_k. If $\hat{\sigma}^k_{\tilde{u}}$ is assumed to be fixed, an explicit expression for $\hat{\theta}_k$ is given by the well-known LS solution

$$\hat{\theta}_k = \hat{G}_k^\dagger \hat{\xi}^k_{\underline{\delta} y}, \quad (34)$$

where $\hat{G}_k^\dagger \triangleq (\hat{G}_k^T \hat{G}_k)^{-1} \hat{G}_k^T$ denotes the Moore-Penrose pseudo inverse. Using the separable LS approach (Ljung, 1999, p. 335), the problem is reduced to an optimisation in one variable only by substituting (34) in (33). Consequently, $\hat{\sigma}^k_{\tilde{u}}$ can be obtained via

$$\hat{\sigma}^k_{\tilde{u}} = \arg\min_{\hat{\sigma}^k_{\tilde{u}}} V_k \quad (35)$$

with

$$\begin{aligned} V_k &= \left|\left| \hat{G}_k \hat{G}_k^\dagger \hat{\xi}^k_{\underline{\delta} y} - \hat{\xi}^k_{\underline{\delta} y} \right|\right|^2 \\ &= \left[\hat{\xi}^k_{\underline{\delta} y}\right]^T \hat{\xi}^k_{\underline{\delta} y} - \left[\hat{\xi}^k_{\underline{\delta} y}\right]^T \hat{G}_k \left[\hat{G}_k^T \hat{G}_k\right]^{-1} \hat{G}_k^T \hat{\xi}^k_{\underline{\delta} y}. \end{aligned} \quad (36)$$

Once $\hat{\sigma}^k_{\tilde{u}}$ is obtained, $\hat{\theta}_k$ is given by (34). Since the solution of (35) should satisfy $V_k = 0$, the value of V_k indicates whether the optimisation algorithm has converged to a global or local minimum (Söderström, 2008).

3.2.2 Step 2

In order to determine the estimates for the auto-correlation sequence $\hat{\rho}_y^k$ the remaining n_a normal equations

$$\begin{bmatrix} \hat{\Sigma}^k_{\varphi_y} - \hat{\Sigma}^k_{\tilde{\varphi}_y} & \hat{\Sigma}^k_{\varphi_y\varphi_u} \end{bmatrix} \hat{\theta}_k = \hat{\xi}^k_{\varphi_y y} - \hat{\xi}^k_{\tilde{\varphi}_y \tilde{y}} \quad (37)$$

together with the Frisch equation (20) are considered. Equation (37) can be expressed as

$$\hat{\Sigma}^k_{\tilde{\varphi}_y} \hat{a}_k - \hat{\xi}^k_{\tilde{\varphi}_y \tilde{y}} = \begin{bmatrix} \hat{\Sigma}^k_{\varphi_y} & \hat{\Sigma}^k_{\varphi_y\varphi_u} \end{bmatrix} \hat{\theta}_k - \hat{\xi}^k_{\varphi_y y}, \quad (38)$$

where only the left hand side depends on $\hat{\rho}_y^k$. In addition, (38) is affine in $\hat{\rho}_y^k$, hence it can be re-expressed as

$$H_k \hat{\rho}_y^k = h_k, \quad (39)$$

where H_k is a $n_a \times n_a+1$ matrix built up from elements of $\hat{\bar{a}}_k$ and h_k is a vector of length n_a given by the right hand side of (38). This is a system of equations with more unknowns than equations, but the set of all possible solutions can be formalised as

$$\rho_y^k = \alpha_k N(H_k) + H_k^\dagger h_k, \quad (40)$$

where $N(\cdot)$ denotes the nullspace and α_k is a scalar factor. It is necessary to distinguish between the input measurement noise variance obtained by (35) in step 1, and the quantity which would be obtained by the Frisch equation (20). Therefore, introduce

$$\hat{\varsigma}_k \triangleq \lambda_{\min}\left(B_k(\alpha_k)\right), \quad (41)$$

where the matrix B_k is now a function of α_k. Using (41) it is possible to search for that α_k which is in best agreement with the previously determined $\hat{\sigma}^k_{\tilde{u}}$, i.e.

$$\hat{\alpha}_k = \arg\min_{\alpha_k} ||J_k||_2^2, \quad (42)$$

where the cost function

$$J_k \triangleq \hat{\sigma}_{\tilde{u}}^k - \hat{\varsigma}_k \tag{43}$$

measures the distance between the input noise variance estimate $\hat{\sigma}_{\tilde{u}}^k$ determined in Step 1 and the input noise variance estimate $\hat{\varsigma}_k$ which is obtained using the n_a normal equations (37) together with the Frisch equation (41) depending on the choice of α_k. Once $\hat{\alpha}_k$ is determined, it is substituted in (40) to obtain $\hat{\rho}_y^k$, the searched estimate of the auto-covariance elements of the coloured output measurement noise $\tilde{y}_k$.

4 RECURSIVE SCHEME

4.1 Step 1

4.1.1 Recursive Update of Covariance Matrices

In order to satisfy the requirements of a recursive algorithm to store all data in a finite dimensional vector, the covariance matrices are updated via

$$\hat{\Sigma}_{\bar{\varphi}}^k = \hat{\Sigma}_{\bar{\varphi}}^{k-1} + \gamma_k \left(\bar{\varphi}_k \bar{\varphi}_k^T - \hat{\Sigma}_{\bar{\varphi}}^{k-1} \right), \quad \hat{\Sigma}_{\bar{\varphi}}^0 = 0, \tag{44a}$$

$$\hat{\Sigma}_{\bar{\zeta}\bar{\varphi}}^k = \hat{\Sigma}_{\bar{\zeta}\bar{\varphi}}^{k-1} + \gamma_k \left(\bar{\zeta}_k \bar{\varphi}_k^T - \hat{\Sigma}_{\bar{\zeta}\bar{\varphi}}^{k-1} \right), \quad \hat{\Sigma}_{\bar{\zeta}\bar{\varphi}}^0 = 0, \tag{44b}$$

where the normalising gain γ_k is given by

$$\gamma_k \triangleq \frac{\gamma_{k-1}}{\lambda + \gamma_{k-1}}, \qquad \gamma_0 = 1 \tag{45}$$

with $0 < \lambda \leq 1$ being the forgetting factor giving exponential forgetting. From (44), the block matrices required in (28) and (37) are readily obtained.

4.1.2 Recursive Update of $\hat{\sigma}_{\tilde{u}}^k$

For the determination of $\hat{\sigma}_{\tilde{u}}^k$, an iterative optimisation procedure can be utilised to minimise (36) where it is iterated once at each step, leading to a recursive scheme (Ljung and Söderström, 1983; Ljung, 1999). Here, an iterative Newton method is utilised for this purpose, however other choices are also possible. The Newton method given by (Ljung, 1999, p. 326) is

$$\sigma_{\tilde{u}}^k = \sigma_{\tilde{u}}^{k-1} - \left[V_k'' \right]^{-1} V_k', \tag{46}$$

where V_k' and V_k'' denote the first and second order derivative of V_k with respect to $\sigma_{\tilde{u}}^k$ evaluated at $\sigma_{\tilde{u}}^{k-1}$. The formulas for the derivatives are given in Appendices A and B, respectively.

Remark 1. *In order to stabilise the algorithm, it might be advantageous to restrict the search for the input measurement noise variance to the interval*

$$0 \leq \sigma_{\tilde{u}} \leq \sigma_{\tilde{u}}^{max}, \tag{47}$$

where $\sigma_{\tilde{u}}^{max}$ is the maximal admissible value for $\sigma_{\tilde{u}}$, which can be computed from the data as discussed in (Beghelli et al., 1990). Alternatively, a positive constant can be chosen for the maximum admissible value, if such a-priori knowledge is available.

4.1.3 Recursive Update of $\hat{\theta}_k$

In order to obtain a recursive expression for $\hat{\theta}_k$, an approach is adopted here, similar to that in (Ding et al., 2006), where the bias of the recursive LS estimate is compensated at each time step k.

Ignoring the influence of $\hat{\sigma}_{\tilde{u}}^k$ in (28), the uncompensated overdetermined IV normal equations can be expressed as

$$\frac{1}{k} \sum_{i=1}^{k} \begin{bmatrix} \varphi_{u_i} \\ \zeta_i \end{bmatrix} \begin{bmatrix} \varphi_{y_i}^T & \varphi_{u_i}^T \end{bmatrix} \hat{\theta}_k^{\mathrm{IV}} = \frac{1}{k} \sum_{i=1}^{k} \begin{bmatrix} \varphi_{u_i} \\ \zeta_i \end{bmatrix} y_i, \tag{48}$$

where $\hat{\theta}_k^{\mathrm{IV}}$ denotes the uncompensated (biased) estimate of θ. Since one unknown, namely $\hat{\sigma}_{\tilde{u}}^k$, has already been obtained, it is sufficient to consider $n_a + n_b$ equations only[1], by disregarding the last equation of (48). Thus the uncompensated IV estimate is given as

$$\hat{\theta}_k^{\mathrm{IV}} = \left[\frac{1}{k} \sum_{i=1}^{k} \delta_i \varphi_i^T \right]^{-1} \frac{1}{k} \sum_{i=1}^{k} \delta_i y_i, \tag{49}$$

where δ_i is obtained by deleting the last entry of $\bar{\delta}_i$. In order to obtain an explicit expression for the bias, the linear regression formulation

$$y_i = \varphi_i^T \theta + e_i \tag{50}$$

is substituted in (49) which gives

$$\begin{aligned} \hat{\theta}_k^{\mathrm{IV}} &= \left[\frac{1}{k} \sum_{i=1}^{k} \delta_i \varphi_i^T \right]^{-1} \frac{1}{k} \sum_{i=1}^{k} \delta_i \left(\varphi_i^T \theta + e_i \right) \\ &= \theta + \left[\frac{1}{k} \sum_{i=1}^{k} \delta_i \varphi_i^T \right]^{-1} \frac{1}{k} \sum_{i=1}^{k} \delta_i e_i \end{aligned} \tag{51}$$

By substituting $e_i = -\tilde{\varphi}_i \theta + \tilde{y}_i$ it follows that

$$\begin{aligned} \hat{\theta}_k^{\mathrm{IV}} = \theta &+ \left[\frac{1}{k} \sum_{i=1}^{k} \delta_i \varphi_i^T \right]^{-1} \frac{1}{k} \sum_{i=1}^{k} \delta_i \tilde{y}_i \\ &- \left[\frac{1}{k} \sum_{i=1}^{k} \delta_i \varphi_i^T \right]^{-1} \frac{1}{k} \sum_{i=1}^{k} \delta_i \tilde{\varphi}_i^T \theta. \end{aligned} \tag{52}$$

The vector δ_i is uncorrelated with $\tilde{y}_i$ which means that the middle part of the sum in (52) diminishes in the

[1] This corresponds to a basic IV estimator where the number of unknowns is equal to the length of the instrument vector.

asymptotic case, whereas

$$\lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k}\begin{bmatrix}\varphi_{u_i}\\ \zeta_i\end{bmatrix}\begin{bmatrix}\tilde{\varphi}_{y_i}^T & \tilde{\varphi}_{u_i}^T\end{bmatrix}^T = \begin{bmatrix}0 & \sigma_{\tilde{u}} I_{n_b}\\ 0 & 0\end{bmatrix}. \quad (53)$$

Consequently, for $k \to \infty$ (52) becomes

$$\theta^{\mathrm{IV}} = \theta - \sigma_{\tilde{u}} \Sigma_{\delta\varphi}^{-1} J\theta, \quad (54)$$

where J is obtained by deleting the last row of $\bar{J}$ in (31). Equation (54) gives rise to the recursive bias compensation update equation for $\hat{\theta}_k$

$$\hat{\theta}_k = \hat{\theta}_k^{\mathrm{IV}} + \hat{\sigma}_{\tilde{u}}^k \left[\hat{\Sigma}_{\delta\varphi}^k\right]^{-1} J\hat{\theta}_{k-1}, \quad (55)$$

where the uncompensated parameter estimate $\hat{\theta}_k^{\mathrm{IV}}$ can be recursively computed via a recursive IV (RIV) algorithm (Ljung, 1999, p. 369) given by

$$\hat{\theta}_k^{\mathrm{IV}} = \hat{\theta}_{k-1}^{\mathrm{IV}} + L_k\left[y_k - \varphi_k^T \hat{\theta}_{k-1}^{\mathrm{IV}}\right], \quad (56a)$$

$$L_k = \frac{P_{k-1}\delta_k}{\frac{1-\gamma_k}{\gamma_k} + \varphi_k^T P_{k-1}\delta_k}, \quad (56b)$$

$$P_k = \frac{1}{1-\gamma_k}\left[P_{k-1} - \frac{P_{k-1}\delta_k\varphi_k^T P_{k-1}}{\frac{1-\gamma_k}{\gamma_k} + \varphi_k^T P_{k-1}\delta_k}\right]. \quad (56c)$$

with the only difference being that P_k is scaled such that

$$\left[\hat{\Sigma}_{\delta\varphi}^k\right]^{-1} = P_k. \quad (57)$$

This avoids the matrix inversion in (55) by substituting (57) in (55).

4.2 Step 2

In order to solve (42) recursively, the Newton method is applied where it is iterated once as new data arrives. Consequently, the first and second order derivative of the cost function J_k in (43) are to be determined w.r.t. α_k, which are denoted J_k' and J_k'', respectively. These are given by

$$J_k' = -2\left(\hat{\sigma}_{\tilde{u}}^k - \hat{\varsigma}_k\right)\hat{\varsigma}_k', \quad (58a)$$

$$J_k'' = \hat{\varsigma}_k', \quad (58b)$$

where $\hat{\varsigma}_k'$ denotes the derivative of $\hat{\varsigma}_k$ w.r.t. α_k and for which an approximation is derived in Appendix C. The recursive update for $\hat{\alpha}_k$ is therefore given by

$$\hat{\alpha}_k = \hat{\alpha}_{k-1} - \left[J_k''\right]^{-1} J_k', \quad (59)$$

whilst

$$\hat{\rho}_y^k = \hat{\alpha}_k N(H_k) + H_k^{\dagger} h_k. \quad (60)$$

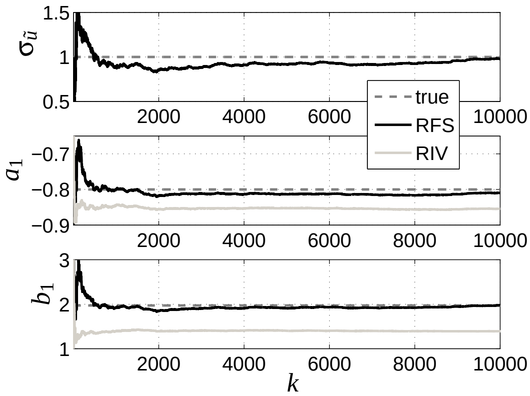

Figure 2: Recursive estimates for θ and $\sigma_{\tilde{u}}$ using the recursive Frisch scheme (RFS) and the biased recursive instrumental-variables (RIV) solution of the uncompensated normal equations.

5 SIMULATION

To compare the results of the recursive Frisch scheme (RFS) with the non-recursive algorithm, the system is chosen similar to that of Example 2 in (Söderström, 2008), i.e. a LTI SISO system with $n_a = n_b = 1$, and characterised, using (12), by

$$\vartheta = \begin{bmatrix}-0.8 & 2 & 1 & 1.96 & 1.37\end{bmatrix}^T. \quad (61)$$

The values for $r_{\tilde{y}}(0)$ and $r_{\tilde{y}}(1)$ arise by generating the output noise by the auto-regressive model

$$\tilde{y}_k = \frac{1}{1-0.7q^{-1}} v_k, \quad (62)$$

where v_k is a zero-mean white process with unity variance. The system is simulated for $10,000$ samples using a zero mean, white and Gaussian distributed input signal of unity variance. The corresponding signal-to-noise ratio for input and output is given by 10.60dB and 39.12dB, respectively.

Choosing $\lambda = 1$, the results for Step 1 are shown in Figure 2. The first subplot shows that the Newton method is able to recursively estimate the input measurement noise variance $\sigma_{\tilde{u}}$. The remaining two subplots compare the RIV solution $\hat{\theta}_k^{\mathrm{IV}}$ of the uncompensated normal equations with the recursively compensated Frisch scheme estimate $\hat{\theta}_k$. As expected, the RIV is biased whilst the the RFS successfully compensates for this.

Figure 3 shows the estimates of ρ_y obtained in Step 2 for both the RFS as well as the off-line case. In contrast to the results obtained in Step 1, the quality of the estimates obtained in Step 2 for $\hat{\rho}_y^k$ is inferior. This is in agreement with the results reported in (Söderström, 2008), where a Monte-Carlo

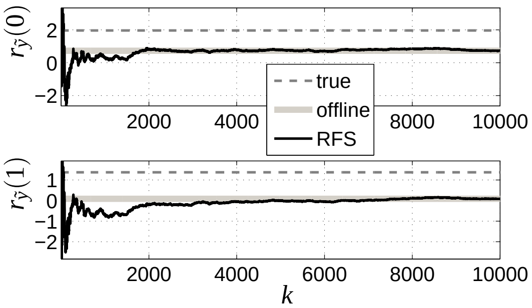

Figure 3: Recursive estimates for $r_{\tilde{y}}(0)$ and $r_{\tilde{y}}(1)$ using the recursive Frisch scheme (RFS).

analysis shows poor performance for $\hat{\rho}^k_y$ in the non-recursive case. The important observation to note here is that the recursively obtained estimates of $r_{\tilde{y}}(0)$ and $r_{\tilde{y}}(0)$ coincide with their off-line counterparts after $k = 10,000$ recursions. It is also observed that the values of $\hat{r}^k_{\tilde{y}}(0)$ (the estimated variance of the output measurement noise) occasionally exhibits a negative sign during the first 500 recursion steps. This could be avoided by projecting the estimates, such that

$$0 < \hat{\Sigma}^k_{\tilde{\varphi}_y} < \hat{\Sigma}^k_{\bar{\varphi}_y} - \hat{\Sigma}^k_{\bar{\varphi}_y\varphi_u}\left[\hat{\Sigma}^k_{\varphi_u}\right]^{-1}\hat{\Sigma}^k_{\varphi_u\bar{\varphi}_y} \tag{63}$$

is satisfied (Söderström, 2008).

6 CONCLUSIONS

The Frisch scheme for the coloured output noise case has been reviewed and a recursive algorithm for its adaptive implementation has been developed. The parameter vector is estimated utilising a recursive bias-compensating instrumental variables approach, where the bias is compensated at each time step. The input measurement noise variance and the output measurement noise auto-covariance elements are obtained via two (distinct) Newton algorithms. A simulation study illustrates the performance of the proposed algorithm.

Further work could concern computational aspects of the algorithm as well as its extension to the bilinear case.

REFERENCES

Beghelli, S., Guidorzi, R. P., and Soverini, U. (1990). The Frisch scheme in dynamic system identification. *Automatica*, 26(1):171–176.

Ding, F., Chen, T., and Qiu, L. (2006). Bias compensation based recursive least-squares identification algorithm for MISO systems. *IEEE Trans. on Circuits and Systems*, 53(5):349–353.

Diversi, R., Guidorzi, R., and Soverini, U. (2003). Algorithms for optimal errors-in-variables filtering. *Systems & Control Letters*, 48:1–13.

Hong, M., Söderström, T., Soverini, U., and Diversi, R. (2007). Comparison of three Frisch methods for errors-in-variables identification. Technical Report 2007-021, Uppsala University.

Linden, J. G., Vinsonneau, B., and Burnham, K. J. (2007). Fast algorithms for recursive Frisch scheme system identification. In *Proc. CD-ROM IAR & ACD Int. Conf.*, Grenoble, France.

Linden, J. G., Vinsonneau, B., and Burnham, K. J. (2008). Gradient-based approaches for recursive Frisch scheme identification. To be published at the 17th IFAC World Congress 2008.

Ljung, L. (1999). *System Identification - Theory for the user*. PTR Prentice Hall Infromation and System Sciences Series. Prentice Hall, New Jersey, 2nd edition.

Ljung, L. and Söderström, T. (1983). *Theory and Practice of Recursive Identification*. M.I.T. Press, Cambridge, MA.

Söderström, T. (2007a). Accuracy analysis of the Frisch scheme for identifying errors-in-variables systems. *Automatica*, 52(6):985–997.

Söderström, T. (2007b). Errors-in-variables methods in system identification. *Automatica*, 43(6):939–958.

Söderström, T. (2008). Extending the Frisch scheme for errors-in-variables identification to correlated output noise. *Int. J. of Adaptive Contr0l and Signal Proc.*, 22(1):55–73.

APPENDIX

A First Order Derivative of V_k

Denoting $(\cdot)'$ the derivative w.r.t. $\hat{\sigma}^k_{\tilde{u}}$ and introducing

$$f_k \triangleq \hat{G}^T_k \hat{\xi}^k_{\bar{\delta}y}, \qquad F_k \triangleq \hat{G}^T_k \hat{G}_k, \tag{64}$$

it holds that

$$f'_k = -\begin{bmatrix} 0 \\ \hat{\xi}^k_{\varphi_u y} \end{bmatrix}, \tag{65a}$$

$$F'_k = \begin{bmatrix} 0 & {\hat{\Sigma}^k_{\varphi_u\varphi_y}}^T \\ \hat{\Sigma}^k_{\varphi_u\varphi_y} & 2\hat{\sigma}^{k-1}_{\tilde{u}} I_{n_b} - 2\hat{\Sigma}^k_{\varphi_u} \end{bmatrix}, \tag{65b}$$

$${F^{-1}_k}' = -F^{-1}_k F'_k F^{-1}_k \tag{65c}$$

and the first order derivative is given by

$$\begin{aligned} V'_k &= -\left(f^T_k F^{-1}_k f_k\right)' \\ &= -\,{f'_k}^T F^{-1}_k f_k - f^T_k\, {F^{-1}_k}'\, f_k - f^T_k F^{-1}_k f'_k. \end{aligned} \tag{66}$$

B Second Order Derivative of V_k

Utilising the product rule, the second order derivative is given by

$$\begin{aligned} V_k'' &= -\, {f_k'}^T\, {F_k^{-1}}'\, f_k - {f_k'}^T F_k^{-1} f_k' \\ &\quad - {f_k'}^T\, {F_k^{-1}}'\, f_k - f_k^T\, {F_k^{-1}}''\, f_k - f_k^T\, {F_k^{-1}}'\, f_k' \\ &\quad - {f_k'}^T F_k^{-1} f_k' - f_k^T\, {F_k^{-1}}'\, f_k' \\ &= -2\, {f_k'}^T\, {F_k^{-1}}'\, f_k - 2\, {f_k'}^T F_k^{-1} f_k' \\ &\quad - f_k^T\, {F_k^{-1}}''\, f_k - 2 f_k^T\, {F_k^{-1}}'\, f_k' \end{aligned} \tag{67}$$

with

$${F_k^{-1}}'' = -\, {F_k^{-1}}' F_k' F_k^{-1} - F_k^{-1} F_k'' F_k^{-1} - F_k^{-1} F_k'\, {F_k^{-1}}', \tag{68a}$$

$$F_k'' = \begin{bmatrix} 0 & 0 \\ 0 & 2I_{n_b} \end{bmatrix}. \tag{68b}$$

C Derivative of $\hat{\varsigma}_k$

The idea is to linearise the Frisch equation (20) using perturbation theory, in order to approximate the derivative of $\hat{\varsigma}_k$ w.r.t. α_k. The derivation here is conceptually similar to that given in Appendix II.B of (Söderström, 2007a), but with the linearisation carried out around $\hat{\vartheta}_{k-1}$ rather than the 'true' parameters.

Assume that at time instance $k-1$, $\hat{\vartheta}_{k-1}$ satisfies the extended compensated normal equations

$$\begin{bmatrix} \hat{\Sigma}^{k-1}_{\bar{\varphi}_y} - \hat{\Sigma}^{k-1}_{\tilde{\varphi}_y} & \hat{\Sigma}^{k-1}_{\bar{\varphi}_y\varphi_u} \\ \hat{\Sigma}^{k-1}_{\varphi_u\bar{\varphi}_y} & \hat{\Sigma}^{k-1}_{\varphi_u} - \hat{\sigma}^{k-1}_{\tilde{u}} I_{n_b} \end{bmatrix} \begin{bmatrix} \bar{\hat{a}}_{k-1} \\ \hat{b}_{k-1} \end{bmatrix} = 0 \tag{69}$$

which are rewritten for ease of notation as

$$\begin{bmatrix} \mathfrak{A} - \mathfrak{B} & \mathfrak{C} \\ \mathfrak{C}^T & \mathfrak{D} - \hat{\sigma}^{k-1}_{\tilde{u}} I \end{bmatrix} \begin{bmatrix} \mathfrak{a} \\ \mathfrak{b} \end{bmatrix} = 0. \tag{70}$$

Similarly, introduce the notation at time instance k as

$$\begin{bmatrix} \mathcal{A} - \mathcal{B} & \mathcal{C} \\ \mathcal{C}^T & \mathcal{D} - \hat{\sigma}^{k}_{\tilde{u}} I \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0. \tag{71}$$

Let $\hat{\sigma}^k_{\tilde{u}}$ denote the estimate obtained via (35). Alternatively, if $\hat{\Sigma}^k_{\tilde{\varphi}_y}$ is known, the input measurement noise could be obtained using (20) and denote this quantity ς^k. Using perturbation theory for eigenvalues yields

$$\begin{aligned} \varsigma^k &= \lambda_{\min}\{B_k(\alpha_k)\} = \lambda_{\min}\{B_{k-1}(\alpha_{k-1}) + \Delta B_k\} \\ &\approx \varsigma^{k-1} + \frac{\mathfrak{b}^T \Delta B_k \mathfrak{b}}{\mathfrak{b}^T \mathfrak{b}}, \end{aligned} \tag{72}$$

where the perturbation is given by (cf. (21))

$$\begin{aligned} \Delta B_k &= B_k(\alpha_k) - B_{k-1}(\alpha_{k-1}) \\ &= \mathcal{D} - \mathcal{C}^T [\mathcal{A} - \mathcal{B}]^{-1} \mathcal{C} - \mathfrak{D} + \mathfrak{C}^T [\mathfrak{A} - \mathfrak{B}]^{-1} \mathfrak{C} \\ &= \mathcal{D} - \mathcal{C}^T \mathcal{F}^{-1} \mathcal{C} - \mathfrak{D} + \mathfrak{C}^T \mathfrak{F}^{-1} \mathfrak{C} \end{aligned} \tag{73}$$

with $\mathcal{F} \triangleq [\mathcal{A} - \mathcal{B}]$ and $\mathfrak{F} \triangleq [\mathfrak{A} - \mathfrak{B}]$. Substituting (73) in (72) yields

$$\begin{aligned} \varsigma^k - \varsigma^{k-1} &\approx \frac{\mathfrak{b}^T}{\mathfrak{b}^T \mathfrak{b}} \left(\mathcal{D} - \mathfrak{D} + \mathfrak{C}^T \mathfrak{F}^{-1} \mathfrak{C} - \mathcal{C}^T \mathcal{F}^{-1} \mathcal{C} \right) \mathfrak{b} \\ &= \frac{\mathfrak{b}^T}{\mathfrak{b}^T \mathfrak{b}} (\mathcal{D} - \mathfrak{D})\, \mathfrak{b} + \frac{\mathfrak{b}^T X \mathfrak{b}}{\mathfrak{b}^T \mathfrak{b}}, \end{aligned} \tag{74}$$

where X can be expressed as

$$\begin{aligned} X &= \mathfrak{C}^T \mathfrak{F}^{-1} \mathfrak{C} - \mathcal{C}^T \mathcal{F}^{-1} \mathcal{C} \\ &\quad + \mathfrak{C}^T \mathcal{F}^{-1} \mathcal{C} - \mathfrak{C}^T \mathcal{F}^{-1} \mathcal{C} \\ &\quad + \mathfrak{C}^T \mathfrak{F}^{-1} \mathcal{C} - \mathfrak{C}^T \mathfrak{F}^{-1} \mathcal{C} \\ &= \left(\mathfrak{C}^T - \mathcal{C}^T \right) \mathcal{F}^{-1} \mathcal{C} + \mathfrak{C}^T \mathfrak{F}^{-1} (\mathfrak{C} - \mathcal{C}) \\ &\quad - \mathfrak{C}^T \mathfrak{F}^{-1} (\mathfrak{F} - \mathcal{F})\, \mathcal{F}^{-1} \mathcal{C} \end{aligned} \tag{75}$$

and by combining (74) and (75), it holds that

$$\begin{aligned} \mathfrak{b}^T \mathfrak{b} \left(\varsigma^k - \varsigma^{k-1} \right) &\approx \mathfrak{b}^T (\mathcal{D} - \mathfrak{D})\, \mathfrak{b} \\ &\quad + \mathfrak{b}^T \left(\mathfrak{C}^T - \mathcal{C}^T \right) \mathcal{F}^{-1} \mathcal{C}\, \mathfrak{b} \\ &\quad + \mathfrak{b}^T \mathfrak{C}^T \mathfrak{F}^{-1} (\mathfrak{C} - \mathcal{C})\, \mathfrak{b} \\ &\quad - \mathfrak{b}^T \mathfrak{C}^T \mathfrak{F}^{-1} (\mathfrak{F} - \mathcal{F})\, \mathcal{F}^{-1} \mathcal{C}\, \mathfrak{b}. \end{aligned} \tag{76}$$

Now, the first row of (70) gives

$$\mathfrak{a} = -\mathfrak{F}^{-1} \mathfrak{C} \mathfrak{b} \tag{77}$$

and by assuming that $\mathcal{F}^{-1} \mathcal{C}\, \mathfrak{b} \approx -\mathfrak{a}$, (76) finally simplifies to

$$\begin{aligned} \mathfrak{b}^T \mathfrak{b} \left(\varsigma^k - \varsigma^{k-1} \right) &\approx \mathfrak{b}^T (\mathcal{D} - \mathfrak{D})\, \mathfrak{b} \\ &\quad - \mathfrak{b}^T \left(\mathfrak{C}^T - \mathcal{C}^T \right) \mathfrak{a} \\ &\quad - \mathfrak{a}^T (\mathfrak{C} - \mathcal{C})\, \mathfrak{b} \\ &\quad - \mathfrak{a}^T (\mathfrak{F} - \mathcal{F})\, \mathfrak{a}, \end{aligned} \tag{78}$$

where $\mathcal{F}$ is the only element depending on α_k. Therefore,

$$\frac{d\varsigma^k}{d\alpha_k} \approx \frac{d}{d\alpha_k} \left(\frac{\mathfrak{a}^T (\mathcal{A} - \mathcal{B})\, \mathfrak{a}}{\mathfrak{b}^T \mathfrak{b}} \right) = -\frac{\mathfrak{a}^T \frac{d\mathcal{B}}{d\alpha_k} \mathfrak{a}}{\mathfrak{b}^T \mathfrak{b}} \tag{79}$$

or equivalently

$$\frac{d}{d\alpha_k} \lambda_{\min}\{B_k(\alpha_k)\} \approx -\frac{\bar{\hat{a}}^T_{k-1}}{\hat{b}^T_{k-1} \hat{b}_{k-1}} \frac{d}{d\alpha_k} \hat{\Sigma}^k_{\tilde{\varphi}_y} \bar{\hat{a}}_{k-1}. \tag{80}$$

Since $\Sigma^k_{\tilde{\varphi}_y}$ consists of the quantities $\hat{r}^k_{\tilde{y}}(0), \ldots, \hat{r}^k_{\tilde{y}}(n_a)$, it remains to determine

$$\frac{d}{d\alpha_k} \hat{\rho}^k_y = \left[\tfrac{d}{d\alpha_k} \hat{r}^k_{\tilde{y}}(0) \quad \cdots \quad \tfrac{d}{d\alpha_k} \hat{r}^k_{\tilde{y}}(n_a) \right]^T \tag{81}$$

which, due to (40), is given by

$$\frac{d}{d\alpha_k} \hat{\rho}^k_y = N(H_k). \tag{82}$$