

COMPUTATIONAL ASPECTS FOR RECURSIVE FRISCH SCHEME SYSTEM IDENTIFICATION

Jens G. Linden, Tomasz Larkowski and Keith J. Burnham

Control Theory and Applications Centre, Coventry University, Priory Street, CV1-5FB Coventry, U.K.

Keywords: System identification, Errors-in-variables, Recursive identification.

Abstract: The implementation of a recursive algorithm for the estimation of parameters of a linear single-input single-output errors-in-variables system is re-considered. The objective is to reduce the computational complexity in order to reduce the computation time per recursion, which, in turn, will allow a wider applicability of the recursive algorithm. The technique of stationary iterative methods for least squares is utilised, in order to reduce the complexity from cubic to quadratic order with respect to the model parameters to be estimated. A numerical simulation underpins the theoretically obtained results.

1 INTRODUCTION

In the case of linear time-invariant (LTI) errors-in-variables (EIV) models not only the output signals of the system, but also the input signals 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; Söderström, 2007a), which yields estimates of the model parameters as well as the measurement noise variances. Recently, recursive Frisch scheme algorithms have been developed in a series of papers by the authors (Linden et al., 2008b; Linden et al., 2007; Linden et al., 2008a). This paper considers a fast implementation of the algorithm presented in (Linden et al., 2008a), which reduces the computational complexity from cubic to quadratic order with respect to the model parameters to be estimated, hence allowing a wider range of applicability of the proposed algorithm.

The paper is outlined as follows. The problem of EIV system identification is formulated in Section 2, where the required notation is also introduced. The Frisch scheme, being one particular EIV system identification approach is reviewed in Section 3, where non-recursive and recursive implementations are discussed. Section 4 develops the novel algorithm which reduces the computational complexity from cubic to

quadratic order, whilst Section 5 presents numerical examples. Section 6 contains concluding remarks as well as direction for further work.

2 PROBLEM STATEMENT

A discrete-time, LTI single-input single-output (SISO) EIV system is considered, which is defined by

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

where i is an integer valued time index and

$$A(q^{-1}) \triangleq 1 + a_1q^{-1} + \dots + a_nq^{-n}, \quad (2a)$$

$$B(q^{-1}) \triangleq b_1q^{-1} + \dots + b_nq^{-n} \quad (2b)$$

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

$$u_i = u_{0i} + \tilde{u}_i, \quad (3a)$$

$$y_i = y_{0i} + \tilde{y}_i \quad (3b)$$

are available, where \tilde{u}_i and \tilde{y}_i denote the input and output measurement noise, respectively. Such an EIV 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.

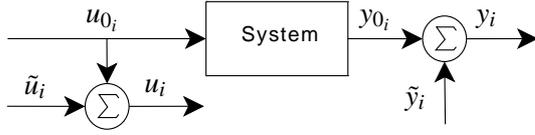


Figure 1: Errors-in-variables setup.

- 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_{0i} is a zero-mean ergodic process and is persistently exciting of sufficiently high order.
- A5** The sequences \tilde{u}_i and \tilde{y}_i are zero-mean, ergodic, white noises with unknown variances $\sigma_{\tilde{u}}$ and $\sigma_{\tilde{y}}$, respectively.
- A6** The sequences \tilde{u}_i and \tilde{y}_i are mutually uncorrelated and uncorrelated with u_{0i} and y_{0i} , respectively.

A notational convention within this paper is that covariance matrices of two column vectors v_k and w_k are denoted

$$\Sigma_{vw} \triangleq E[v_i w_i^T], \quad \Sigma_v \triangleq E[v_i v_i^T], \quad (4)$$

where $E[\cdot]$ denotes the expected value operator. In addition, vectors consisting of covariance elements are denoted

$$\xi_{vc} \triangleq E[v_i c_i] \quad (5)$$

with c_i being a scalar. The parameters are elements within a vector, which is defined by

$$\theta \triangleq [a^T \ b^T]^T = [a_1 \ \dots \ a_{n_a} \ b_1 \ \dots \ b_{n_b}]^T, \quad (6a)$$

$$\bar{\theta} \triangleq [\bar{a}^T \ b^T]^T = [1 \ \theta^T]^T. \quad (6b)$$

This allows an alternative formulation of (1)-(3) given by

$$\bar{\varphi}_{0i}^T \bar{\theta} = 0, \quad (7a)$$

$$\bar{\varphi}_i = \bar{\varphi}_{0i} + \tilde{\varphi}_i, \quad (7b)$$

where the regression vector is defined by

$$\begin{aligned} \varphi_i &\triangleq [\varphi_{y_i}^T \ \varphi_{u_i}^T]^T \\ &\triangleq [-y_{i-1} \ \dots \ -y_{i-n_a} \ u_{i-1} \ \dots \ u_{i-n_b}]^T, \end{aligned} \quad (8)$$

$$\bar{\varphi}_i \triangleq [\bar{\varphi}_{y_i}^T \ \varphi_{u_i}^T]^T \triangleq [-y_i \ \varphi_i^T]^T. \quad (9)$$

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

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

$$\begin{aligned} \vartheta &\triangleq [\theta^T \ \sigma^T]^T \\ &= [a_1 \ \dots \ a_{n_a} \ b_1 \ \dots \ b_{n_b} \ \sigma_{\tilde{y}} \ \sigma_{\tilde{u}}]^T. \end{aligned} \quad (10)$$

Throughout this paper, the convention is made that estimated quantities are marked by a $\hat{\cdot}$ 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 REVIEW OF THE FRISCH SCHEME

One possibility to address Problem 1 is the so called Frisch scheme (Beghelli et al., 1990; Söderström, 2006), which defines a set of admissible solutions for the estimates of the input and output measurement noise variances as well as the parameter vector. In order to single out one particular solution, different model selection criteria have been proposed within the literature, leading to different variants of the Frisch scheme (Hong et al., 2007). The criterion which is considered here is the Yule-Walker (YW) model selection criterion described in (Diversi et al., 2006) and the corresponding Frisch scheme algorithm is denoted Frisch-YW.

3.1 Non-recursive Frisch Scheme

The estimates of the (non-recursive) Frisch-YW are characterised by the input measurement noise variance $\sigma_{\tilde{u}}$, whose estimate, denoted $\hat{\sigma}_{\tilde{u}}^k$, is obtained by the nonlinear set of equations (Beghelli et al., 1990; Diversi et al., 2006)

$$\theta = \left(\hat{\Sigma}_{\varphi}^k - \Sigma_{\bar{\varphi}}(\sigma) \right)^{-1} \hat{\xi}_{\varphi y}^k, \quad (11a)$$

$$\sigma_{\tilde{y}} = \lambda_{\min}(\hat{A}_k), \quad (11b)$$

$$\hat{\sigma}_{\tilde{u}}^k = \arg \min_{\sigma_{\tilde{u}}} V_k, \quad (11c)$$

where

$$\Sigma_{\bar{\varphi}}(\sigma) = \begin{bmatrix} \sigma_{\tilde{y}} I_{n_a} & 0 \\ 0 & \sigma_{\tilde{u}} I_{n_b} \end{bmatrix}, \quad (12a)$$

$$\hat{A}_k \triangleq \hat{\Sigma}_{\bar{\varphi} y}^k - \hat{\Sigma}_{\bar{\varphi} y}^k \left[\hat{\Sigma}_{\varphi u}^k - \sigma_{\tilde{u}} I_{n_b} \right]^{-1} \hat{\Sigma}_{\varphi u}^k, \quad (12b)$$

$$V_k = \frac{1}{2} \|r_k(\theta)\|_2^2 = \frac{1}{2} \|\hat{\Sigma}_{\varphi}^k \theta - \hat{\xi}_{\varphi y}^k\|_2^2, \quad (12c)$$

and λ_{\min} denotes the minimum eigenvalue operator. The instrument vector, denoted ζ_k , is defined by

$$\zeta_k = [u_{k-n_b-1} \ \cdots \ u_{k-n_b-n_\zeta}]^T \quad (13)$$

where $n_\zeta \geq n_a + n_b + 1$ denotes the number of instruments which is user specified. The quantity $r_k(\theta)$ denotes the nonlinear least squares residual corresponding to a certain θ . Once the $\sigma_{\bar{u}}$ has been estimated, this value is substituted in (11b) and (11a), in order to obtain $\hat{\sigma}_{\bar{y}}^k$ and $\hat{\theta}_k$, respectively. Note that (11a)-(11b) form the core of the Frisch scheme, whilst (11c) is the YW model selection criterion with the corresponding YW cost function, denoted V_k . Also note that $\hat{\theta}_k$ depends on $\hat{\sigma}_{\bar{u}}^k$ and $\hat{\sigma}_{\bar{y}}^k$, where the latter is also a function of $\hat{\sigma}_{\bar{u}}^k$ defined by (11b), hence, V_k is nonlinear in $\hat{\sigma}_{\bar{u}}^k$.

3.2 Recursive Frisch Scheme

Update of θ . The recursive Frisch scheme presented in (Linden et al., 2008b) is based on the iterative/recursive bias compensating least squares (RBCLS) approach (Sagara and Wada, 1977; Söderström, 2007b; Zheng and Feng, 1989). Assuming the noise covariances have already been obtained, the parameter vector is computed via

$$\hat{\theta}_k = \hat{\theta}_k^{\text{LS}} + P_k \Sigma_{\bar{\phi}}(\hat{\sigma}_k) \hat{\theta}_{k-1}, \quad (14)$$

where $\hat{\theta}_k^{\text{LS}}$ and P_k are the least squares (LS) estimate and corresponding (scaled) error covariance matrix, respectively. Both quantities are computed via the well known recursive least squares (RLS) algorithm (Ljung, 1999)

$$\hat{\theta}_k^{\text{LS}} = \hat{\theta}_{k-1}^{\text{LS}} + L_k \left(y_k - \Phi_k^T \hat{\theta}_{k-1}^{\text{LS}} \right), \quad (15a)$$

$$L_k = \frac{P_{k-1} \Phi_k}{\Phi_k^T P_{k-1} \Phi_k + \frac{1-\gamma_k}{\gamma_k}}, \quad (15b)$$

$$P_k = \frac{1}{1-\gamma_k} \left(P_{k-1} - \frac{P_{k-1} \Phi_k \Phi_k^T P_{k-1}}{\Phi_k^T P_{k-1} \Phi_k + \frac{1-\gamma_k}{\gamma_k}} \right). \quad (15c)$$

The quantity P_k is scaled such that $P_k = [\hat{\Sigma}_{\bar{\phi}}^k]^{-1}$, whilst the scaling factor γ_k is chosen to be $1 - \lambda$ in the case of exponential forgetting, with λ being the forgetting factor.

Update of $\sigma_{\bar{y}}$. For the determination of $\sigma_{\bar{y}}$, a conjugate gradient subspace tracking algorithm (cf. (Feng and Owen, 1996)) has been utilised in (Linden et al., 2008b). In order to reduce the computational complexity from cubic to quadratic complexity¹, an ap-

¹The complexity with respect to the number of parameters to be estimated is considered. Note that the conjugate gradient algorithm in (Linden et al., 2008b) is of cubic order due to the inverse computation within the Schur complement (12b) and not due to the subspace tracking algorithm.

proximate algorithm based on the Rayleigh quotient has been proposed in (Linden et al., 2007). This leads to

$$\hat{\theta}_{k-\frac{1}{2}} = \hat{\theta}_k^{\text{LS}} + P_k \begin{bmatrix} \hat{\sigma}_{\bar{y}}^{k-1} I_{n_a} & 0 \\ 0 & \hat{\sigma}_{\bar{u}}^k I_{n_b} \end{bmatrix} \hat{\theta}_{k-1}, \quad (16a)$$

$$\hat{\sigma}_{\bar{y}}^k = \frac{\hat{a}_{k-\frac{1}{2}}^T}{\hat{a}_{k-\frac{1}{2}}^T \hat{a}_{k-\frac{1}{2}}} \left(\hat{\Sigma}_{\bar{\phi}_y}^k \hat{a}_{k-\frac{1}{2}} + \hat{\Sigma}_{\bar{\phi}_y \bar{\phi}_u}^k \hat{b}_{k-\frac{1}{2}} \right), \quad (16b)$$

where $\hat{\theta}_{k-\frac{1}{2}}$ denotes an intermediate parameter estimate, which makes use of the most recent estimate of $\sigma_{\bar{u}}$ (which is determined before the update of $\hat{\sigma}_{\bar{y}}^k$ takes place).

Update of $\sigma_{\bar{u}}$. For the update of the input measurement noise variance $\sigma_{\bar{u}}$, a steepest-gradient algorithm has been proposed in (Linden et al., 2008b). Recently, an alternative approach for the (approximate) minimisation of (12c) has been suggested in (Linden et al., 2008a). There, the cost function V_k is modified by replacing θ in (12c), which is nonlinear in $\sigma_{\bar{u}}$ due to (11a) and (11b), by the approximation $L_{\theta}(\hat{\vartheta}_{k-1})$, which is obtained by making use of linearisations of (11a) and (11b) around the latest estimates $\hat{\vartheta}_{k-1}$. These linearisations have been developed in (Söderström, 2007a) and are given by

$$\hat{\theta}_k \approx L_{\theta}(\hat{\vartheta}_{k-1}) \triangleq \hat{\theta}_{k-1} + \left(\hat{\Sigma}_{\bar{\phi}}^k - \Sigma_{\bar{\phi}}(\hat{\sigma}_{k-1}) \right)^{-1} \quad (17a)$$

$$\times \left(\hat{\xi}_{\bar{\phi}_y}^k - \hat{\Sigma}_{\bar{\phi}}^k \hat{\theta}_{k-1} + \begin{bmatrix} \hat{\sigma}_{\bar{y}}^k a_{k-1} \\ \hat{\sigma}_{\bar{u}}^k b_{k-1} \end{bmatrix} \right),$$

$$\hat{\sigma}_{\bar{y}}^k \approx L_{\sigma_{\bar{y}}}(\hat{\vartheta}_{k-1}) \triangleq \hat{\sigma}_{\bar{y}}^{k-1} + \frac{\hat{b}_{k-1}^T \hat{b}_{k-1}}{\hat{a}_{k-1}^T \hat{a}_{k-1}} \left(\hat{\sigma}_{\bar{u}}^{k-1} - \hat{\sigma}_{\bar{u}}^k \right). \quad (17b)$$

For a convenient notation, introduce

$$\mathbf{v}(\hat{\vartheta}_{k-1}) \triangleq \hat{\xi}_{\bar{\phi}_y}^k - \hat{\Sigma}_{\bar{\phi}}^k \hat{\theta}_{k-1} \quad (18a)$$

$$+ \begin{bmatrix} \hat{\sigma}_{\bar{y}}^{k-1} + \frac{\hat{b}_{k-1}^T \hat{b}_{k-1}}{\hat{a}_{k-1}^T \hat{a}_{k-1}} \hat{\sigma}_{\bar{u}}^{k-1} \\ 0 \end{bmatrix} \begin{bmatrix} \hat{a}_{k-1} \\ 0 \end{bmatrix},$$

$$\mathbf{\kappa}(\hat{\vartheta}_{k-1}) \triangleq \begin{bmatrix} -\frac{\hat{b}_{k-1}^T \hat{b}_{k-1}}{\hat{a}_{k-1}^T \hat{a}_{k-1}} \hat{a}_{k-1} \\ \hat{b}_{k-1} \end{bmatrix}, \quad (18b)$$

$$\Sigma_{\bar{\phi}_0}(\hat{\sigma}_{k-1}) \triangleq \hat{\Sigma}_{\bar{\phi}}^k - \Sigma_{\bar{\phi}}(\hat{\sigma}_{k-1}). \quad (18c)$$

Using this notation, the quantity $\sigma_{\bar{y}}$ given by (17b) can be eliminated in (17a) yielding a linear expression for θ which only depends on $\hat{\sigma}_{\bar{u}}^k$

$$L_{\theta}(\hat{\vartheta}_{k-1}) = \hat{\theta}_{k-1} + \Sigma_{\bar{\phi}_0}^{-1}(\hat{\sigma}_{k-1}) \mathbf{v}(\hat{\vartheta}_{k-1}) + \Sigma_{\bar{\phi}_0}^{-1}(\hat{\sigma}_{k-1}) \mathbf{\kappa}(\hat{\vartheta}_{k-1}) \hat{\sigma}_{\bar{u}}^k. \quad (19)$$

Substituting θ in (12c) with $L_\theta(\hat{\vartheta}_{k-1})$ allows the approximate cost function to be defined

$$V_k^{\text{lin}} \triangleq \frac{1}{2} \|r_k(L_\theta(\hat{\vartheta}_{k-1}))\|_2^2, \quad (20)$$

which can be minimised analytically at each time instance k . Differentiating with respect to θ and setting equal to zero gives

$$\hat{\sigma}_u^k = \frac{J^T(\hat{\sigma}_u^k) \left(\hat{\Sigma}_{\zeta\varphi}^k [\hat{\theta}_{k-1} + \Sigma_{\varphi_0}^{-1}(\hat{\sigma}_{k-1})\mathbf{1}(\hat{\vartheta}_{k-1})] - \hat{\xi}_{\zeta y}^k \right)}{-J^T(\hat{\sigma}_u^k) \hat{\Sigma}_{\zeta\varphi}^k \Sigma_{\varphi_0}^{-1}(\hat{\sigma}_{k-1}) \kappa(\hat{\vartheta}_{k-1})}, \quad (21)$$

where the Jacobian $J(\hat{\sigma}_u^k)$ is given by

$$J(\hat{\sigma}_u^k) = \hat{\Sigma}_{\zeta\varphi}^k \frac{dL_\theta}{d\hat{\sigma}_u^k}, \quad (22)$$

whilst the total derivative of L_θ is obtained from (19) as

$$\frac{dL_\theta}{d\hat{\sigma}_u^k} = \Sigma_{\varphi_0}^{-1}(\hat{\sigma}_{k-1}) \kappa(\hat{\vartheta}_{k-1}). \quad (23)$$

The resulting algorithm, which consists of (14)-(16) and (21) is referred to as recursive Frisch scheme (RFS) within the subsequent development.

4 FAST RECURSIVE FRISCH SCHEME ALGORITHM

It is observed that for the computation of the input measurement noise variance in (21), the matrix Σ_{φ_0} is required to be inverted. A matrix inversion is generally of cubic complexity with respect to its dimension, which is here equal to $n_a + n_b$, the number of parameters to be estimated. Indeed, this matrix inversion is the bottleneck within the RFS approach described in Section 3.2, since the remaining operations are only of quadratic complexity with respect to the model parameters. Since the intended use for such a recursive scheme lies in an online computation of the system parameters, it would certainly be attractive to reduce the computational burden of the input measurement noise computation to quadratic order. This would allow a wider application of the algorithm for cases where less computational power is available. The development of such an algorithm is the topic of this section.

4.1 First Bottleneck

The first bottleneck is due the computation of the inverse within the total derivative of L_θ , which has been

given in Section 3.2 by (23). However, by making use of stationary iterative methods for solving LS problems (Björck, 1996, Chapter 7), Equation (23) can be re-expressed as

$$\hat{\Sigma}_\varphi^k \frac{dL_\theta}{d\hat{\sigma}_u^k} - \Sigma_{\bar{\varphi}}(\hat{\sigma}_{k-1}) \frac{dL_\theta}{d\hat{\sigma}_u^k} = \kappa(\hat{\vartheta}_{k-1}), \quad (24)$$

where the matrix splitting is given naturally by (18c). An iterative/recursive way to compute $dL_\theta/d\hat{\sigma}_u^k$ could therefore be given by

$$L_\theta^{k'} \triangleq P_k \left[\kappa(\hat{\vartheta}_{k-1}) + \Sigma_{\bar{\varphi}}(\hat{\sigma}_{k-1}) L_\theta^{k-1'} \right], \quad (25)$$

where $L_\theta^{k'}$ denotes the recursively computed derivative and $P_k = [\hat{\Sigma}_\varphi^k]^{-1}$ is given by the matrix inversion lemma of the RLS algorithm, which is already computed for the determination of $\hat{\theta}_k$.

4.2 Second Bottleneck

The second bottleneck is due to the matrix inverse within the computation of (19), therefore, an (approximate) recursive expression for $L_\theta(\hat{\vartheta}_{k-1})$ is required, which is of quadratic complexity only. Firstly, introduce the notation $L_\theta(\hat{\vartheta}_{k-1}) \triangleq L_\theta^k$, where the index k is chosen to reflect the fact that L_θ^k corresponds to the linearisation at time instance k (although it depends on the estimate $\hat{\vartheta}_{k-1}$ with time index $k-1$). Secondly, assume that all past $\hat{\theta}_k$ have been computed using the expression (19), which means that $\hat{\theta}_{k-1}$ can be replaced with L_θ^{k-1} in (19). Thirdly, from (18a) and (18b) it holds

$$\begin{aligned} \mathbf{1}(\hat{\vartheta}_{k-1}) + \kappa(\hat{\vartheta}_{k-1}) \hat{\sigma}_u^k &= \hat{\xi}_{\varphi y}^k - \hat{\Sigma}_\varphi^k \hat{\theta}_{k-1} + \begin{bmatrix} \hat{a}_{k-1} \\ 0 \end{bmatrix} \hat{\sigma}_y^{k-1} \\ &+ \begin{bmatrix} \hat{b}_{k-1}^T \hat{b}_{k-1} \\ \hat{a}_{k-1}^T \hat{a}_{k-1} \\ 0 \end{bmatrix} \hat{\sigma}_u^{k-1} + \begin{bmatrix} -\hat{b}_{k-1}^T \hat{b}_{k-1} \\ \hat{a}_{k-1}^T \hat{a}_{k-1} \\ 0 \end{bmatrix} \hat{\sigma}_u^k \\ &+ \begin{bmatrix} 0 \\ \hat{b}_{k-1} \end{bmatrix} \hat{\sigma}_u^k, \end{aligned} \quad (26)$$

and by assuming that $\hat{\sigma}_u^k \approx \hat{\sigma}_u^{k-1}$, $\hat{\sigma}_y^k \approx \hat{\sigma}_y^{k-1}$ and using $\hat{\theta}_{k-1} = L_\theta^{k-1}$, one obtains

$$\begin{aligned} \mathbf{1}(\hat{\vartheta}_{k-1}) + \kappa(\hat{\vartheta}_{k-1}) \hat{\sigma}_u^k &\approx \hat{\xi}_{\varphi y}^k - \hat{\Sigma}_\varphi^k L_\theta^{k-1} \\ &+ \begin{bmatrix} \hat{\sigma}_y^k I_{n_a} & 0 \\ 0 & \hat{\sigma}_u^k I_{n_b} \end{bmatrix} L_\theta^{k-1}. \end{aligned} \quad (27)$$

Finally, by substituting (18c) and (27) into (19), it holds

$$\begin{aligned} \left[\hat{\Sigma}_\varphi^k - \Sigma_{\bar{\varphi}}(\hat{\sigma}_{k-1}) \right] L_\theta^k &= \left[\hat{\Sigma}_\varphi^k - \Sigma_{\bar{\varphi}}(\hat{\sigma}_{k-1}) \right] L_\theta^{k-1} \\ &+ \hat{\xi}_{\varphi y}^k - \hat{\Sigma}_\varphi^k L_\theta^{k-1} + \Sigma_{\bar{\varphi}}(\hat{\sigma}_k) L_\theta^{k-1}, \end{aligned} \quad (28)$$

which simplifies to

$$\begin{aligned} \left[\hat{\Sigma}_{\varphi}^k - \Sigma_{\hat{\varphi}}(\hat{\sigma}_{k-1}) \right] L_{\hat{\theta}}^k &= -\Sigma_{\hat{\varphi}}(\hat{\sigma}_{k-1}) L_{\hat{\theta}}^{k-1} + \hat{\xi}_{\varphi}^k \\ &+ \Sigma_{\hat{\varphi}}(\hat{\sigma}_k) L_{\hat{\theta}}^{k-1}. \end{aligned} \quad (29)$$

Thus, by using $L_{\hat{\theta}}^{k-1} \approx L_{\hat{\theta}}^k$, a recursive computation of the linearised θ -equation (17a) is given by

$$L_{\hat{\theta}}^k \approx [\hat{\Sigma}_{\varphi}^k]^{-1} \hat{\xi}_{\varphi}^k + [\hat{\Sigma}_{\varphi}^k]^{-1} \Sigma_{\hat{\varphi}}(\hat{\sigma}_k) L_{\hat{\theta}}^{k-1}, \quad (30)$$

which interestingly is, indeed, the RBCLS algorithm given in (14) (i.e. simply replace $\hat{\theta}_k$ with $L_{\hat{\theta}}^k$ in (14)). Since the recursive computation of $L_{\hat{\theta}}^k$ is identical to the RBCLS computation of $\hat{\theta}_k$, the latter, more familiar, notation can be utilised. Substituting the linearised λ_{\min} -equation (17b), the RBCLS equation becomes

$$\begin{aligned} \hat{\theta}_k &= \hat{\theta}_k^{\text{LS}} + P_k \begin{bmatrix} \hat{\sigma}_y^k \hat{a}_{k-1} \\ \hat{\sigma}_u^k \hat{b}_{k-1} \end{bmatrix} \\ \Leftrightarrow \hat{\theta}_k &= \hat{\theta}_k^{\text{LS}} + P_k \begin{bmatrix} 0 \\ \hat{b}_{k-1} \end{bmatrix} \hat{\sigma}_u^k + P_k \begin{bmatrix} \hat{a}_{k-1} \\ 0 \end{bmatrix} \\ &\times \left[\hat{\sigma}_y^{k-1} + \frac{\hat{b}_{k-1}^T \hat{b}_{k-1}}{\hat{a}_{k-1}^T \hat{a}_{k-1}} \hat{\sigma}_u^{k-1} - \frac{\hat{b}_{k-1}^T \hat{b}_{k-1}}{\hat{a}_{k-1}^T \hat{a}_{k-1}} \hat{\sigma}_u^k \right], \end{aligned} \quad (31)$$

which simplifies to

$$\hat{\theta}_k = P_k \bar{\iota}(\hat{\vartheta}_{k-1}) + P_k \kappa(\hat{\vartheta}_{k-1}) \hat{\sigma}_u^k, \quad (32)$$

where $\kappa(\hat{\vartheta}_{k-1})$ is defined by (18b) and

$$\bar{\iota}(\hat{\vartheta}_{k-1}) \triangleq \hat{\xi}_{\varphi}^k + \begin{bmatrix} \hat{a}_{k-1} \\ 0 \end{bmatrix} \left[\hat{\sigma}_y^{k-1} + \frac{\hat{b}_{k-1}^T \hat{b}_{k-1}}{\hat{a}_{k-1}^T \hat{a}_{k-1}} \hat{\sigma}_u^{k-1} \right]. \quad (33)$$

4.3 Fast Update of $\hat{\sigma}_u^k$

Using the previous results, a fast implementation for the update of $\hat{\sigma}_u^k$ can be realised. With the Jacobian at time instance k being given by (cf. (22))

$$J_k \triangleq \hat{\Sigma}_{\zeta\varphi}^k L_{\hat{\theta}}^k, \quad (34)$$

it is therefore possible to compute $\hat{\sigma}_u^k$ as

$$0 = J_k^T \left[\hat{\Sigma}_{\zeta\varphi}^k \hat{\theta}_k - \hat{\xi}_{\zeta\varphi}^k \right] \quad (35)$$

and by substituting $\hat{\theta}_k$ given in (32), the fast update for $\hat{\sigma}_u^k$ is finally given by

$$\hat{\sigma}_u^k = \frac{J_k^T \left[\hat{\xi}_{\zeta\varphi}^k - \hat{\Sigma}_{\zeta\varphi}^k P_k \bar{\iota}(\hat{\vartheta}_{k-1}) \right]}{J_k^T \hat{\Sigma}_{\zeta\varphi}^k P_k \kappa(\hat{\vartheta}_{k-1})}. \quad (36)$$

Note that only matrix vector multiplications are required for the fast computation of $\hat{\sigma}_u^k$, hence the computational effort is reduced towards quadratic complexity. The fast RFS algorithm, which consists of (14)-(16) and (36) is referred to as FRFS within the subsequent development.

5 NUMERICAL EXAMPLES

It is of interest to compare the RFS estimates with those obtained by the FRFS and also to compare the computation time of both algorithms.

5.1 Estimation of $\sigma_{\bar{u}}$

A LTI SISO system with $n_a = n_b = 2$ and given by

$$\theta = [-1.5 \quad 0.7 \quad 1 \quad 0.5]^T \quad (37a)$$

$$\sigma = [2.1 \quad 0.1]^T \quad (37b)$$

is simulated for 1000 samples using a zero mean, white and Gaussian distributed input signal of unity variance. The RFS and FRFS algorithms are applied to estimate ϑ using $n_{\zeta} = n_a + n_b + 1$, whilst $\lambda = 1$ is chosen (i.e. no forgetting). The estimates of $\sigma_{\bar{u}}$ and $\sigma_{\bar{y}}$ are projected into the intervals $[0, \sigma_{\bar{u}}^{\max}]$ and $[0, \sigma_{\bar{y}}^{\max}]$, where the maximal admissible values for the input and output measurement noise variances are chosen to be $\sigma_{\bar{u}}^{\max} = 2\sigma_{\bar{u}} = 0.2$ and $\sigma_{\bar{y}}^{\max} = 2\sigma_{\bar{y}} = 4.2$, respectively. The estimates of $\sigma_{\bar{u}}$ are compared in Figure 2. Here it is observed that the projection facility (which

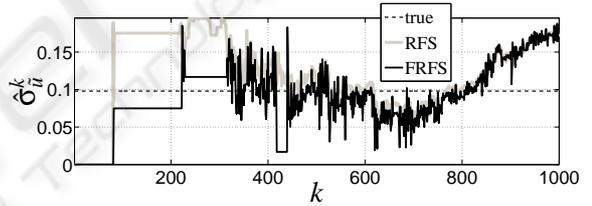


Figure 2: Estimates of $\sigma_{\bar{u}}$ for using the RFS, and FRFS.

sets $\hat{\sigma}_u^k = \hat{\sigma}_u^{k-1}$ if the estimate is not within the specified interval) seems to be more often active for the fast algorithm (see around $k = 420$). After approximately 500 recursions, however, the FRFS estimate is barely distinguishable from the RFS, although the FRFS solution seems to be slightly more erratic. Hence, at least in the example considered here, the FRFS appears to be able to approximate the estimate of $\sigma_{\bar{u}}$ obtained by the more computationally demanding RFS algorithm.

5.2 Comparison of Computation Time

Naturally, it is of major interest to compare the computation time per recursion of the FRFS algorithm with that of the RFS scheme. Therefore, the algorithms are applied to systems with an incrementally increasing model order $m = n_a = n_b = 1, \dots, 30$ and the computation time per single recursion is recorded for each identification task. The results are presented

in Figure 3, which clearly shows the relative reduction of computational complexity for the FRFS approach. For a model order of $m = 30$, the RFS requires around

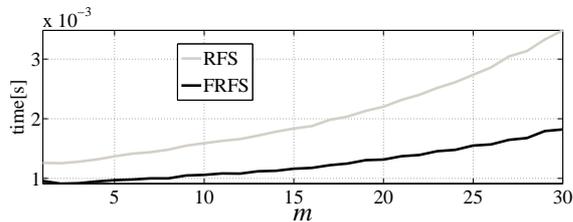


Figure 3: Computation time per single recursion with increasing model order m .

3.5ms, whilst the FRFS requires less than 2.0ms. The fact that the slope of the curve corresponding to the FRFS algorithm is lower than that of the RFS approach illustrates that the computational complexity is reduced from cubic to quadratic order; this underpins the theoretical results obtained in this paper.

6 CONCLUSIONS AND FURTHER WORK

The Frisch scheme for the identification of linear time-invariant single-input single-output errors-in-variables systems has been reviewed. The well-known non-recursive case as well as a recently developed recursive algorithm has been discussed. Since the latter is of cubic computational complexity with respect to the number of parameters to be estimated, several approximations have been introduced, in order to reduce the complexity from cubic to quadratic order. This theoretical result is in agreement with the measured computation time which has been obtained for a numerical simulation. This simulation has also shown that the fast algorithm is able to approximate the solution of the computationally more demanding algorithm satisfactorily.

Further work could concern the convergence properties of the recursive algorithm.

REFERENCES

- Beghelli, S., Guidorzi, R. P., and Soverini, U. (1990). The Frisch scheme in dynamic system identification. *Automatica*, 26(1):171–176.
- Björck, Å. (1996). *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia.
- Diversi, R., Guidorzi, R., and Soverini, U. (2006). Yule-Walker equations in the Frisch scheme solution of errors-in-variables identification problems. In *Proc.*

17th Int. Symp. on Math. Theory of Networks and Systems, pages 391–395.

- Feng, Y. T. and Owen, D. R. J. (1996). Conjugate gradient methods for solving the smallest eigenpair of large symmetric eigenvalue problems. *Int. J. for Numerical Methods in Engineering*, 39:2209–2229.
- 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, Uppsala, Sweden.
- Linden, J. G., Larkowski, T., and Burnham, K. J. (2008a). An improved recursive Frisch scheme identification algorithm. In *Proc. 19th Int. Conf. on Systems Engineering*, pages 65–70, Las Vegas, USA.
- Linden, J. G., Vinsonneau, B., and Burnham, K. J. (2007). Fast algorithms for recursive Frisch scheme system identification. In *Proc. CD-ROM 22nd IAR & ACD Workshop*, Grenoble, France.
- Linden, J. G., Vinsonneau, B., and Burnham, K. J. (2008b). Gradient-based approaches for recursive Frisch scheme identification. In *Preprints of the 17th IFAC World Congress*, pages 1390–1395, Seoul, Korea.
- Ljung, L. (1999). *System Identification - Theory for the user*. PTR Prentice Hall Information and System Sciences Series. Prentice Hall, New Jersey, 2nd edition.
- Sagara, S. and Wada, K. (1977). On-line modified least-squares parameter estimation of linear discrete dynamic systems. *Int. J. Control*, 25(3):329–343.
- Söderström, T. (2006). Statistical analysis of the Frisch scheme for identifying errors-in-variables systems. Technical report 2006-002, Uppsala University, Department of Information Technology, Uppsala, Sweden.
- Söderström, T. (2007a). Accuracy analysis of the Frisch scheme for identifying errors-in-variables systems. *IEEE Trans. Autom. Contr.*, 52(6):985–997.
- Söderström, T. (2007b). Errors-in-variables methods in system identification. *Automatica*, 43(6):939–958.
- Zheng, W. X. and Feng, C. B. (1989). Unbiased parameter estimation of linear systems in the presence of input and output noise. *Int. J. of Adaptive Control and Signal Proc.*, 3:231–251.