Novel algorithms based on conjunction of the Frisch scheme and extended compensated least squares

Tomasz Larkowski * **Jens G. Linden** * **Benoit Vinsonneau** * **Keith J. Burnham** *

* *Control Theory and Applications Centre, Coventry, UK (Email: larkowst@coventry.ac.uk)*

Abstract: The paper presents a general framework for the Frisch scheme and the extended compensated least squares technique within which two new algorithms for the identification of single-input single-output linear time-invariant errors-in-variables models are proposed. The first algorithm is essentially the Frisch scheme using a novel model selection criterion. The second method is a modification of the extended compensated least squares technique, which utilizes not only the set of overdetermined normal equations, but also the Frisch equation to solve the parameter estimation problem. An extensive Monte-Carlo simulation compares the novel algorithms with existing errors-in-variables identification approaches.

Keywords: Bias elimination, Errors-in-variables, Frisch scheme, Least squares, Parameter estimation

1. INTRODUCTION

The identification of systems in the errors-in-variables (EIV) framework has recently attracted considerable attention. The EIV problem can be perceived as a generalization of the output error model structure (Ljung [1999]) in which measurement uncertainties are assumed to appear at the output exclusively. The EIV formulation treats the system in a symmetric manner allowing, therefore, both the input and output signals to be subjected to potential disturbances. This additional degree of freedom leads to a significant increase of the overall complexity of the identification problem giving rise to various techniques. For more details see Söderström et al. [2002], Söderström [2007] or Markovsky and Van Huffel [2007]. Two interesting techniques based on the bias compensating least squares principle (Stoica and Söderström [1982]) that proved to be robust and reliable in the EIV framework are: the Frisch scheme (FS) (Beghelli et al. [1990]) and the extended compensated least squares (ECLS) (Ekman [2005a]). Whilst the former method shows a significant noise robustness for a low signal to noise ratio (SNR) (high noise contamination) the latter technique can provide more accurate estimates in the cases of moderate or high SNR (low noise contamination), see Ekman [2005a].

This paper presents a generalized framework within which two new approaches are developed by combining the FS and the ECLS method. The resulting algorithms seem to exhibit the relatively high noise robustness from the FS, while retaining the precision of the ECLS technique, especially for the moderate SNR cases.

The paper is organized as follows: in the second section the notation used in the EIV framework together with the assumptions made and the problem statement are introduced. The third section reviews the ECLS algorithm together with the FS. Section four proposes generalizations of the FS and the ECLS approach and two novel algorithms are proposed within these frameworks. All algorithms are compared in a numerical simulation study given in Section five, whilst overall conclusions and further work are summarized in Section six.

2. NOTATION AND PROBLEM STATEMENT

Consider a discrete linear time-invariant (LTI) single-input single-output (SISO) system represented by the difference equation

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

where the polynomials $A(q^{-1})$ and $B(q^{-1})$ are given by

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

$$B(q^{-1}) \triangleq b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \tag{2b}$$

with q^{-1} being the backward shift operator, defined by $x_k q^{-1} \triangleq x_{k-1}$. The unknown noise free input and output signals denoted u_{0_k} and y_{0_k}, respectively, are related to the available noisy variables, denoted u_k and y_k, such that

$$u_k = u_{0_k} + \tilde{u}_k, \tag{3a}$$

$$y_k = y_{0_k} + \tilde{y}_k, \tag{3b}$$

where $\tilde{u}_k$ and $\tilde{y}_k$ denote the input and output measurement noise sequences, respectively. The following standard assumptions, see e.g. Söderström [2006], are introduced:

A1. The LTI 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 system structure, i.e. n_a and n_b are known *a priori* and $n_b \leq n_a$.

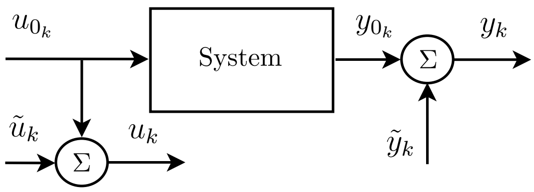


Fig. 1. Typical EIV SISO setup.

A3. The true input u_{0_k} is a zero mean, ergodic random sequence persistently exciting and of sufficiently high order, i.e. at least of order $n_a + n_b$.

A4. The input/output noise sequences are ergodic, zero mean, white processes with unknown variances $\sigma_{\tilde{u}}$ and $\sigma_{\tilde{y}}$, respectively, mutually uncorrelated and uncorrelated with the noise free signals u_{0_k} and y_{0_k}, respectively.

A diagrammatic illustration of the typical EIV setup for SISO systems is depicted in Figure 1.

The system parameter vector is denoted as

$$\theta^T \triangleq \left[a^T \;\; b^T\right] \in \mathbb{R}^{n_\theta}, \tag{4a}$$

$$a^T \triangleq [a_1 \; \dots \; a_{n_a}] \in \mathbb{R}^{n_a}, \tag{4b}$$

$$b^T \triangleq [b_1 \; \dots \; b_{n_b}] \in \mathbb{R}^{n_b}, \tag{4c}$$

where $n_\theta = n_a + n_b$. The extended regressor vectors for the measured data are defined as

$$\bar{\varphi}_k^T \triangleq \left[-y_k \;\; \varphi_k^T\right] \in \mathbb{R}^{n_\theta+1}, \tag{5a}$$

$$\bar{\varphi}_{y_k}^T \triangleq \left[-y_k \;\; \varphi_{y_k}^T\right] \in \mathbb{R}^{n_a+1}, \tag{5b}$$

where

$$\varphi_k^T \triangleq \left[\varphi_{y_k}^T \;\; \varphi_{u_k}^T\right] \in \mathbb{R}^{n_\theta}, \tag{5c}$$

$$\varphi_{y_k}^T \triangleq [-y_{k-1} \dots - y_{k-n_a}] \in \mathbb{R}^{n_a}, \tag{5d}$$

$$\varphi_{u_k}^T \triangleq [u_{k-1} \dots u_{k-n_b}] \in \mathbb{R}^{n_b}. \tag{5e}$$

The noise contributions in the corresponding regressor vectors are denoted by a tilde, i.e. $\tilde{[\cdot]}$, whereas the noise-free signals are denoted by a zero subscript, i.e. $[\cdot]_0$. From A4 it follows that

$$\bar{\varphi}_k = \bar{\varphi}_{0_k} + \tilde{\bar{\varphi}}_k. \tag{6}$$

The notation Σ_{cd} is used as a general notion for the covariance matrix of the vectors c_k and d_k, whereas ξ_{cf} is utilized for a covariance vector with f being a scalar, i.e.

$$\Sigma_{cd} \triangleq E\left[c_k d_k^T\right], \quad \Sigma_c \triangleq E\left[c_k c_k^T\right], \quad \xi_{cf} \triangleq E\left[c_k f_k\right]. \tag{7}$$

The corresponding estimates are denoted

$$\hat{\Sigma}_{cd} \triangleq \frac{1}{N}\sum_{k=1}^{N} c_k d_k^T, \quad \hat{\Sigma}_c \triangleq \frac{1}{N}\sum_{k=1}^{N} c_k c_k^T, \quad \hat{\xi}_{cf} \triangleq \frac{1}{N}\sum_{k=1}^{N} c_k f_k. \tag{8}$$

In addition, $0_{c\times d}$ denotes the null matrix of arbitrary dimension $c \times d$ and a single index is used in the case of a column vector as well as in the case of a square matrix, e.g. the identity matrix I_c.

The dynamic identification problem in the EIV framework considered here is formulated as:

Problem 1. Given N samples of the measured signals $\{u_k\}_{k=1}^N$ and $\{y_k\}_{k=1}^N$, determine the vector

$$\vartheta^T \triangleq \left[\theta^T \;\; \sigma_{\tilde{u}} \;\; \sigma_{\tilde{y}}\right] \in \mathbb{R}^{n_\theta+2}. \tag{9}$$

3. REVIEW

3.1 Extended bias compensated least squares (EBCLS)

The system (1-3) can be reformulated in the equation error form as

$$y_k = \varphi_k^T \theta + e_k, \tag{10}$$

where the residual is given by

$$e_k = \tilde{y}_k - \tilde{\varphi}_k^T \theta. \tag{11}$$

Denoting an estimate by $\hat{[\cdot]}$, the generalized solution of the system (10) in the least squares (LS) sense is given by

$$\hat{\theta} = \hat{\Sigma}_{x\varphi}^{\dagger} \hat{\xi}_{xy} \tag{12}$$

where $[\cdot]^\dagger$ is the Moore-Penrose pseudo inverse operator defined by $A^\dagger \triangleq (A^T A)^{-1} A^T$, $x_k \in \mathbb{R}^{n_x}$ denotes an arbitrary instrumental vector with $n_x \geq n_\theta$ and the covariance elements are defined according to (8). Note that for the case when $x_k = \varphi_k$ the basic LS estimate is obtained, whilst an arbitrary x_k gives rise to the instrumental variable (IV) estimator (see Ekman [2005b]). Due to the measurement noise, it is the case that unless the elements of x_k are uncorrelated with $\tilde{\varphi}_k$, the solution obtained is biased. In order to achieve a consistent estimate of θ, a bias compensation procedure must be carried out Söderström [2007]).

This yields the EBCLS estimator given by

$$\hat{\theta}_{\mathrm{EBCLS}} \triangleq \left(\hat{\Sigma}_{x\varphi} - \hat{\Sigma}_{\tilde{x}\tilde{\varphi}}\right)^{\dagger} \left(\hat{\xi}_{xy} - \hat{\xi}_{\tilde{x}\tilde{y}}\right), \tag{13}$$

Note that $\hat{\Sigma}_{\tilde{x}\tilde{\varphi}}$ and $\hat{\xi}_{\tilde{x}\tilde{y}}$, in general, are functions of $\hat{\sigma}_{\tilde{u}}$ and $\hat{\sigma}_{\tilde{y}}$ depending on the elements contained in the instrument vector x. It remains to determine the estimated input and output measurement noise variances $\hat{\sigma}_{\tilde{u}}$ and $\hat{\sigma}_{\tilde{y}}$. Different approaches are discussed in the subsequent development.

3.2 The separable nonlinear least squares (SNLS)

One possibility to determine $\hat{\sigma}_{\tilde{u}}$ and $\hat{\sigma}_{\tilde{y}}$ is the SNLS approach. With reference to Ekman [2005b], the following instruments are postulated

$$x_k^T \triangleq \left[\varphi_k^T \;\; \phi_k^T\right] \in \mathbb{R}^{n_x}, \tag{14}$$

where

$$\phi_k^T \triangleq \left[\phi_{y_k}^T \;\; \phi_{u_k}^T\right] \in \mathbb{R}^{2n_\phi+2}, \tag{15a}$$

$$\phi_{y_k}^T \triangleq \left[-y_k \;\; -y_{k-n_a-1} \; \dots \; -y_{k-n_a-n_\phi}\right] \in \mathbb{R}^{n_\phi+1}, \tag{15b}$$

$$\phi_{u_k}^T \triangleq \left[u_k \;\; u_{k-n_b-1} \; \dots \; u_{k-n_b-n_\phi}\right] \in \mathbb{R}^{n_\phi+1} \tag{15c}$$

with $n_x = n_\theta + 2n_\phi + 2$ and n_ϕ being defined as the maximum delay used in ϕ_{y_k} and ϕ_{u_k}. Note that n_ϕ determines the number of additional equations that are utilised for solving the identification problem. Consequently, the data covariance matrices are given by

$$\hat{\Sigma}_{x\varphi} \triangleq \begin{bmatrix} \hat{\Sigma}_{\varphi} \\ \hat{\Sigma}_{\phi\varphi} \end{bmatrix}, \qquad \hat{\xi}_{xy} \triangleq \begin{bmatrix} \hat{\xi}_{\varphi y} \\ \hat{\xi}_{\phi y} \end{bmatrix} \tag{16}$$

and the corresponding noise compensation matrices become in the asymptotic case

$$\Sigma_{\tilde{x}\tilde{\varphi}} \triangleq \begin{bmatrix} \Sigma_{\tilde{\varphi}} \\ 0_{(2n_\phi+2)\times n_\theta} \end{bmatrix}, \qquad \xi_{\tilde{x}\tilde{y}} \triangleq \begin{bmatrix} 0_{n_\theta} \\ \sigma_{\tilde{y}} \\ 0_{2n_\phi+1} \end{bmatrix} \tag{17}$$

with

$$\Sigma_{\tilde{\varphi}} \triangleq \begin{bmatrix} \sigma_{\tilde{y}} I_{n_a} & 0_{n_a \times n_b} \\ 0_{n_b \times n_a} & \sigma_{\tilde{u}} I_{n_b} \end{bmatrix}. \tag{18}$$

The underpinning idea of the SNLS is to estimate the input and output measurement noise variances by minimizing the norm of the resulting residuals

$$\{\hat{\sigma}_{\tilde{u}}, \hat{\sigma}_{\tilde{y}}\} \triangleq \arg\min_{\sigma_{\tilde{u}}, \sigma_{\tilde{y}}} V_1(\sigma_{\tilde{u}}, \sigma_{\tilde{y}}) \tag{19}$$

with

$$V_1(\sigma_{\tilde{u}}, \sigma_{\tilde{y}}) \triangleq \|\hat{\xi}_{xy} - \hat{\xi}_{\tilde{x}\tilde{y}} - \big(\hat{\Sigma}_{x\varphi} - \hat{\Sigma}_{\tilde{x}\tilde{\varphi}}\big) \times \big(\hat{\Sigma}_{x\varphi} - \hat{\Sigma}_{\tilde{x}\tilde{\varphi}}\big)^{\dagger} \big(\hat{\xi}_{xy} - \hat{\xi}_{\tilde{x}\tilde{y}}\big)\|_2^2. \tag{20}$$

The problem given by (20) is also referred to as a variable projection problem (see Golub and Peryera [1973] and Osborne [2007] for more details). Moreover, the search space of the optimization problem (20) can be restricted substantially by utilizing the results of the FS regarding the maximum admissible values of the input/output noise variances (Ekman [2005b]).

In the subsequent development the EBCLS algorithm which makes use of the input and output measurement noise variances computed by the SNLS, is referred to as ECLS.

3.3 The Frisch scheme (FS)

The FS approach is based on a particular case of the EBCLS technique (13) with $x_k = \varphi_k$, namely the standard bias compensating least squares (BCLS) approach. This means that the FS aims to remove the bias embedded in the estimates calculated via the ordinary LS method (Beghelli et al. [1990]), hence the BCLS estimate is given by

$$\hat{\theta}_{\text{BCLS}} \triangleq \left(\hat{\Sigma}_{\varphi} - \hat{\Sigma}_{\tilde{\varphi}}\right)^{-1} \hat{\xi}_{\varphi y}. \tag{21}$$

The major characteristic of the FS is that the output (or input) measurement noise variance is expressed as a nonlinear function of the input (or output) measurement noise variance. The so-called Frisch equation is given by

$$\hat{\sigma}_{\tilde{y}}^{\text{FS}} = \lambda_{\min}\Big[\hat{\Sigma}_{\bar{\varphi}_y} - \hat{\Sigma}_{\bar{\varphi}_y \varphi_u}\Big(\hat{\Sigma}_{\varphi_u} - \sigma_{\tilde{u}} I_{n_b}\Big)^{-1} \hat{\Sigma}_{\varphi_u \bar{\varphi}_y}\Big], \tag{22}$$

where $\lambda_{\min}[\cdot]$ is the least eigenvalue operator. In addition, a maximal admissible value of $\sigma_{\tilde{u}}$ can be determined by

$$\hat{\sigma}_{\tilde{u}}^{\max} = \lambda_{\min}\left[\hat{\Sigma}_{\varphi_u} - \hat{\Sigma}_{\varphi_u \bar{\varphi}_y} \hat{\Sigma}_{\bar{\varphi}_y}^{-1} \hat{\Sigma}_{\bar{\varphi}_y \varphi_u}\right]. \tag{23}$$

Equation (22) defines a whole set of Frisch models characterized by a convex curve in the noise space. In order to solve the identification problem, it remains to select a particular model which is uniquely characterized by an estimate of $\sigma_{\tilde{u}}$. Three common model selection criteria are:

Extended model criterion (EM): The Frisch equation is evaluated for the nominal and the extended model structure within the range $0 \leq \hat{\sigma}_{\tilde{u}} \leq \hat{\sigma}_{\tilde{u}}^{\max}$. This results in two curves in the so-called noise plane that theoretically intersect at a unique point, which corresponds to their true values (Beghelli et al. [1990]). The algorithm is denoted as FS_{EM}.

Covariance match criterion (CM): Statistical properties of the residuals computed from the system are compared with those predicted from a certain model (Diversi et al. [2003]). This algorithm is denoted as FS_{CM}.

Yule-Walker criterion (YW): The set of high order Yule-Walker equations can be exploited, which is equivalent to the utilization of an additional IV estimator that assesses the quality of the admissible solutions. In Diversi et al. [2006], the instrument vector is given by

$$\zeta_k^T \triangleq \left[u_{k-n_b-n_\zeta} \;\cdots\; u_{k-n_b-1}\right], \tag{24}$$

where $n_\zeta \geq n_\theta + 1$ is user specified. The corresponding algorithm is denoted as FS_{YW}.

A comprehensive study of these different methods can be found in Hong et al. [2007].

4. NOVEL ALGORITHMS

4.1 Generalized ECLS framework

This section presents a generalization of the ECLS algorithm, denoted the GCLS framework.

Choice of the same instrument vector x_k within (20) and (13) is unnecessarily restrictive. A more general ECLS scheme is given by

$$\hat{\theta}_{\text{GCLS}} \triangleq \left(\hat{\Sigma}_{z_1\varphi} - \hat{\Sigma}_{\tilde{z}_1\tilde{\varphi}}\right)^{\dagger} \left(\hat{\xi}_{z_1 y} - \hat{\xi}_{\tilde{z}_1\tilde{y}}\right) \tag{25}$$

with $\sigma_{\tilde{u}}$ and $\sigma_{\tilde{y}}$ being obtained by minimizing

$$V_2(\sigma_{\tilde{u}}, \sigma_{\tilde{y}}) \triangleq \|\hat{\xi}_{z_2 y} - \hat{\xi}_{\tilde{z}_2\tilde{y}} - \big(\hat{\Sigma}_{z_2\varphi} - \hat{\Sigma}_{\tilde{z}_2\tilde{\varphi}}\big) \times \big(\hat{\Sigma}_{z_3\varphi} - \hat{\Sigma}_{\tilde{z}_3\tilde{\varphi}}\big)^{\dagger} \big(\hat{\xi}_{z_3 y} - \hat{\xi}_{\tilde{z}_3\tilde{y}}\big)\|_p^2, \tag{26}$$

where p denotes a user chosen norm and z_{1_k}, z_{2_k} and z_{3_k} are arbitrarily chosen instrument vectors. The GCLS framework is hence given by:

Framework 1. (GCLS).

(1) Choose p and the instruments z_{1_k}, z_{2_k} and z_{3_k}.
(2) Compute the input and output measurement noise variances, i.e.

$$\{\hat{\sigma}_{\tilde{u}}, \hat{\sigma}_{\tilde{y}}\} = \arg\min_{\sigma_{\tilde{u}}, \sigma_{\tilde{y}}} V_2(\sigma_{\tilde{u}}, \sigma_{\tilde{y}}). \tag{27}$$

(3) Determine the parameter vector from (25).

Note that the ECLS is a particular case with $p = 2$ and $z_{1_k} = z_{2_k} = z_{3_k} = x_k$ as defined in (14).

4.2 Algorithm framework based on GCLS and the FS

Inspired by the GCLS approach, it is possible to reduce the optimization problem given by (27) to a minimization over a single variable only, using the Frisch equation (22). This leads to a novel framework, denoted GCLS-FS, which is given by:

Framework 2. (GCLS-FS).

(1) Choose p and the instruments z_{1_k}, z_{2_k} and z_{3_k}.
(2) Compute the input measurement noise variance, i.e.

$$\hat{\sigma}_{\tilde{u}} = \arg\min_{\sigma_{\tilde{u}}} V_2(\sigma_{\tilde{u}}, \hat{\sigma}_{\tilde{y}}^{\text{FS}}), \tag{28}$$

where dependency on $\hat{\sigma}_{\tilde{y}}^{\text{FS}}$ is substituted via (22).
(3) Compute the resulting output measurement noise variance using (22).
(4) Determine the parameter vector from (25).

Note that the algorithms within the GCLS-FS framework only belong to the family of FS algorithms if $z_{1_k} = \varphi_k$ is

Algorithm	z_{1_k}	z_{2_k}	z_{3_k}	Framework	FS	Reference
$\mathrm{FS_{YW}}$	φ_k	ζ_k	φ_k	GCLS-FS (Frisch equation utilized)	yes	Diversi et al. [2006]
Algorithm 1	φ_k	x_k	x_k	GCLS-FS (Frisch equation utilized)	yes	novel
Algorithm 2	x_k	x_k	x_k	GCLS-FS (Frisch equation utilized)	no	novel
ECLS	x_k	x_k	x_k	GCLS (Frisch equation not utilized)	no	Ekman [2005b]

Table 1. Overview of estimators within the proposed general frameworks.

selected. In general, z_{1_k} may contain arbitrary instruments with the consequence that the estimated parameter vector is not exclusively dependent on the computed input and output noise variances. This means that $\{\sigma_{\tilde{u}}, \sigma_{\tilde{y}}\}$ does not uniquely map into the parameter space, which is one of the major characteristics of the FS (Beghelli et al. [1990]). In the case where $z_{1_k} = \varphi_k$ holds, the 'Frisch-character' of the solution is retained and (28) can be interpreted as a general model selection criterion for the FS. Hence, this encompasses the YW criterion as a special case by choosing $z_{2_k} = \zeta_k$ defined in (24) and selecting $z_{3_k} = \varphi_k$. Furthermore, it is outlined in Hong et al. [2007], that under certain conditions the $\mathrm{FS_{YW}}$ and the $\mathrm{FS_{EM}}$ are equivalent, hence the $\mathrm{FS_{EM}}$ can also be interpreted in the GCLS-FS framework.

4.3 Two particular GCLS-FS realizations

Based on the choice of the instrument vectors in (25) and (26), an arbitrary number of estimators can be created. Two particular choices are proposed in this section. The first algorithm belongs to the family of FS algorithms utilizing a novel model selection criteria. It is denoted Algorithm 1 and the model selection criterion is defined by selecting $z_{1_k} = \varphi_k$ and $z_{2_k} = z_{3_k} = x_k$. This means that Algorithm 1 is equivalent to the $\mathrm{FS_{YW}}$ where the instrumental vector ζ_k is replaced by x_k given in (14). Consequently this algorithm is the FS with a novel model selection criterion. The algorithm can be summarized as:

Algorithm 1.

(1) Set $p = 2$ and specify the instruments as $z_{1_k} = \varphi_k$, $z_{2_k} = z_{3_k} = x_k$.
(2) Compute the input measurement noise variance by (28) where dependency on $\hat{\sigma}_{\tilde{y}}^{\mathrm{FS}}$ is substituted via (22).
(3) Compute the resulting output measurement noise variance using (22).
(4) Determine the parameter vector from (25) (which is reduced to (21)).

The second algorithm, denoted Algorithm 2, does not belong to the family of FS algorithms, since $z_{1_k} = x_k$ is chosen. In addition, as in the case of Algorithm 1, $z_{2_k} = z_{3_k} = x_k$ is used. This means Algorithm 2 is essentially the ECLS algorithm which, additionally, makes use of the Frisch equation.

Algorithm 2.

(1) Set $p = 2$ and specify the instruments as $z_{1_k} = z_{2_k} = z_{3_k} = x_k$.
(2) Compute the input measurement noise variance by (28) where dependency on $\hat{\sigma}_{\tilde{y}}^{\mathrm{FS}}$ is substituted via (22).
(3) Compute the resulting output measurement noise variance using (22).
(4) Determine the parameter vector from (25).

Remark 1. Note that the Frisch equation (22) is equivalent to

$$\left(\Sigma_{\bar{\varphi}} - \begin{bmatrix} \hat{\sigma}_{\tilde{y}}^{FS} I_{n_a+1} & 0 \\ 0 & \hat{\sigma}_{\tilde{u}} I_{n_b} \end{bmatrix}\right) \hat{\bar{\theta}} = 0. \tag{29}$$

Therefore, if z_1 already contains $\bar{\varphi}$, as in the case of Algorithm 1, the additional usage of the Frisch equation (22) may seem redundant. However, by computing $\hat{\sigma}_{\tilde{y}}$ via (22), the relation (29) is forced to hold exactly. In contrast, although the set of equations (29) has been implicitly utilized to estimate θ, in the case of Algorithm 2 this relation holds only approximately.

4.4 Overview

A general overview of the algorithms that can be interpreted in the proposed GCLS as well as in the GCLS-FS framework is given in Table 1. The fifth column basically allows the algorithms to be differentiated according to the particular framework to which they belong. In addition, as emphasized in Section 4.2, it is important to note that, although Algorithm 2 is a member of the GCLS-FS framework by virtue of utilizing the Frisch equation, it does not, however, exhibit the FS properties.

5. SIMULATION STUDIES

This section addresses a numerical analysis of the two proposed algorithms when employed for the identification of a standard SISO LTI second order system within the EIV framework. The system is given by the following discrete-time transfer function

$$G(q^{-1}) = \frac{1.0q^{-1} + 0.5q^{-1}}{1 - 1.5q^{-1} + 0.7q^{-2}}, \tag{30}$$

where the input u_{0_k} is a white zero mean sequence with variance 0.5. In terms of accuracy, the overall quality of the estimators is assessed via the following two performance criteria:

$$e_1 \triangleq \|\hat{\theta}_j - \theta\|_2^2, \tag{31a}$$

$$e_2 \triangleq \|\,[\hat{\sigma}_{\tilde{u},j}\ \hat{\sigma}_{\tilde{y},j}] - [\sigma_{\tilde{u}}\ \sigma_{\tilde{y}}]\,\|_2^2 \tag{31b}$$

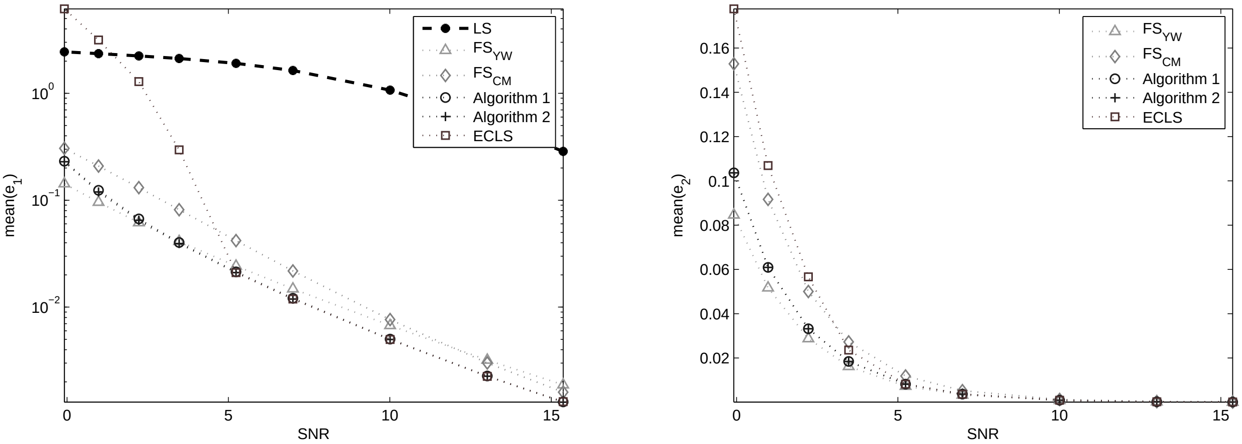
where j denotes the $j-th$ Monte-Carlo simulation from which the estimates are computed. It is remarked that e_1 assesses the quality of the estimated parameter vector, whilst e_2 assesses the quality of the estimated input/output noise variances.

5.1 Experiment 1

In the first experiment Algorithm 1 and Algorithm 2 are compared with other EIV techniques, namely $\mathrm{FS_{YW}}$, $\mathrm{FS_{CM}}$, ECLS as well as with the standard LS for completeness. In the case of the $\mathrm{FS_{YW}}$ method the additional instruments of length $n_\theta + 1$ are constructed as described

Table 2. Results of the estimation of parameters and noise variances with SNR $\approx$ 10dB and SNR $\approx$ 3.5dB.

	true	FS_{YW}	FS_{CM}	Algorithm 1	Algorithm 2	ECLS	LS
				SNR $\approx$ 10dB			
a_1	-1.500	-1.497 ± 0.020	-1.497 ± 0.020	-1.500 ± 0.011	-1.500 ± 0.011	-1.500 ± 0.011	-0.855 ± 0.015
a_2	0.700	0.697 ± 0.018	0.697 ± 0.018	0.700 ± 0.010	0.700 ± 0.009	0.700 ± 0.009	0.110 ± 0.014
b_1	1.000	0.995 ± 0.040	0.996 ± 0.053	0.990 ± 0.052	0.990 ± 0.052	0.991 ± 0.052	0.905 ± 0.028
b_2	0.500	0.511 ± 0.065	0.512 ± 0.063	0.505 ± 0.045	0.505 ± 0.045	0.505 ± 0.045	1.044 ± 0.029
$\sigma_{\tilde{y}}$	0.950	0.948 ± 0.028	0.948 ± 0.031	0.949 ± 0.028	0.949 ± 0.028	0.949 ± 0.028	–
$\sigma_{\tilde{u}}$	0.050	0.051 ± 0.011	0.051 ± 0.019	0.048 ± 0.013	0.048 ± 0.013	0.048 ± 0.014	–
e_1	–	0.007 ± 0.007	0.008 ± 0.007	0.005 ± 0.005	0.005 ± 0.005	0.005 ± 0.005	1.070 ± 0.053
e_2	–	0.001 ± 0.001	0.001 ± 0.002	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	–
				SNR $\approx$ 3.5dB			
a_1	-1.500	-1.488 ± 0.049	-1.494 ± 0.073	-1.497 ± 0.041	-1.498 ± 0.036	-1.500 ± 0.041	-0.493 ± 0.014
a_2	0.700	0.689 ± 0.045	0.695 ± 0.063	0.698 ± 0.036	0.699 ± 0.030	0.700 ± 0.034	-0.140 ± 0.012
b_1	1.000	0.998 ± 0.099	0.987 ± 0.182	0.985 ± 0.146	0.985 ± 0.149	1.049 ± 0.461	0.684 ± 0.045
b_2	0.500	0.538 ± 0.160	0.532 ± 0.196	0.516 ± 0.125	0.514 ± 0.122	0.479 ± 0.282	1.042 ± 0.046
$\sigma_{\tilde{y}}$	4.275	4.265 ± 0.124	4.266 ± 0.137	4.267 ± 0.126	4.267 ± 0.126	4.262 ± 0.134	–
$\sigma_{\tilde{u}}$	0.225	0.231 ± 0.028	0.212 ± 0.092	0.219 ± 0.050	0.219 ± 0.050	0.228 ± 0.074	–
e_1	–	0.041 ± 0.044	0.082 ± 0.073	0.040 ± 0.041	0.039 ± 0.041	0.296 ± 1.817	2.117 ± 0.071
e_2	–	0.016 ± 0.022	0.027 ± 0.031	0.019 ± 0.023	0.019 ± 0.023	0.024 ± 0.043	–



(a) The mean of the performance criterion e_1.

(b) The mean of the performance criterion e_2.

Fig. 2. The performance of algorithms as a function of incrementally increasing SNR.

in (24). For the FS_{CM} the time shift and the weighting matrix parameters are set to unity and the identity matrix, respectively. The number of instruments for the ECLS, Algorithm 1 and Algorithm 2 is set to $n_x = 10$, i.e. $n_\phi = 2$. Two scenarios with approximately equal SNR on the input and output are considered, i.e. SNR $\approx$ SNR_u $\approx$ SNR_y. The first scenario exhibits SNR $\approx$ 10dB ($\sigma_{\tilde{u}} = 0.050$, $\sigma_{\tilde{y}} = 0.950$) whereas the second setup uses SNR $\approx$ 3.5dB ($\sigma_{\tilde{u}} = 0.225$, $\sigma_{\tilde{y}} = 4.275$). The system is simulated using 200 Monte-Carlo runs and $N = 5000$ samples. The results of this experiment are given in Table 2, where the mean values and standard deviations of $\hat{\vartheta}$, e_1 and e_2 with respect to the Monte-Carlo simulations are recorded.

It it observed that for the higher SNR (low noise contamination) the ECLS, Algorithm 1 and Algorithm 2 produce virtually identical results, which, with respect to e_1 and its associated standard deviation, are the better among the algorithms evaluated. Considering the lower SNR case (high noise contamination), a substantial decrease in the quality of the estimates produced by ECLS is observed, particularly for the estimates of the $B(q^{-1})$ coefficients. On the contrary, Algorithm 1 and Algorithm 2 retain their relatively superior performance yielding the best performance for the estimates of the unknown system parameters and hence the lowest value for the index e_1 in this case. Note that Algorithm 2 is characterized by slightly lower values for the associated standard deviations than Algorithm 1 in this case.

The index e_2 reveals that the FS_{YW} method achieves a slightly better performance for the estimates of the input/output noise variances. Therefore, it is suspected that in the case of Algorithm 2, it is the utilization of the extended normal equations within the final bias compensation phase that leads to an improved accuracy of the estimated model parameters $\hat{\theta}$.

5.2 *Experiment 2*

The performances of the algorithms are evaluated considering a wider range of the potential SNR. The same system setup is used as in Experiment 1, however the SNR is incrementally increased starting from about 0dB up to about 15dB and preserving the condition $SNR_u \approx SNR_y$.

The performance indices e_1 and e_2 are recorded and their mean values are plotted against the SNR in Figure 2(a) and Figure 2(b), respectively.

Considering e_1, it is observed that whilst the ECLS, Algorithm1 and Algorithm2 are characterized by the better estimates with similar qualities being obtained for the high SNR cases (low noise contamination), the performance of the ECLS deteriorates significantly for the SNR levels lower than about 5dB. A similar observation is apparent in Figure 2(b), when a decline in the quality of the ECLS method is noted for the same cases. The performances of Algorithm 1 and Algorithm 2 are virtually identical yielding the most accurate estimates for the cases greater than about 2.5dB. For the lower SNR cases, i.e. when SNR $<$ 2.5dB the $\mathrm{FS}_{\mathrm{YW}}$ method seems to produce slightly smaller values of e_1. Considering Figure 2(b) it is observed that the $\mathrm{FS}_{\mathrm{YW}}$ obtains the most precise estimates of the input/output noise variances, whilst Algorithm 1 and Algorithm 2 in general produce the joint second bests.

6. CONCLUSIONS

The extended compensated least squares and the Frisch scheme for the identification of dynamical linear time-invariant single-input single-output errors-in-variables models have been reviewed within generalized frameworks. It is shown that the well known Frisch scheme - Yule-Walker algorithm as well as the extended compensated least squares method can be interpreted as members of these general frameworks. Moreover, two novel algorithms have been proposed. Whilst the first method belongs to the family of Frisch scheme algorithms using a novel model selection criterion, the second method is a modification of the extended compensated least squares technique. An extensive Monte-Carlo simulation, which compares the proposed algorithms for different signal-to-noise ratios, has been carried out. For the cases considered, the new algorithms appear to combine the advantages of the Frisch scheme and the extended compensated least squares techniques, which can yield an improved accuracy of the estimated parameters.

Future work could consider the consistency properties where these features are suspected to be assured by the properties inherited from the constituent algorithms. Moreover, potential extensions to handle the case of coloured noise on the system output measurements is to be considered together with the recursive implementation of the two new algorithms.

REFERENCES

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

R. Diversi, R. Guidorzi, and U. Soverini. A new criterion in EIV identification and filtering applications. In *13th IFAC on Symp. on System Identification*, pages 27–29, Rotterdam, The Netherlands, 2003.

R. Diversi, R. Guidorzi, and U. Soverini. Yule-Walker equations in the Frisch scheme solution of errors-in-variables identification problems. In *Proc. of the 17th Int. Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan, 2006.

M. Ekman. *Modeling and Control of Bilinear Systems: Applications to the Activated Sludge Process.* PhD thesis, Uppsala University, Sweden, 2005a.

M. Ekman. Identification of linear systems with errors in variables using separable nonlinear least squares. In *Proc. of 16th IFAC World Congress*, Prague, Czech Republic, 2005b.

G. H. Golub and V. Peryera. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM J. of Numerical Analysis*, 10 (2):413–432, 1973.

M. Hong, T. Söderström, U. Soverini, and R Diversi. Comparison of three Frisch methods for errors-in-variables identification. Internal Report 2007-021, Department of Information Technology, Uppsala University, Sweden, 2007.

L. Ljung. *System Identification - Theory for the User.* Prentice Hall PTR, New Jersey, USA, 2nd edition, 1999.

I. Markovsky and S. Van Huffel. Overview of total least-squares methods. In *Signal Proc.*, pages 2283–2302. Elsevier, 2007.

M. R. Osborne. Separable least squares, variable projection, and the Gauss-Newton algorithm. *Electronic Trans. on Numerical Analysis.*, 28:1–15, 2007.

T. Söderström. Statistical analysis of the Frisch scheme for identifying errors-in-variables systems. Technical report, Uppsala University, Department of Information Technology, Uppsala, Sweden, 2006.

T. Söderström. Errors-in-variables methods in system identification. In *Automatica*, volume 43, pages 939–958, 2007.

T. Söderström, U. Soverini, and K. Mahata. Perspectives on errors-in-variables estimation for dynamic systems. In *Signal Proc.*, volume 82(8), pages 1139–1154, 2002.

P. Stoica and T. Söderström. Bias correction in least-squares identification. *Int. J. of Control*, 35(3):449–457, 1982.