HAL
open science

Survey of analytical IV estimates for errors-in-variables model identification

Stéphane Thil, Marion Gilson

To cite this version:

Stéphane Thil, Marion Gilson. Survey of analytical IV estimates for errors-in-variables model identification. 18th IFAC World Congress, IFAC WC'2011, Aug 2011, Milan, Italy. pp.13098-13103. hal-00691860v2

HAL Id: hal-00691860

https://hal.science/hal-00691860v2

Submitted on 27 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Survey of analytical IV estimates for errors-in-variables model identification

Stéphane Thil *, **Marion Gilson** **

* *Laboratoire PROMES (UPR CNRS 8521), Université de Perpignan Via Domitia, 52 avenue Paul Alduy, 66860 Perpignan, France. E-mail address: stephane.thil@univ-perp.fr*
** *Centre de Recherche en Automatique de Nancy (CRAN UMR 7039), Nancy-Université, CNRS, BP 239, 54506, Vandœuvre-lès-Nancy Cedex, France. E-mail address: marion.gilson@cran.uhp-nancy.fr*

Abstract: This paper deals with discrete-time errors-in-variables model identification, where output and input data are both perturbed by additive noises. The goal of this paper is to show how instrumental variable techniques may handle this situation; more precisely, the focus is put on instrumental variable methods having closed-form solutions. Several methods using second- and high-order statistics are reviewed, and some improvements and new results are also introduced. The discussions are illustrated by means of a numerical simulation.

Keywords: System identification; errors-in-variables; instrumental variable; bias compensation.

1. INTRODUCTION

System identification is an established field in the area of system analysis and control. It aims at determining mathematical models for dynamical systems based on observed inputs and outputs. Many solutions have been presented for system identification of linear dynamic systems from noise-corrupted output measurements. The identification of models where uncertainties are present on both input and output observations – called 'errors-in-variables' (EIV) models – have received far less attention. This type of models plays an important role when the identification purpose is the determination of the inner laws that describe the process, rather than the prediction of its future behaviour. Numerous scientific disciplines use such errors-in-variables models, including time series modelling, array signal processing for direction-of-arrival estimation, blind channel equalisation, multivariate calibration in analytical chemistry, image processing, or environmental modelling Van Huffel (2002).

Identification of EIV models has been an active domain of research in the recent years (see the survey paper Söderström (2007)). Among the 'classical' approaches for EIV model identification, the instrumental variable (IV) methods have not yet witnessed considerable interest. An explanation can be that the asymptotic accuracy of the *basic* IV method cannot reach the Cramér-Rao lower bound Söderström (2007). However, there is a considerable freedom in the choice of the instrument vector and many ideas stem from the basic IV principle, *e.g.* iterative methods Young (1970, 1984); Gilson and Van den Hof (2005); Young et al. (2008); Thil et al. (2008b) or the use of higher-order statistics Inouye and Tsuchiya (1991); Inouye and Suga (1994); Thil et al. (2009).

The aim of this paper is first to enumerate the different possibilities offered by IV-based methods to solve the EIV model identification problem (with the focus on methods having closed-form solutions) and offer a comparison between these IV-based methods; second to introduce a few improvements and some new results.

The paper organised as follows. The problem is formally stated in Section 2. In Sections 3–5, different IV variants are exposed. Finally, before concluding, the performances of the methods are assessed by means of a simulation example in Section 6.

2. ERRORS-IN-VARIABLES MODEL

Consider a discrete-time, linear, time-invariant EIV system. The data-generating system is given by:

$$\mathcal{S} : \begin{cases} y_o(t) = G_o(q)u_o(t) \\ u(t) = u_o(t) + \tilde{u}(t) \\ y(t) = y_o(t) + \tilde{y}(t) \end{cases} \tag{1}$$

where q is the forward operator, $G_o(\cdot)$ is the transfer operator of the 'true' system and $\tilde{u}$, $\tilde{y}$ are the input and output noises, respectively. The system is then parameterised as follows:

$$\begin{cases} y(t) = G(q,\boldsymbol{\theta})\,(u(t) - \tilde{u}(t)) + \tilde{y}(t) \\ G(q,\boldsymbol{\theta}) = B(q^{-1},\boldsymbol{\theta})/A(q^{-1},\boldsymbol{\theta}) \\ A(q^{-1},\boldsymbol{\theta}) = 1 + a_1 q^{-1} + ... + a_{n_a} q^{-n_a} \\ B(q^{-1},\boldsymbol{\theta}) = b_1 q^{-1} + ... + b_{n_b} q^{-n_b} \end{cases} \tag{2}$$

with $n_a \geqslant n_b$ and $\boldsymbol{\theta}^{\mathrm{T}} = [a_1 \ \ldots \ a_{n_a} \ b_1 \ \ldots \ b_{n_b}]$. Equation (2) can be rewritten as:

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t,\boldsymbol{\theta}) \tag{3}$$

$$v(t,\boldsymbol{\theta}) = \tilde{y}(t) - \tilde{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{\theta} \tag{4}$$

where the regression vector is given as:

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = \big[-y(t-1) \ \dots \ -y(t-n_a) \quad u(t-1) \ \dots \ u(t-n_b)\big] \tag{5}$$

and $\tilde{\boldsymbol{\varphi}}(t)$ is defined in a similar way to $\boldsymbol{\varphi}(t)$, but with u and y being replaced by $\tilde{u}$ and $\tilde{y}$, respectively.

The problem of identifying this errors-in-variables model is concerned with consistently estimating the parameter vector $\boldsymbol{\theta}$ from the noisy data $\{u(t), y(t)\}_{t=1}^{N}$.

2.1 Notations

As the input/output noises are additive, linear functions of the measured signals can be broken down into two parts, one made up of the noise-free signal contribution (denoted with a 'o' subscript) and the other of the noises contribution (denoted with the '˜' sign). For conciseness' sake, it is also convenient to define:

$$\bar{\boldsymbol{a}}^{\mathrm{T}} = \big[1 \ \boldsymbol{a}^{\mathrm{T}}\big] = [1 \ a_1 \ \dots \ a_{n_a}] \tag{6}$$

$$\boldsymbol{b}^{\mathrm{T}} = [b_1 \ \dots \ b_{n_b}] \tag{7}$$

$$\bar{\boldsymbol{y}}^{\mathrm{T}}(t,n) = \big[y(t) \ \boldsymbol{y}^{\mathrm{T}}(t,n)\big] \tag{8}$$

$$\boldsymbol{y}^{\mathrm{T}}(t,n) = [y(t-1) \ \dots \ y(t-n)] \tag{9}$$

$$\boldsymbol{u}^{\mathrm{T}}(t,n) = [u(t-1) \ \dots \ u(t-n)] \tag{10}$$

For example, the regression vector can now be written as:

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = \big[-\boldsymbol{y}(t,n_a) \ \boldsymbol{u}(t,n_b)\big] \tag{11}$$

The following notations are used in the sequel for the correlation vectors and matrices:

$$\boldsymbol{R}_{\varphi\varphi} = \mathrm{E}\big\{\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\big\} \qquad \boldsymbol{r}_{\varphi y} = \mathrm{E}\{\boldsymbol{\varphi}(t)y(t)\} \tag{12}$$

Lastly, an estimate of the parameter vector $\boldsymbol{\theta}$ from N samples of input/output data and its asymptotic version are written as:

$$\hat{\boldsymbol{\theta}} \xrightarrow[N\to\infty]{} \boldsymbol{\theta}^{\star} \tag{13}$$

2.2 Assumptions

The following assumptions are needed:

A1. The system (1) is asymptotically stable, and all the system modes are observable and controllable;
A2. The signals u_o, $\tilde{u}$ and $\tilde{y}$ are stationary, ergodic and zero-mean;
A3. The signals $\tilde{u}$ and $\tilde{y}$ are assumed to be uncorrelated with the input u_o.

In system identification literature, it is usually assumed (often implicitly) that the 'true' system belongs to the considered model set, a situation referred to as $\mathcal{S} \in \mathcal{G}^{\star}$ Ljung (1999). However, this notation is not properly suited for errors-in-variables models and some additional notations must be introduced Thil et al. (2009); the whole 'true' system includes:

- the 'true' process G_o and its associated model set $\mathcal{G}^{\star} = \{G(\cdot,\boldsymbol{\theta})\}$;
- the 'true' noise processes $H_o^{\tilde{u}}$, $H_o^{\tilde{y}}$ and their associated model set $\mathcal{H}^{\star} = \big\{\big(H^{\tilde{u}}(\cdot,\boldsymbol{\eta}),\ H^{\tilde{y}}(\cdot,\boldsymbol{\eta})\big)\big\}$;
- the 'true' noise-free input process $H_o^{u_o}$ and its associated model set $\mathcal{E}^{\star} = \{H^{u_o}(\cdot,\boldsymbol{\eta})\}$;

where $\boldsymbol{\eta}$ is a vector gathering the parameters of the noise-free input and noise models. In this paper it is assumed that the 'true' process belongs to the model set and that the noises' models belong to the model set as well:

A4. $G_o \in \mathcal{G}^{\star}$ and $\big(H_o^{\tilde{u}}, H_o^{\tilde{y}}\big) \in \mathcal{H}^{\star}$.

It is also supposed that:

A6. $\tilde{u}$ is a white noise with a symmetric probability density function (pdf);
A7. $\tilde{y}$ is a white noise with a symmetric pdf;
A8. u_o has a skewed pdf.

The assumption of a symmetric pdf is rather mild (a uniformly distributed or gaussian noise is a very common assumption – with reason) and will prove useful when using methods based on higher-order statistics (HOS).

3. CLASSIC (EXTENDED) IV

Let $\boldsymbol{z}(t)$ be a vector of instruments of dimension $n_z \geqslant n_a + n_b$. An eXtended IV estimator of $\boldsymbol{\theta}$ is given by Söderström and Stoica (1983):

$$\boldsymbol{\theta}^{\star}_{\mathsf{xiv}} = \big(\boldsymbol{R}_{z\varphi}^T\boldsymbol{R}_{z\varphi}\big)^{-1}\boldsymbol{R}_{z\varphi}^T\boldsymbol{r}_{zy} \tag{14a}$$

$$= \boldsymbol{\theta} + \big(\boldsymbol{R}_{z\varphi}^T\boldsymbol{R}_{z\varphi}\big)^{-1}\boldsymbol{r}_{zv} \tag{14b}$$

For the estimator (14) to be unbiased, the vector of instruments $\boldsymbol{z}(t)$ must be uncorrelated with the composite error $v(t,\boldsymbol{\theta})$, while being 'as correlated as possible' with the regression vector $\boldsymbol{\varphi}(t)$, so that the matrix to be inverted in (14) is not ill-conditioned, which is summarised as:

$$\boldsymbol{r}_{zv} = 0 \tag{15}$$

$$\boldsymbol{R}_{z\varphi} \text{ is nonsingular} \tag{16}$$

Remark 1
In the classic IV method, these two requirements are generally contradictory: the delays used to obtain an instrument vector uncorrelated with the noise also decrease the correlation between the regressor and the instrument vector.

3.1 First approach: $\mathsf{xiv_{uu}}$ method

In the first approach the instrument vector contains delayed inputs only Söderström and Mahata (2002):

$$\boldsymbol{z}^T(t) = \boldsymbol{u}^T(t-d_1, n_{z_u}) \tag{17}$$

$$= \big[u(t-1-d_1) \ \dots \ u(t-n_{z_u}-d_1)\big] \tag{18}$$

Since $\tilde{u}$, $\tilde{y}$ and u_o are uncorrelated:

$$\boldsymbol{r}_{zv} = \mathrm{E}\{\boldsymbol{z}(t)v(t,\boldsymbol{\theta})\} \tag{19}$$

$$= \mathrm{E}\Big\{\boldsymbol{z}(t)\Big(\tilde{\bar{\boldsymbol{y}}}^T(t,n_a)\bar{\boldsymbol{a}} - \tilde{\boldsymbol{u}}^T(t,n_b)\boldsymbol{b}\Big)\Big\} \tag{20}$$

$$= -\mathrm{E}\big\{\tilde{\boldsymbol{u}}(t-d_1,n_{z_u})\tilde{\boldsymbol{u}}^T(t,n_b)\big\}\,\boldsymbol{b} \tag{21}$$

Thus, if the delay is such that $d_1 \geqslant n_b$, the vector of instruments $\boldsymbol{z}(t)$ satisfies (15).

3.2 Second approach: $\mathsf{xiv_{yu}}$ method

Let us define an instrument vector containing delayed inputs and delayed outputs as Thil et al. (2008b):

$$\boldsymbol{z}^T(t) = \big[-\boldsymbol{y}^T(t-d_2,n_{z_y}) \ \boldsymbol{u}^T(t-d_3,n_{z_u})\big] \tag{22}$$

Since $\tilde{u}$, $\tilde{y}$ and u_o are uncorrelated:

$$\boldsymbol{r}_{zv} = \mathrm{E}\{\boldsymbol{z}(t)v(t,\boldsymbol{\theta})\} \tag{23}$$

$$= \mathrm{E}\left\{\begin{array}{l} -\tilde{\boldsymbol{y}}(t-d_2,n_{z_y})\tilde{\bar{\boldsymbol{y}}}^T(t,n_a)\bar{\boldsymbol{a}} \\ -\tilde{\boldsymbol{u}}(t-d_3,n_{z_u})\tilde{\boldsymbol{u}}^T(t,n_b)\boldsymbol{b} \end{array}\right\} \tag{24}$$

If the delays satisfy $d_2 \geqslant n_a + 1$ and $d_3 \geqslant n_b$, the vector of instruments $\boldsymbol{z}(t)$ satisfies (15).

The autocorrelation of $y(t)$ should be higher than the cross-correlation between $u(t)$ and $y(t)$. Thus, in this approach the correlation between the regressor and the instrument vector should be increased, which should improve the conditioning of the matrix $\boldsymbol{R}_{z\varphi}$; in turn, the $\mathsf{xiv}_{\mathsf{yu}}$ method should yield more accurate estimates than the $\mathsf{xiv}_{\mathsf{uu}}$ method.

4. HOS-BASED METHODS

The freedom inherent to IV methods is such that one can actually obtain higher-order statistics (HOS) methods from the simple IV principle. Only a few methods have been developed Inouye and Tsuchiya (1991); Chen and Chen (1994); Thil et al. (2008a), all of which have been showed to stem from the same set of equations in Thil et al. (2009). Before exposing the unified method, let us recall a few of the numerous properties of higher-order cumulants. For a more complete treatment see *e.g.* Brillinger (1981); Mendel (1991).

The third-order cumulants of stationary signals x_1, x_2, x_3 are defined and written as:

$$C_{x_1x_2x_3}(\tau_1,\tau_2) = \mathrm{E}\{x_1(t)x_2(t+\tau_1)x_3(t+\tau_2)\} \quad (25)$$
$$= C_{x_1x_2x_3}(\boldsymbol{\tau}) \quad (26)$$

where $\boldsymbol{\tau}$ denotes $[\tau_1,\tau_2]$.

4.1 Properties of HOS

The properties used in the sequel are:

P1. Multilinearity: cumulants are linear with respect to each of their arguments;
P2. Additivity: if two random vectors are independent, then the cumulant of their sum equals the sum of their cumulants;
P3. The third-order cumulant of a random variable with a symmetric pdf is equal to zero.

From assumptions A3, A6, A7, A8 and using properties P2, P3, the following holds:

$$C_{uuy}(\boldsymbol{\tau}) = C_{u_0u_0y_0}(\boldsymbol{\tau}) + C_{\tilde{u}\tilde{u}\tilde{y}}(\boldsymbol{\tau}) = C_{u_0u_0y_0}(\boldsymbol{\tau})$$
$$C_{uuu}(\boldsymbol{\tau}) = C_{u_0u_0u_0}(\boldsymbol{\tau}) + C_{\tilde{u}\tilde{u}\tilde{u}}(\boldsymbol{\tau}) = C_{u_0u_0u_0}(\boldsymbol{\tau})$$

The third-order (cross-)cumulants of the input and output signals are thus insensitive to symmetrically distributed noises. Note that this result is still valid for the third-order (cross-)cumulant of any combination of input and output signals.

The so-called third-order cumulant-based extended instrumental variable method (tocxiv) can now be reviewed.

4.2 The tocxiv method

The EIV model can be shown to satisfy the following third-order cumulants equation Thil et al. (2008a):

$$C_{uuy}(\boldsymbol{\tau}) = \frac{B(q^{-1},\boldsymbol{\theta})}{A(q^{-1},\boldsymbol{\theta})} C_{uuu}(\boldsymbol{\tau}) \quad (27)$$

which can be expressed in a linear regression form as:

$$C_{uuy}(\boldsymbol{\tau}) = \boldsymbol{\varphi}_{uuy}^{\mathrm{T}}(\boldsymbol{\tau})\boldsymbol{\theta} \quad (28)$$

where:

$$\boldsymbol{\varphi}_{uuy}^{\mathrm{T}}(\boldsymbol{\tau}) = \big[-C_{uuy}(\tau_1,\tau_2-1) \ \ldots \ -C_{uuy}(\tau_1,\tau_2-n_a) \quad C_{uuu}(\tau_1,\tau_2) \ \ldots \ C_{uuu}(\tau_1,\tau_2-n_b)\big] \quad (29)$$

Due to the properties of third-order cumulants, this equation is (asymptotically) noise-free. After having chosen a cumulant slice and avoided the use of third-order cumulants with large time-lags, the parameter vector is obtained by solving the following system of equations in a least squares sense. Written in a compact matrix form it yields (see Thil et al. (2008a, 2009)):

$$\boldsymbol{C} = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\theta} \quad (30)$$

The tocxiv estimator is then given as:

$$\boldsymbol{\theta}^{\star}_{\mathsf{tocxiv}} = \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}}\right)^{-1}\boldsymbol{\Phi}\boldsymbol{C} \quad (31)$$

Again, two versions of this method will be tested: the $\mathsf{tocxiv}_{\mathsf{uuy}}$ method (using (27)), and the $\mathsf{tocxiv}_{\mathsf{yyy}}$ method using the following equation (see Thil et al. (2009)):

$$C_{yyy}(\boldsymbol{\tau}) = \frac{B(q^{-1},\boldsymbol{\theta})}{A(q^{-1},\boldsymbol{\theta})} C_{yyu}(\boldsymbol{\tau}) \quad (32)$$

It is yet unclear which equation should be used. Exploring how the properties of the system and the signals influence this choice is a difficult task and an interesting topic for future research.

5. BIAS COMPENSATED IV

The idea of compensating the IV bias has been used in Yang et al. (1993) and more recently in Ekman (2005); Ekman et al. (2006) and Thil et al. (2008b). Since only the methods having a closed-form solution are studied in this paper, the method presented in Ekman (2005); Ekman et al. (2006) is not reviewed here: indeed, while it can be seen as an extension of the IV method, it relies on a numerical optimisation to estimate the parameters.

5.1 The bciv method

Define a vector of instruments leading to a *biased* estimate of $\boldsymbol{\theta}$:

$$\boldsymbol{z}^{\mathrm{T}}(t) = \left[\boldsymbol{x}^{\mathrm{T}}(t,n_a) \ \ \boldsymbol{u}^{\mathrm{T}}(t,n_b)\right] \quad (33)$$

where $\boldsymbol{x}(t,n_a)$ is chosen such that $\tilde{\boldsymbol{x}}(t,n_a)$ is uncorrelated with $\tilde{\boldsymbol{y}}(t,n_a)$. The biased instrumental variable (BIV) estimator of $\boldsymbol{\theta}$ is then given by:

$$\boldsymbol{\theta}^{\star}_{\mathsf{biv}} = \boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{r}_{zy} = \boldsymbol{\theta} + \boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{r}_{zv} \quad (34)$$

We have:

$$\boldsymbol{r}_{zv} = -\lambda_{e_{\tilde{u}}} \begin{bmatrix} \boldsymbol{0}_{n_a\times n_a} & \boldsymbol{0}_{n_a\times n_b} \\ \boldsymbol{0}_{n_b\times n_a} & \boldsymbol{I}_{n_b\times n_b} \end{bmatrix} \boldsymbol{\theta} = -\lambda_{e_{\tilde{u}}}\boldsymbol{M}\boldsymbol{\theta} \quad (35)$$

Thus finally:

$$\boldsymbol{\theta}^{\star}_{\mathsf{biv}} = \boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{r}_{zy} = \boldsymbol{\theta} - \lambda_{e_{\tilde{u}}}\boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{M}\boldsymbol{\theta} \quad (36)$$

If an estimate of the input noise variance $\lambda_{e_{\tilde{u}}}$ is available, then the bias of the estimator (34) can be subtracted to obtain a consistent estimate of $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}^{\star}_{\mathsf{bciv}} = \boldsymbol{\theta}^{\star}_{\mathsf{biv}} + \lambda_{e_{\tilde{u}}}\boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{M}\boldsymbol{\theta} \quad (37)$$

Estimation of $\lambda_{e_{\tilde{u}}}$ The key problem to solve now is the estimation of $\lambda_{e_{\tilde{u}}}$. To achieve this, the idea is to use the correlation between the composite noise $v(t,\boldsymbol{\theta})$ and the 'residual' r, defined as:

$$r_1(t,\boldsymbol{\theta}) = \bar{\boldsymbol{x}}^{\mathrm{T}}(t,n_a)\bar{\boldsymbol{a}} - \boldsymbol{u}^{\mathrm{T}}(t,n_b)\boldsymbol{b} \quad (38)$$

where $\bar{\boldsymbol{x}}(t,n_a)$ is chosen such that $\tilde{\bar{\boldsymbol{x}}}(t,n_a)$ is uncorrelated with $\tilde{\bar{\boldsymbol{y}}}(t,n_a)$. Then:

$$J_1(\boldsymbol{\theta}) = \mathrm{E}\{r_1(t,\boldsymbol{\theta})v(t,\boldsymbol{\theta})\} \tag{39}$$

$$= \bar{\boldsymbol{a}}^{\mathrm{T}}\mathrm{E}\Big\{\bar{\tilde{\boldsymbol{x}}}(t,n_a)\bar{\tilde{\boldsymbol{y}}}^{\mathrm{T}}(t,n_a)\Big\}\,\bar{\boldsymbol{a}} + \boldsymbol{b}^{\mathrm{T}}\mathrm{E}\Big\{\tilde{\boldsymbol{u}}(t,n_b)\tilde{\boldsymbol{u}}^{\mathrm{T}}(t,n_b)\Big\}\,\boldsymbol{b} \tag{40}$$

$$= \lambda_{e_{\tilde{u}}}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} \tag{41}$$

If an estimate of $\boldsymbol{b}$ is available, then an estimate of the input noise variance is obtained as:

$$\hat{\lambda}_{e_{\tilde{u}}} = \frac{J_1(\hat{\boldsymbol{\theta}})}{\hat{\boldsymbol{b}}^{\mathrm{T}}\hat{\boldsymbol{b}}} \tag{42}$$

5.2 *The* fbciv *method*

The input noise estimation may be improved by using a filtered version of (42) as suggested in Thil et al. (2008b).

Define a moving average filter $F(q^{-1})$ as:

$$F(q^{-1}) = \sum_{k=-n_f}^{n_f} f_k q^{-k}, \qquad f_0 = 1 \tag{43}$$

Note that in Thil et al. (2008b) this filter was assumed to be causal, while here it is not. A second residual is introduced:

$$r_2(t,\boldsymbol{\theta}) = \bar{\boldsymbol{x}}^{\mathrm{T}}(t,n_a)\bar{\boldsymbol{a}} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{u}_f(t,n_b) \tag{44}$$

where $\boldsymbol{u}_f(t,n_b) = F(q^{-1})\boldsymbol{u}(t,n_b)$. Then, if $\bar{\boldsymbol{x}}(t,n_a)$ is chosen such that $\bar{\tilde{\boldsymbol{x}}}(t,n_a)$ is uncorrelated with $\bar{\tilde{\boldsymbol{y}}}(t,n_a)$:

$$J_2(\boldsymbol{\theta}) = \mathrm{E}\{r_2(t,\boldsymbol{\theta})v(t,\boldsymbol{\theta})\} \tag{45}$$

$$= \bar{\boldsymbol{a}}^{\mathrm{T}}\mathrm{E}\Big\{\bar{\tilde{\boldsymbol{x}}}(t,n_a)\bar{\tilde{\boldsymbol{y}}}^{\mathrm{T}}(t,n_a)\Big\}\,\bar{\boldsymbol{a}} + \boldsymbol{b}^{\mathrm{T}}\mathrm{E}\Big\{\tilde{\boldsymbol{u}}_f(t,n_b)\tilde{\boldsymbol{u}}^{\mathrm{T}}(t,n_b)\Big\}\,\boldsymbol{b} \tag{46}$$

$$= \lambda_{e_{\tilde{u}}}\boldsymbol{b}^{\mathrm{T}}\boldsymbol{M}_F\,\boldsymbol{b} \tag{47}$$

where $\boldsymbol{M}_F$ is the following Tœplitz matrix (assuming that $|n_f| \geqslant n_b - 1$):

$$\boldsymbol{M}_F = \begin{bmatrix} f_0 & f_1 & \cdots & f_{n_b-1} \\ f_{-1} & f_0 & \cdots & f_{n_b-2} \\ \vdots & \ddots & \ddots & \vdots \\ f_{-n_b+1} & \cdots & \cdots & f_0 \end{bmatrix} \tag{48}$$

Hence, if an estimate of $\boldsymbol{b}$ is available, then an estimate of the input noise variance is obtained as:

$$\hat{\lambda}_{e_{\tilde{u}}} = \frac{J_2(\hat{\boldsymbol{\theta}})}{\hat{\boldsymbol{b}}^{\mathrm{T}}\boldsymbol{M}_F\hat{\boldsymbol{b}}} \tag{49}$$

5.3 *The* bciv *algorithm*

1. Initialization: $\hat{\boldsymbol{\theta}}_{\mathsf{bciv}}(0) = \hat{\boldsymbol{\theta}}_{\mathsf{biv}}$
2. Iteration until convergence:
   2.1. Estimation of $\lambda_{e_{\tilde{u}}}$ using (42) or (49)
   2.2. Computation of $\hat{\boldsymbol{\theta}}_{\mathsf{bciv}}(i+1)$

$$\hat{\boldsymbol{\theta}}_{\mathsf{bciv}}(i+1) = \hat{\boldsymbol{\theta}}_{\mathsf{biv}} + \hat{\lambda}_{e_{\tilde{u}}}(i)\hat{\boldsymbol{R}}_{z\varphi}^{-1}\boldsymbol{M}\hat{\boldsymbol{\theta}}_{\mathsf{bciv}}(i)$$

Remark 2

- *Depending on the choice for $\bar{\boldsymbol{x}}(t,n_a)$, equations (42) and (49) will be valid for several delays starting from a minimal value. The algorithms tested in the numerical simulations therefore use several values of the delay to improve the variance estimation.*
- *The algorithms using (42) to obtain the input variance estimate are referred to as* bciv*. The vector $\boldsymbol{x}(t,n_a)$ is chosen such that $\tilde{\boldsymbol{x}}(t,n_a)$ is uncorrelated with $\tilde{\boldsymbol{y}}(t,n_a)$. Two examples of such a vector are:*
  - $\boldsymbol{u}(t-d,n_a)$ *with* $d \geqslant n_b+1$
  - $\boldsymbol{y}(t-d,n_a)$ *with* $d \geqslant n_a+1$

  The corresponding algorithms are denoted $\mathsf{bciv}_{\mathsf{uu}}$ *and* $\mathsf{bciv}_{\mathsf{yu}}$ *respectively.*
- *The algorithms using (49) are denoted* fbciv*, for 'Filtered Bias Compensated Instrumental Variable'.*

5.4 *Consistency and convergence analysis*

The asymptotic convergence of the iterative algorithms is discussed below. In the discussion equation (42) is used, but the developments are valid for equation (49) as well. It should also be noted that the algorithms can diverge when used with a finite number of data (see, *e.g.*, the numerical example section).

When $N = \infty$ and at the ith iteration of the algorithm, equation (42) becomes:

$$\lambda_{e_{\tilde{u}}}^{\star}(i) = \frac{J_1(d,\boldsymbol{\theta}_{\mathsf{bciv}}^{\star}(i))}{\boldsymbol{b}_{\mathsf{bciv}}^{\star,\mathrm{T}}(i)\boldsymbol{b}_{\mathsf{bciv}}^{\star}(i)} \tag{50}$$

$$= \lambda_{e_{\tilde{u}}}\frac{\boldsymbol{b}_{\mathsf{bciv}}^{\star,\mathrm{T}}(i)\boldsymbol{b}_{\mathsf{bciv}}^{\star}(i)}{\boldsymbol{b}_{\mathsf{bciv}}^{\star,\mathrm{T}}(i)\boldsymbol{b}_{\mathsf{bciv}}^{\star}(i)} \tag{51}$$

$$= \lambda_{e_{\tilde{u}}} \tag{52}$$

Thus:

$$\forall i \in \mathbb{N}^*,\ \hat{\lambda}_{e_{\tilde{u}}}(i) \xrightarrow[N\to\infty]{} \lambda_{e_{\tilde{u}}} \tag{53}$$

It is then possible to write:

$$\boldsymbol{\theta}_{\mathsf{bciv}}^{\star}(i+1) = \boldsymbol{\theta}_{\mathsf{biv}}^{\star} + \lambda_{e_{\tilde{u}}}\boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{M}\boldsymbol{\theta}_{\mathsf{bciv}}^{\star}(i) \tag{54}$$

Define $\boldsymbol{K} = \lambda_{e_{\tilde{u}}}\boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{M}$. We have:

$$\boldsymbol{R}_{z\varphi} = \mathrm{E}\big\{\boldsymbol{z}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\big\} \tag{55}$$

$$= \mathrm{E}\big\{\big(\boldsymbol{z_o}(t) + \tilde{\boldsymbol{z}}(t)\big)\big(\boldsymbol{\varphi_o}^{\mathrm{T}}(t) + \tilde{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\big)\big\} \tag{56}$$

$$= \mathrm{E}\big\{\boldsymbol{z_o}(t)\boldsymbol{\varphi_o}^{\mathrm{T}}(t)\big\} + \mathrm{E}\big\{\tilde{\boldsymbol{z}}(t)\tilde{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\big\} \tag{57}$$

$$= \mathrm{E}\big\{\boldsymbol{z_o}(t)\boldsymbol{\varphi_o}^{\mathrm{T}}(t)\big\} + \lambda_{e_{\tilde{u}}}\boldsymbol{M} \tag{58}$$

Therefore, provided the cross-covariance matrix of $\boldsymbol{z_o}$ and $\boldsymbol{\varphi_o}$ is positive definite, the spectral radius of the matrix $\boldsymbol{K}$ is strictly less than unity. It implies that $\boldsymbol{K}$ defines a contraction mapping; the fixed point theorem guarantees the convergence of the iterative scheme (see Banach (1922) for a historical reference, or *e.g.* Ortega and Rheinboldt (2000)).

Lastly, when $i \to \infty$ equation (54) leads to:

$$\boldsymbol{\theta}_{\mathsf{bciv}}^{\star} = \big(\boldsymbol{I} - \lambda_{e_{\tilde{u}}}\boldsymbol{R}_{z\varphi}^{-1}\boldsymbol{M}\big)^{-1}\boldsymbol{\theta}_{\mathsf{biv}}^{\star} = \boldsymbol{\theta} \tag{59}$$

the last equality being obtained using (36).

6. NUMERICAL SIMULATIONS

The goal of the numerical simulations is to compare the methods studied in the previous sections and assess the validity of the discussions. The following system is considered:

$$G_o(q) = \frac{q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \tag{60}$$

The noise-free input is defined as Söderström and Mahata (2002):

$$u_o(t_k) = \frac{1}{1 - 0.2q^{-1} + 0.5q^{-2}}e_{u_o}(t_k) \tag{61}$$

where e_{u_o} is a white noise source following a χ^2 distribution. The input/output noises are white gaussian noises whose variances are set to $\lambda_{e_{\tilde{u}}} = 0.36$ and $\lambda_{e_{\tilde{y}}} = 4$ to obtain a signal-to-noise ratio (SNR) equal to about 15 dB on both input and output, with:

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{P_{x_o}}{P_{\tilde{x}}} \right) \tag{62}$$

where P_x represents the average power of the signal x. The system parameters are estimated on the basis of a data set of length $N = 5000$; a Monte Carlo simulation of $n_{mc} = 200$ runs is performed.

In all the iterative algorithms the same stop criterion is used, that is:

$$\frac{\left\| \hat{\boldsymbol{\theta}}(i+1) - \hat{\boldsymbol{\theta}}(i) \right\|}{\left\| \hat{\boldsymbol{\theta}}(i) \right\|} < 10^{-3} \quad \text{or} \quad i \geqslant 50 \tag{63}$$

where $\| \cdot \|$ is the Euclidian norm and $\hat{\boldsymbol{\theta}}(i)$ the estimate obtained at the ith iteration. Hence, if an algorithm has not converged before 50 iterations, it is automatically stopped. Two different filters are used in the fbciv method. The first is causal while the second is not:

$$F_1(q) = 1 + q^{-1} + q^{-2} \tag{64}$$

$$F_2(q) = q^{1} + 1 + q^{-1} \tag{65}$$

The corresponding algorithms are denoted $\mathsf{fbciv_c}$ and $\mathsf{fbciv_{nc}}$ (for causal and non-causal).

The estimates of the parameter vector and the input noise variance are given in Table 1. Besides, since some methods do not always converge before 50 iterations, the Table 1 also contains 'Conv^{ce}', the percentage of simulations for which the algorithms have converged. Finally, Table 1 also contains the normalized root mean square error, defined as:

$$\mathrm{NRMSE} = \sqrt{\frac{1}{n_{mc}} \sum_{i=1}^{n_{mc}} \frac{\left\| \hat{\boldsymbol{\theta}}(i) - \boldsymbol{\theta}_o \right\|^2}{\left\| \boldsymbol{\theta}_o \right\|^2}} \tag{66}$$

It should be noted that when there are two numbers in the NRMSE column, the first is calculated using all the runs, while for the second only the runs in which the algorithms have converged have been taken into account.

6.1 Discussion

A first observation is that all of the methods give consistent and accurate results, even in an EIV framework – at least on this example.

The methods using delayed input and delayed outputs in the instrument vector are more accurate than those using only delayed inputs, showing the importance of having a good correlation between the instruments and the regressor (compare *e.g.* $\mathsf{xiv_{uu}}$ and $\mathsf{xiv_{yu}}$).

Regarding bias compensated IV methods, $\mathsf{bciv_{uu}}$ and $\mathsf{bciv_{yu}}$ both give good results. However, the fact that they do not always converge is problematic. This is in all likelihood due to the poor estimation of the input noise variance $\lambda_{e_{\tilde{u}}}$. The filtered version, fbciv, greatly increase the quality of this estimation, which in turn leads to no convergence deficiency and accurate parameter estimates. In particular the $\mathsf{fbciv_{nc}}$ method, which uses a non-causal filter, yields the lower NRMSE. An interesting topic for future work is to derive an optimal filter to be used in these algorithms.

A byproduct of this example is also the comparison between second-order and higher-order statistics techniques. It can be seen that the HOS-based methods yield results comparable to the second-order techniques in terms of NRMSE. Remember that the HOS-based methods would still be applicable with no change if the noises were to be coloured, contrary to the second-order methods.

7. CONCLUSION

In this paper, different analytical IV methods for EIV model identification have been reviewed. Three approaches have been considered: the classical extended IV solution using delayed versions of the input/output data as instruments, the bias compensated IV method and higher-order statistics techniques.

Although relaying on simple linear regressions, all of the methods give consistent and accurate results on the example considered in the numerical simulation. This fact gives an idea of the potential of IV techniques, even in an EIV framework. Of course, no definitive conclusion can be drawn from one example only; future work on the subject will include a thorough theoretic study and comparisons between IV and non-IV methods, on several examples.

REFERENCES

S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3:133–181, 1922.

D.R. Brillinger. *Time Series, Data Analysis and Theory.* Holden Day, San Fransisco, 1981.

J.-M. Chen and B.-S. Chen. A higher-order correlation method for model-order and parameter estimation. *Automatica*, 30(8):1339–1344, August 1994.

M. Ekman. Identification of linear systems with errors in variables using separable nonlinear least squares. In *16th IFAC World Congress on Automatic Control*, pages 178–183, Prague, Czech Republic, July 2005.

M. Ekman, M. Hong, and T. Söderström. A separable nonlinear least squares approach for identification of linear systems with errors in variables. In *14th IFAC Symposium on System Identification*, Newcastle, Australia, March 2006.

M. Gilson and P. Van den Hof. Instrumental variable methods for closed-loop system identification. *Automatica*, 41(2):241–249, February 2005.

Y. Inouye and Y. Suga. Identification of linear systems with noisy input using input-output cumulants. *International Journal of Control*, 59(5):1231–1253, May 1994.

Y. Inouye and H. Tsuchiya. Identification of linear systems using input-output cumulants. *International Journal of Control*, 53(6):1431–1448, June 1991.

L. Ljung. *System Identification: Theory for the User.* Prentice Hall, Englewood Cliffs, 2nd edition, 1999.

J.M. Mendel. Tutorial on high-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79(3): 278–305, March 1991.

J.M. Ortega and W.C. Rheinboldt. *Iterative solution of nonlinear equations in several variables.* Society for

Table 1. Monte Carlo simulation results. NRMSE is calculated using all the runs. If applicable, the NRMSE between parentheses is calculated using only the runs for which convergence has occurred.

	$a_1 = -1.5$	$a_2 = 0.7$	$b_1 = 1.0$	$b_2 = 0.5$	$\lambda_{e_{\bar{u}}} = 0.36$	$\mathrm{Conv}^{\mathrm{ce}}$	NRMSE
$\mathsf{xiv}_{\mathsf{uu}}$	-1.4976 ± 0.0001	0.6978 ± 0.0001	0.9939 ± 0.0018	0.5010 ± 0.0009	–	–	0.304
$\mathsf{xiv}_{\mathsf{yu}}$	-1.4997 ± 0.0001	0.6996 ± 0.0001	0.9976 ± 0.0010	0.5004 ± 0.0007	–	–	0.243
$\mathsf{bciv}_{\mathsf{uu}}$	-1.5011 ± 0.0006	0.6998 ± 0.0007	0.9977 ± 0.0006	0.5012 ± 0.0008	0.7654 ± 0.6141	73.5%	0.624(0.238)
$\mathsf{bciv}_{\mathsf{yu}}$	-1.5018 ± 0.0002	0.6997 ± 0.0001	1.0001 ± 0.0004	0.4991 ± 0.0002	0.7409 ± 0.5926	77.5%	0.368(0.147)
$\mathsf{fbciv}_{\mathsf{c}}$	-1.4983 ± 0.0001	0.6995 ± 0.0001	1.0068 ± 0.0012	0.5040 ± 0.0004	0.4555 ± 0.1212	100%	0.239
$\mathsf{fbciv}_{\mathsf{nc}}$	-1.4999 ± 0.0001	0.6996 ± 0.0001	0.9963 ± 0.0008	0.4989 ± 0.0003	0.3251 ± 0.0604	100%	0.202
$\mathsf{tociv}_{\mathsf{uuy}}$	-1.4897 ± 0.0001	0.6900 ± 0.0001	0.9938 ± 0.0008	0.5092 ± 0.0010	–	–	0.269
$\mathsf{tociv}_{\mathsf{yyy}}$	-1.4941 ± 0.0002	0.6956 ± 0.0001	0.9722 ± 0.0037	0.5040 ± 0.0010	–	–	0.416

Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

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

T. Söderström and K. Mahata. On instrumental variable and total least squares approaches for identification of noisy systems. *International Journal of Control*, 75(6): 381–389, April 2002.

T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer, Berlin, 1983.

S. Thil, H. Garnier, and M. Gilson. Third–order cumulants based methods for continuous–time errors–in–variables model identification. *Automatica*, 44(3):647–658, March 2008a.

S. Thil, M. Gilson, and H. Garnier. On instrumental variable methods for errors-in-variables model identification. In *17th IFAC World Congress on Automatic Control*, Seoul, Korea, 2008b.

S. Thil, W.X. Zheng, H. Garnier, and M. Gilson. Unifying some higher-order statistic-based methods for errors-in-variables model identification. *Automatica*, 45(8):1937–1942, August 2009.

S. Van Huffel. *Total least squares and errors-in-variables modeling*. Kluwer Academic Publishers, Dordrecht, 2002.

Z.-J. Yang, S. Sagara, and K. Wada. Identification of continuous-time systems from sampled input-output data using bias eliminating techniques. *Control Theory and Advanced Technology*, 9(1):53–75, March 1993.

P.C. Young. An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6(2): 271–287, March 1970.

P.C. Young. *Recursive Estimation and Time Series Analysis*. Springer-Verlag, Berlin, 1984.

P.C. Young, H. Garnier, and M. Gilson. Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models. In H. Garnier and L. Wang, editors, *Identification of Continuous-time Models from Sampled Data*. Springer-Verlag, London, 2008.