

Unifying some higher-order statistic-based methods for errors-in-variables model identification [★]

Stéphane Thil ^a, Wei Xing Zheng ^b, Marion Gilson ^c, Hugues Garnier ^c

^aLaboratoire ELIAUS, Université de Perpignan Via Domitia, 52 avenue Paul Alduy, 66860 Perpignan, France.
E-mail address: stephane.thil@univ-perp.fr

^bSchool of Computing and Mathematics, University of Western Sydney, Perish South DC NW 1797, Australia.
E-mail address: w.zheng@uws.edu.au

^cCentre de Recherche en Automatique de Nancy (CRAN UMR 7039),
Nancy-Université, CNRS, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France.
E-mail addresses: givenname.familyname@cran.uhp-nancy.fr

Abstract

In this paper, the problem of identifying linear discrete-time systems from noisy input and output data is addressed. Several existing methods based on higher-order statistics are presented. It is shown that they stem from the same set of equations and can thus be united from the viewpoint of extended instrumental variable methods. A numerical example is presented which confirms the theoretical results. Some possible extensions of the methods are then given.

Key words: System identification; errors-in-variables; instrumental variable; higher-order statistics.

1 Introduction

Identification of errors-in-variables (EIV) models has been a very active domain of research in the past few years (see *e.g.* (Mahata and Garnier 2006, Mahata 2007, Diversi *et al.* 2007, Hong *et al.* 2007, Pintelon and Schoukens 2007, Thil *et al.* 2007, Söderström 2008, Thil *et al.* 2008b)), and a survey paper gathering most of the known developments has been recently published (Söderström 2007).

Most of the research has been concerned with estimating the parameters of discrete-time EIV models with the help of second-order statistics. Nonetheless, a recently published paper (Thil *et al.* 2008a) has shown that continuous-time EIV model identification can be successfully handled using higher-order statistics (HOS). Although much work has been conducted in the HOS field for EIV model identification in the 90's, it seems that several questions concerning the practical use of HOS for system identification remain to be answered.

The aim of this paper is to present some HOS-based

[★] This work was supported by a research grant from the Australian Research Council.

methods for EIV model identification in a unified way. More precisely, the links between the methods developed in (Inouye and Tsuchiya 1991, Chen and Chen 1994) and the discrete-time version of an algorithm presented in (Thil *et al.* 2008a) are explored. It is shown that these methods stem from the same set of equations. Simulation results support the theoretical analysis, and some possible extensions for future work are given.

2 Errors-in-variables framework

Consider a discrete-time, linear, time-invariant EIV system. The noise-free input/output signals are related by

$$y_o(t) = G_o(q)u_o(t) \quad (1)$$

where q is the forward operator and $G_o(\cdot)$ is the transfer operator of the 'true' system. The input and output signals are both contaminated by noise sequences, denoted as \tilde{u} and \tilde{y} , respectively. The data-generating system is thus given by

$$\begin{cases} y_o(t) = G_o(q)u_o(t) \\ u(t) = u_o(t) + \tilde{u}(t) \\ y(t) = y_o(t) + \tilde{y}(t) \end{cases} \quad (2)$$

It is then parameterized as follows:

$$\begin{cases} y(t) = G(q, \boldsymbol{\theta})(u(t) - \tilde{u}(t)) + \tilde{y}(t) \\ G(q, \boldsymbol{\theta}) = B(q^{-1}, \boldsymbol{\theta})/A(q^{-1}, \boldsymbol{\theta}) \\ A(q^{-1}, \boldsymbol{\theta}) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \\ B(q^{-1}, \boldsymbol{\theta}) = b_0 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \end{cases} \quad (3)$$

with $n_a \geq n_b$ and $\boldsymbol{\theta}^T = [a_1 \dots a_{n_a} b_0 \dots b_{n_b}]$. Equation (3) can be rewritten as

$$\begin{aligned} y(t) &= \boldsymbol{\varphi}^T(t) \boldsymbol{\theta} + v(t, \boldsymbol{\theta}) \\ v(t, \boldsymbol{\theta}) &= \tilde{y}(t) - \tilde{\boldsymbol{\varphi}}^T(t) \boldsymbol{\theta} \end{aligned} \quad (4)$$

where the regression vector is given by

$$\boldsymbol{\varphi}^T(t) = \begin{bmatrix} -y(t-1) & \dots & -y(t-n_a) \\ u(t) & \dots & u(t-n_b) \end{bmatrix} \quad (6)$$

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 input/output data $\{u(t), y(t)\}_{t=1}^N$.

2.1 Notations

As the input and output noises are additive, linear functions of the measured signals can be broken down into two parts: one part made up of the noise-free signals' contribution (denoted with an 'o' subscript) and the other made up of the noises' contribution (denoted with the '~' sign). For example, the regression vector $\boldsymbol{\varphi}$ can be decomposed into

$$\boldsymbol{\varphi}(t) = \boldsymbol{\varphi}_o(t) + \tilde{\boldsymbol{\varphi}}(t) \quad (7)$$

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

$$\mathbf{R}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} = \bar{E}\{\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^T(t)\}, \quad \mathbf{r}_{\boldsymbol{\varphi}y} = \bar{E}\{\boldsymbol{\varphi}(t)y(t)\} \quad (8)$$

where $\bar{E}\{\cdot\}$ stands for (see (Ljung 1999))

$$\bar{E}\{f(t)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E\{f(t)\} \quad (9)$$

The notation used for the third-order cumulants is

$$\begin{aligned} C_{x_1 x_2 x_3}(\tau_1, \tau_2) &= \bar{E}\{x_1(t)x_2(t+\tau_1)x_3(t+\tau_2)\} \\ &= C_{x_1 x_2 x_3}(\boldsymbol{\tau}) \end{aligned} \quad (10)$$

where, for the sake of conciseness, $\boldsymbol{\tau}$ denotes $[\tau_1, \tau_2]$.

2.2 Assumptions and elements of structure

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 .

For methods based on second-order statistics to give unbiased estimates, it is usually assumed (and often implicitly) that the 'true' system belongs to the considered model set, a situation referred to as $\mathcal{S} \in \mathcal{M}^*$ (Ljung 1999). However, this notation has been introduced for systems with noise-free inputs, and – being too general – is not properly suited for errors-in-variables models. Indeed, more often than not, the input and output noises and the noise-free input must be modeled. Thus, some additional notations must be introduced. The whole 'true' system includes

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

where $\boldsymbol{\eta}$ is a vector gathering the parameters of noises models and noise-free input models.

In addition to assuming that the 'true' process belongs to the model set, *i.e.*, $G_o \in \mathcal{G}^*$, most methods based on second-order statistics require that the noises' models belong to the model set, *i.e.*, $(H_o^{\tilde{u}}, H_o^{\tilde{y}}) \in \mathcal{H}^*$. A few even require that the noise-free input is adequately modeled, and thus that $H_o^{u_o} \in \mathcal{E}^*$. For example, the maximum likelihood and prediction error methods require such assumptions (Söderström 1981, Söderström 2007).

On the contrary, methods based on higher-order statistics do not require *structural assumptions* on the input and output noises \tilde{u} , \tilde{y} , and on the noise-free input u_o . The only structural assumption needed is

- A4. The true process belongs to the model set: $G_o \in \mathcal{G}^*$.

The input and output noises can thus be arbitrarily coloured (and even mutually correlated), and there is no structural assumption on the noise-free input. However, for the higher-order cumulants of the noises to be zero and for the higher-order cumulants of the noise-free input not to be zero, *distributional assumptions* are needed. These distributional assumptions differ whether third- or fourth-order cumulants are used. For the third-order cumulants,

- A5a. The input and output noises \tilde{u} , \tilde{y} have symmetric probability density functions (pdf's),
A6a. The noise-free input u_o has a skewed pdf.

For the fourth-order cumulants,

- A5b. The input and output noises \tilde{u} , \tilde{y} have Gaussian pdf's,
A6b. The noise-free input u_o has a non-Gaussian pdf.

In the sequel, the focus will be placed on the case of third-order cumulants, and consequently the assumptions A1-A4 and A5a-A6a are supposed to be satisfied.

3 Identification methods using HOS

3.1 Properties of HOS

The identification techniques presented in this paper are based on higher-order statistics (see *e.g.* (Brillinger 1981, Mendel 1991)). Here we recall a few of the numerous properties of higher-order cumulants.

- 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, A5a, A6a and using properties P2, P3, the following holds

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

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.

3.2 Higher-order statistic-based LS method

A simple least-squares (LS) method based on higher-order statistics has been introduced in (Thil *et al.* 2008a) for continuous-time EIV model identification, and its statistical analysis has been conducted in (Thil *et al.* 2008c). Here we give a brief overview of the method.

The EIV model can be shown to satisfy the following third-order cumulants equation

$$C_{uuy}(\boldsymbol{\tau}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uuu}(\boldsymbol{\tau}) \quad (12)$$

which can be expressed in a linear regression form as

$$C_{uuy}(\boldsymbol{\tau}) = \boldsymbol{\varphi}_{uuy}^T(\boldsymbol{\tau}) \boldsymbol{\theta} \quad (13)$$

where

$$\boldsymbol{\varphi}_{uuy}^T(\boldsymbol{\tau}) = [-C_{uuy}(\tau_1, \tau_2 - 1) \dots - C_{uuy}(\tau_1, \tau_2 - n_a) \\ C_{uuu}(\tau_1, \tau_2) \dots C_{uuu}(\tau_1, \tau_2 - n_b)] \quad (14)$$

Due to the stationarity of the signals, the linear regression model (13) does not depend on the time t , but on τ_1 and τ_2 , the time lags appearing in the third-order cumulants. Should all that information be used? If not, what part should be chosen? These questions must be answered with the properties of cumulants in mind (symmetries), as well as the problem of their estimation (larger time lags imply less accuracy). In (Thil *et al.* 2008a), the chosen cumulant slice was $\tau_1 = 0$ and $0 \leq \tau_2 \leq M - 1$; and the purpose of the user-selected parameter M is to avoid the use of third-order cumulants with large time-lags. The reason for this is that as the time-lag increases, the cumulants estimates are calculated from less and less data, thus being less and less reliable (Thil *et al.* 2008a). It yields

$$\begin{bmatrix} C_{uuy}(0, 0) \\ \vdots \\ C_{uuy}(0, M - 1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varphi}_{uuy}^T(0, 0) \\ \vdots \\ \boldsymbol{\varphi}_{uuy}^T(0, M - 1) \end{bmatrix} \boldsymbol{\theta} \quad (15)$$

which may be written in a compact form as

$$\mathbf{C} = \boldsymbol{\Phi}^T \boldsymbol{\theta} \quad (16)$$

An estimate of the parameter vector $\boldsymbol{\theta}$ can then be obtained by solving the system of equations (16) in a least squares sense, yielding the so-called third-order cumulant-based LS (called **tocls** for short) estimator

$$\boldsymbol{\theta}_{\text{tocls}}^* = (\boldsymbol{\Phi} \boldsymbol{\Phi}^T)^{-1} \boldsymbol{\Phi} \mathbf{C} \quad (17)$$

It should also be noted that, instead of (12), the following third-order cumulants equations can also be used:

$$C_{uyy}(\boldsymbol{\tau}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uyu}(\boldsymbol{\tau}) \quad (18)$$

$$C_{yyy}(\boldsymbol{\tau}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{yyu}(\boldsymbol{\tau}) \quad (19)$$

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

3.3 Higher-order statistic-based IV method

Another solution is to use an instrumental variable (IV) approach instead of a least squares method to handle the parameter estimation. Indeed, the classical IV approach can be modified in order to use third-order cumulants. Here we briefly recall the higher-order IV method introduced in (Inouye and Tsuchiya 1991).

The starting point is the model equation written in a linear regression form (4)-(5). Define an instrumental vector as

$$\mathbf{z}^T(t) = \begin{bmatrix} -y^2(t-1) & \dots & -y^2(t-n_a) \\ u^2(t) & \dots & u^2(t-n_b) \end{bmatrix} \quad (20)$$

that is, a vector containing the square of each element of the regression vector $\varphi(t)$. Now multiplying (4) by $\mathbf{z}(t)$ and taking the expectation of the result yields

$$\mathbf{r}_{zy} = \mathbf{R}_{z\varphi}\boldsymbol{\theta} + \mathbf{r}_{zv} \quad (21)$$

where $\mathbf{R}_{z\varphi} = \bar{E}\{\mathbf{z}(t)\varphi^T(t)\}$. The vectors and matrices appearing in (21) contain third-order cumulants. Taking a closer look at \mathbf{r}_{zv} , we can see that

$$\mathbf{r}_{zv} = \bar{E}\{\mathbf{z}(t)v(t, \boldsymbol{\theta})\} = \bar{E}\{\tilde{\mathbf{z}}(t)v(t, \boldsymbol{\theta})\} = \mathbf{0} \quad (22)$$

Here we first used the fact that the noise-free input signal u_o is uncorrelated with the input and output noises \tilde{u} and \tilde{y} [A3], and then the fact that the noises have a symmetric distribution [A5a]. The parameter vector $\boldsymbol{\theta}$ can thus be estimated using (21). Moreover, if the input signal u_o is third-order white, the matrix $\mathbf{R}_{z\varphi}$ is generically non-singular (Inouye and Tsuchiya 1991). An IV estimator based on third-order cumulants is therefore obtained as

$$\boldsymbol{\theta}_{\text{tociv}}^* = \mathbf{R}_{z\varphi}^{-1}\mathbf{r}_{zy} \quad (23)$$

where **tociv** stands for *Third-Order Cumulant-based Instrumental Variable*.

3.4 Higher-order correlation method

In (Chen and Chen 1994) a higher-order correlation method (**hocm**) is described to estimate the order and the parameters of an EIV model. It is striking to note that, in that paper, the parameters are estimated in the exact same way as in the **tociv** method presented in (Inouye and Tsuchiya 1991), recalled in the preceding subsection. Indeed, although the presentation differs, their parameter estimate is exactly given by (23), the instrumental vector being identical in both papers.

4 A unifying analysis

As stated before, the **tocls** method originates from the following equations

$$C_{uuy}(\boldsymbol{\tau}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uuu}(\boldsymbol{\tau}) \quad (24)$$

$$C_{uyy}(\boldsymbol{\tau}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{uyu}(\boldsymbol{\tau}) \quad (25)$$

$$C_{yyy}(\boldsymbol{\tau}) = \frac{B(q^{-1}, \boldsymbol{\theta})}{A(q^{-1}, \boldsymbol{\theta})} C_{yyu}(\boldsymbol{\tau}) \quad (26)$$

Now we take a closer look at equation (21), which forms the basis of the **tociv** method. It can be noticed that it may be broken down into two parts (see (29)):

- (1) the first n_a lines stem from the linear regression based on (26), with $\tau_1 = 0$ and $1 \leq \tau_2 \leq n_a$;
- (2) the other $n_b + 1$ lines stem from the linear regression based on (24), with $\tau_1 = 0$ and $0 \leq \tau_2 \leq n_b$.

Note that in (29) the stationarity of the involved signals is necessary, so this result is only true for $N \rightarrow \infty$. The **tociv** method of (Inouye and Tsuchiya 1991) is thus closely linked to the **tocls** method in terms of using equations (24) and (26), and with an implicit choice of the cumulant slice.

Inversely, equation (16) of the **tocls** method can be written as

$$\mathbf{r}_{zy} = \mathbf{R}_{z\varphi}\boldsymbol{\theta} \quad (27)$$

with

$$\mathbf{z}^T = [u^2(t) \ u^2(t-1) \ \dots \ u^2(t-M+1)] \quad (28)$$

The **tocls** method can thus be viewed as an extended version of the **tociv** method (called **tocxiv** for short), since the instrumental vector (28) is of dimension M , *i.e.*, larger than the dimension of $\boldsymbol{\theta}$.

The three methods presented in the preceding section use the same set of equations to estimate the parameter vector. The only differences between them are which equations (or combination of equations) are used, and how many of those equations are used.

5 Numerical example

The objective of this section is to assess the results of Section 4. To this end, several variants of those algorithms are used to estimate the parameters of the two

$$\begin{bmatrix} -C_{yyyy}(0,1) \\ \vdots \\ -C_{yyyy}(0,n_a) \\ C_{uuuy}(0,0) \\ \vdots \\ C_{uuuy}(0,n_b) \end{bmatrix} = \begin{bmatrix} C_{yyyy}(0,0) & \cdots & C_{yyyy}(0,1-n_a) & -C_{yyyu}(0,1) & \cdots & -C_{yyyu}(0,1-n_b) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ C_{yyyy}(0,n_a-1) & \cdots & C_{yyyy}(0,0) & -C_{yyyu}(0,n_a) & \cdots & -C_{yyyu}(0,n_a-n_b) \\ -C_{uuuy}(0,-1) & \cdots & -C_{uuuy}(0,-n_a) & C_{uuuu}(0,0) & \cdots & C_{uuuu}(0,-n_b) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -C_{uuuy}(0,n_b-1) & \cdots & -C_{uuuy}(0,n_b-n_a) & C_{uuuu}(0,n_b) & \cdots & C_{uuuu}(0,0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{n_a} \\ b_0 \\ \vdots \\ b_{n_b} \end{bmatrix} \quad (29)$$

following systems, that have different properties

$$\text{System 1: } G_o(q) = \frac{1q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \quad (30)$$

$$\text{System 2: } G_o(q) = \frac{0.5 - 0.25q^{-1} + 0.15q^{-2}}{1 + 0.5q^{-1} - 0.125q^{-2}} \quad (31)$$

The Bode diagrams of these systems are given on Figure 1. While the system 1 is commonly used as a bench-

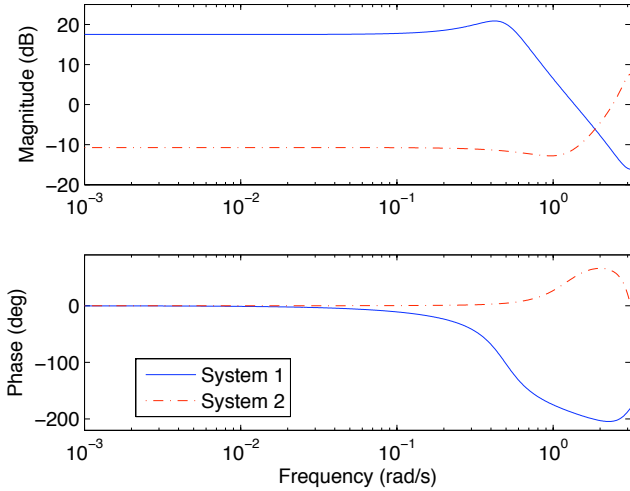


Fig. 1. Bode diagrams of System 1 and System 2.

mark (it exhibits properties one would expect from an adequately tuned second-order system), the second system is a bit less ‘ideal’. The simulation conditions are the same for both systems. The noise-free input is obtained as the output of

$$H_o^{u_o}(q) = \frac{1}{1 - 0.2q^{-1} + 0.5q^{-2}} \quad (32)$$

which is excited by a white exponential noise. The input and output noises are defined as the outputs of the following MA filters

$$H_o^{\tilde{u}}(q) = 1 - 2.33q^{-1} + 0.75q^{-2} + 0.5q^{-3} + 0.3q^{-4} - 1.4q^{-5}$$

$$H_o^{\tilde{y}}(q) = 1 + 0.2q^{-1} - 0.3q^{-2} + 0.4q^{-3}$$

which are excited by white uniform noises. The variances of these white noises are then set to obtain a signal-to-noise ratio equal to about 10 dB on both input and output of each system. The length of the data set is $N = 5000$. The results of Monte Carlo simulations of $n_{mc} = 200$ runs are displayed in Table 1 and Table 2 respectively for the two systems, where the mean and standard deviation of the estimates, together with the normalized root mean square error (NRMSE) on the parameter estimates, are given, where

$$\text{NRMSE} = \sqrt{\frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} \frac{\|\hat{\theta}_j - \theta\|^2}{\|\theta\|^2}} \quad (33)$$

The algorithms appearing in Tables 1-2 are ¹

- `tociv` and `hocm` are the – identical – methods developed in (Inouye and Tsuchiya 1991) and (Chen and Chen 1994), respectively;
- `tocxiv` is an extension of `tociv`, using an instrumental vector of dimension M , larger than θ , and solving the over-determined system in a least-squares sense;
- the ‘variants’ of the algorithms are linked to the equations used to estimate the parameter vector. Those labelled as ‘u only’ use only (24), while those labelled as ‘u&y’ use (24) and (26).

Discussion

Although their actual implementations are different, the `tociv` and `hocm` methods give exactly the same results, which confirms that these two methods are identical. Like the basic instrumental variable methods using second-order statistics, these methods are prone to numerical instability when the matrix to be inverted is ill-conditioned; see especially the results of `tociv` obtained for the second system. However, their performances are improved by allowing the instrumental vector to be of a dimension larger than θ (compare the results of `tociv` and `tocxiv`, especially for the second system in Table 2).

¹ It should be noted that algorithms using different combinations of (24)–(26) have been tested. Since no version appeared to be uniformly best, only the results of the compared methods have been included.

	Variant	M	$a_1 = -1.5$	$a_2 = 0.7$	$b_0 = 0$	$b_1 = 1.0$	$b_2 = 0.5$	NRMSE
hocm		–	Same as <i>tociv</i> below					
<i>tociv</i>	u&y	–	-1.5048 ± 0.0041	0.7052 ± 0.0046	-0.0069 ± 0.0027	1.0071 ± 0.0033	0.4891 ± 0.0123	0.91
<i>tocxiv</i>	u&y	30	-1.4948 ± 0.0002	0.6959 ± 0.0001	0.0038 ± 0.0092	0.9878 ± 0.0080	0.5029 ± 0.0069	0.88
	u only	30	-1.4852 ± 0.0003	0.6859 ± 0.0002	-0.0053 ± 0.0024	0.9983 ± 0.0026	0.5101 ± 0.0039	0.57
<i>tocls</i>	u&y	30	-1.4948 ± 0.0001	0.6958 ± 0.0001	0.0055 ± 0.0083	0.9791 ± 0.0093	0.5099 ± 0.0042	0.84
	u only	30	-1.4847 ± 0.0003	0.6853 ± 0.0002	-0.0055 ± 0.0024	0.9988 ± 0.0026	0.5099 ± 0.0039	0.57

Table 1. Monte Carlo simulation results for System 1 (see (30)).

	Variant	M	$a_1 = 0.5$	$a_2 = -0.125$	$b_0 = 0.5$	$b_1 = -0.25$	$b_2 = 0.15$	NRMSE
hocm		–	Same as <i>tociv</i> below					
<i>tociv</i>	u&y	–	0.6148 ± 0.3565	-0.0191 ± 0.2745	0.4998 ± 0.0001	-0.1936 ± 0.0886	0.1481 ± 0.0008	6.31
<i>tocxiv</i>	u&y	30	0.5087 ± 0.0041	-0.1191 ± 0.0031	0.5005 ± 0.0001	-0.2462 ± 0.0010	0.1494 ± 0.0001	0.87
	u only	30	0.5160 ± 0.0042	-0.1136 ± 0.0032	0.4995 ± 0.0001	-0.2420 ± 0.0010	0.1477 ± 0.0001	0.90
<i>tocls</i>	u&y	30	0.5092 ± 0.0041	-0.1189 ± 0.0031	0.5003 ± 0.0001	-0.2461 ± 0.0009	0.1493 ± 0.0001	0.86
	u only	30	0.5168 ± 0.0042	-0.1132 ± 0.0032	0.4993 ± 0.0001	-0.2418 ± 0.0010	0.1476 ± 0.0001	0.90

Table 2. Monte Carlo simulation results for System 2 (see (31)).

Regarding the comparison of the *tocxiv* and *tocls* methods, since the dimension of the instrumental vector in *tocxiv* and the number of additional equations in *tocls* are chosen to be the same, the ‘u only’ variant of these two methods should give the same results, just as the ‘u&y’ variant. Their results are indeed quite similar, especially for the ‘u only’ variant. However, slight differences may be noticed between them. That can be explained by the way the estimation is conducted. In the case of the *tocls* method, the third-order cumulants are first estimated, and then the parameter vector is estimated using those estimates. In the *tocxiv* method, both the third-order cumulant estimation and the parameter vector estimation are conducted in one step. This leads the third-order cumulant estimates to be less accurate since they are not obtained from the maximum amount of available data. Also, none of the ‘u&y’ and the ‘u only’ variants is uniformly best. Depending on the system and on the signals, they alternatively give the best results.

6 Conclusion and perspectives

This paper has been concerned with exhibiting the links between the parameter estimation methods introduced in (Inouye and Tsuchiya 1991, Chen and Chen 1994, Thil *et al.* 2008a). It has been shown that these higher-order statistic-based methods are based on the same set of equations, and thus can be interpreted as extended instrumental variable methods. The numerical simulations have confirmed these results. It should also be noted that the same analysis can be conducted for the methods using fourth-order cumulants (see (Inouye and Suga 1994) for the instrumental variable method and (Thil *et*

al. 2007) for the least-squares method).

Now that the similarities between the methods have been pointed out, future work will investigate the ‘classical extensions’ of IV estimators, such as weighting and filtering. It would also be interesting to explore how asymptotic results derived for second-order IV methods can be adapted to the higher-order case.

References

- Brillinger, D.R. (1981). *Time Series, Data Analysis and Theory*. Holden Day. San Francisco.
- Chen, J.-M. and B.-S. Chen (1994). A higher-order correlation method for model-order and parameter estimation. *Automatica* **30**(8), 1339–1344.
- Diversi, R., R. Guidorzi and U. Soverini (2007). Maximum likelihood identification of noisy input-output models. *Automatica* **43**(2), 464–472.
- Hong, M., T. Söderström and W.X. Zheng (2007). Accuracy analysis of bias-eliminating least squares estimates for errors-in-variables identification. *Automatica* **43**(9), 1590–1596.
- Inouye, Y. and H. Tsuchiya (1991). Identification of linear systems using input-output cumulants. *International Journal of Control* **53**(6), 1431–1448.
- Inouye, Y. and Y. Suga (1994). Identification of linear systems with noisy input using input-output cumulants. *International Journal of Control* **59**(5), 1231–1253.
- Ljung, L. (1999). *System Identification: Theory for the User*. 2nd ed.. Prentice Hall. Englewood Cliffs.
- Mahata, K. (2007). An improved bias-compensation approach for errors-in-variables model identification. *Automatica* **43**(8), 1339–1354.

- Mahata, K. and H. Garnier (2006). Identification of continuous-time errors-in-variables models. *Automatica* **49**(9), 1470–1490.
- Mendel, J.M. (1991). 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.
- Pintelon, R. and J. Schoukens (2007). Frequency domain maximum likelihood estimation of linear dynamic errors-in-variables models. *Automatica* **43**(4), 621–630.
- Söderström, T. (1981). Identification of stochastic linear systems in presence of input noise. *Automatica* **17**(5), 713–725.
- Söderström, T. (2007). Errors-in-variables methods in system identification. *Automatica* **43**(6), 939–958.
- Söderström, T. (2008). Extending the Frisch scheme for errors-in-variables identification to correlated output noise. *International Journal of Adaptive Control and Signal Processing* **22**(1), 55–73.
- Thil, S., H. Garnier and M. Gilson (2008a). Third-order cumulants based methods for continuous-time errors-in-variables model identification. *Automatica* **44**(3), 647–658.
- Thil, S., H. Garnier, M. Gilson and K. Mahata (2007). Continuous-time model identification from noisy input/output measurements using fourth-order cumulants. In: *46th IEEE Conference on Decision and Control (CDC'2007)*. New Orleans, LA, USA.
- Thil, S., M. Gilson and H. Garnier (2008b). On instrumental variable methods for errors-in-variables model identification. In: *17th IFAC World Congress*. Seoul, Korea.
- Thil, S., M. Hong, T. Söderström, M. Gilson and H. Garnier (2008c). Statistical analysis of a third-order cumulants based algorithm for discrete-time errors-in-variables identification. In: *17th IFAC World Congress*. Seoul, Korea.