IFAC

Statistical Analysis of a Third-Order Cumulants Based Algorithm for Discrete-Time Errors-in-Variables Identification [1]

Stéphane Thil [*], **Mei Hong** [**], **Torsten Söderström** [**],
Marion Gilson [*], **Hugues Garnier** [*]

[*] *Centre de Recherche en Automatique de Nancy
(CRAN UMR 7039), Nancy-Université, CNRS,
BP 239, 54506, Vandœuvre-lès-Nancy Cedex, France.*
[**] *Dept. of Information Technology, Uppsala University, P.O. Box 337,
SE 75105 Uppsala, Sweden.*

Abstract: This paper deals with identification of dynamic discrete-time errors-in-variables systems. The statistical accuracy of a least squares estimator based on third-order cumulants is analyzed. In particular, the asymptotic covariance matrix of the estimated parameters is derived. The results are supported by numerical simulation studies.

Keywords: System identification, errors-in-variables, higher-order statistics, accuracy analysis.

1. INTRODUCTION

Consider the linear time-invariant errors-in-variables (EIV) system represented in Figure 1. The noise-free input and output signals are related by:

$$y_0(t_k) = G_0(q)u_0(t_k) \tag{1}$$

where q is the forward operator and $G_0(q)$ is the transfer function of the 'true' system. The input/output signals are both contaminated by noise sequences, denoted as $\tilde{u}$ and $\tilde{y}$ respectively.
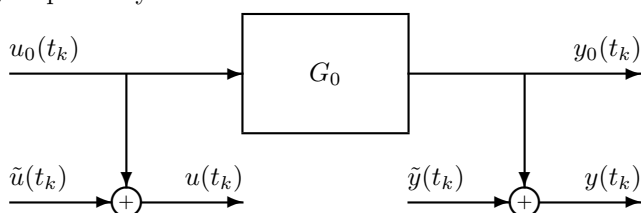

Fig. 1. Discrete-time EIV model

The data-generating system is thus characterized by:

$$\mathcal{S}: \begin{cases} y_0(t_k) = G_0(q)u_0(t_k) \\ u(t_k) = u_0(t_k) + \tilde{u}(t_k) \\ y(t_k) = y_0(t_k) + \tilde{y}(t_k) \end{cases} \tag{2}$$

It is then parameterized as follows:

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

where $\boldsymbol{\theta} = [a_1 \ \dots \ a_{n_a} \ b_0 \ \dots \ b_{n_b}]^T$. The problem of identifying this errors-in-variables system is concerned with consistently estimating the parameter vector $\boldsymbol{\theta}$ from the measured noisy data $\{u(t_k), y(t_k)\}_{k=1}^N$.

It is well known Anderson and Deistler [1984], Scherrer and Deistler [1998] that if only the second-order statistics are exploited and without introducing more explicit assumptions, a unique solution to identification of the errors-in-variables systems does not exist. It is thus natural to study alternative methods based on higher-order statistics (HOS). Under different assumptions on the distributional properties of the noises and noise-free signals, we could remove the effect of the disturbances from the measured data to get consistent estimates. Several identification methods using higher-order statistics have been proposed for discrete-time EIV systems, for example, in Friedlander and Porat [1990], Tugnait [1992], Delopoulos and Giannakis [1994] and Tugnait and Ye [1995]. Recently a new HOS-based method for identification of continuous-time EIV models has been proposed, using the fact that the equation of the model is satisfied by the third-order cumulants. For details, see Thil et al. [2008] and references therein.

Here we are interested in deriving the asymptotic properties of (the discrete-time equivalent of) the third-order cumulants-based estimator proposed in that paper. An explicit expression for the asymptotic covariance matrix of the estimated parameters is derived. To apply the identification method, the following assumptions are needed:

A1. $G_0(z)$ is asymptotically stable and there is no common factor between $A_0(z)$ and $B_0(z)$;

A2. The polynomial degrees n_a and n_b are *a priori* known;

A3. The noise-free signal u_0 is a zero-mean, stationary, ergodic stochastic process such that its third-order cumulants are non-zero. Its probability density function (pdf) is thus not symmetric;

A4. $\tilde{u}$ and $\tilde{y}$ are zero-mean, stationary, ergodic mutually independent stochastic processes [2] with symmetric pdf's, and both are independent of u_0;

A5. The moments of second, fourth- and sixth-order of the noise-free input u_0 and the noises $\tilde{u}$, $\tilde{y}$ decay exponentially, that is $E\{x(t_k)x(t_{k+\tau_1})\ldots x(t_{k+\tau_p})\}$ decays exponentially as $\max_j |\tau_j| \to \infty$.

The outline of the paper is as follows. In the next section we recall the main definitions and properties of the third-order cumulants, and the identification method is quickly introduced. In Section 3 the asymptotic covariance matrix of the estimator is derived. Finally, before concluding in Section 5, a simulation example is given in Section 4 to assess the agreement between the theoretical and the simulated results.

2. THIRD-ORDER CUMULANTS BASED ESTIMATOR

2.1 Properties of HOS

The identification technique developed in this paper is based on higher-order statistics Brillinger [1981], Mendel [1991], Lacoume et al. [1997]. The main definitions and properties used in the proposed approach are recalled in this section.

The third-order cumulant of a real-valued, zero-mean stationary random process x is defined as:

$$C_{xxx}(\tau_1, \tau_2) = Cum[x(t_k), x(t_{k+\tau_1}), x(t_{k+\tau_2})]$$
$$= E\{x(t_k)x(t_{k+\tau_1})x(t_{k+\tau_2})\}$$

The cumulants of order higher than two have many properties amongst which we only recall those used in the proposed estimation scheme.

Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ be two random vectors.

P1. Multilinearity: the cumulants are linear with respect to each of their arguments. If α_i, β_j, γ_k are scalars:

$$Cum\left[\sum_i \alpha_i x(t_i), \sum_j \beta_j x(t_j), \sum_k \gamma_k x(t_k)\right]$$
$$= \sum_{i,j,k} \alpha_i \beta_j \gamma_k Cum[x(t_i), x(t_j), x(t_k)]$$

P2. Additivity: if $\boldsymbol{x}$ and $\boldsymbol{y}$ are *independent*, the cumulant of their sum equals the sum of their cumulants:

$$Cum[x(t_1)+y(t_1),\ldots,x(t_n)+y(t_n)]$$
$$= Cum[x(t_1),\ldots,x(t_n)]$$
$$+ Cum[y(t_1),\ldots,y(t_n)]$$

P3. The third-order cumulant of a random variable with a symmetric pdf is equal to zero.

From assumptions A3, A4 and using properties P2, P3 we have:

$$C_{uyu}(\tau_1, \tau_2) = C_{u_0 y_0 u_0}(\tau_1, \tau_2)$$
$$C_{uuu}(\tau_1, \tau_2) = C_{u_0 u_0 u_0}(\tau_1, \tau_2)$$

The third-order (cross-)cumulants of the input and output signals are thus insensitive to the symmetrically distributed noises. In the sequel, the notation $\boldsymbol{\tau} = [\tau_1\ \tau_2]$ will be used.

---

[2] The independence simplifies the analysis. The method also works when $\tilde{u}$ and $\tilde{y}$ are mutually correlated.

2.2 Third-order cumulants based system identification algorithm

The identification algorithm is based on the following result: it can be proven that the equation of the system is satisfied for the third-order cross-cumulants (see Thil et al. [2008] and references therein). More specifically:

$$A(q^{-1})C_{uyu}(\tau_1, \tau_2) = B(q^{-1})C_{uuu}(\tau_1, \tau_2) \quad (4)$$
$$\text{i.e.} \quad C_{uyu}(\tau_1, \tau_2) = \boldsymbol{\varphi}_{cum}^T(\tau_1, \tau_2)\boldsymbol{\theta} \quad (5)$$

where:

$$\boldsymbol{\varphi}_{cum}^T(\tau_1, \tau_2) = \big[-C_{uyu}(\tau_1-1, \tau_2)\ \ldots\ -C_{uyu}(\tau_1-n_a, \tau_2)$$
$$C_{uuu}(\tau_1, \tau_2)\ \ldots\ C_{uuu}(\tau_1-n_b, \tau_2)\big]$$

Setting $\tau_2 = 0$ and taking $0 \leqslant \tau_1 \leqslant M-1$ yields:

$$\begin{pmatrix} C_{uyu}(0,0) \\ \vdots \\ C_{uyu}(M-1,0) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi}_{cum}^T(0,0) \\ \vdots \\ \boldsymbol{\varphi}_{cum}^T(M-1,0) \end{pmatrix} \boldsymbol{\theta}$$

which is expressed in a compact form as:

$$\boldsymbol{C} = \boldsymbol{\Phi}^T \boldsymbol{\theta} \quad (6)$$

An estimate of the parameter vector $\boldsymbol{\theta}$ can then be obtained by solving (6) in a least squares sense, yielding the so-called Third-Order Cumulant based Least Squares estimator:

$$\boldsymbol{\theta}_{\text{tocls}} = \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1}\boldsymbol{\Phi}\boldsymbol{C} \quad (7)$$

Remark 1. There are several questions that raise, such as:

- how to choose the user parameter M;
- which cumulant slice should be chosen (here $0 \leqslant \tau_1 \leqslant M-1$ and $\tau_2 = 0$);
- is it better to use (4) or another equation, such as $A(q^{-1})C_{yyy}(\boldsymbol{\tau}) = B(q^{-1})C_{yuy}(\boldsymbol{\tau})$.

Some answers to the first two issues are given in Thil et al. [2008], and some further insight is given in the present work.

Besides, note that the third-order cumulants have to be estimated with a finite number of available data, replacing expectations by sample averages:

$$\hat{C}_{xxx}(\boldsymbol{\tau}) = \frac{1}{N}\sum_{k=1}^{N} x(t_k)x(t_k+\tau_1)x(t_k+\tau_2)$$

Since the estimate of the expectation is unbiased and consistent, third-order cumulants estimates are unbiased and consistent as well.

3. STATISTICAL ANALYSIS

A statistical analysis of the accuracy of the HOS-based method described in Section 2 is valuable. It can highly facilitate evaluation of and comparison with other identification approaches, especially methods based on second-order properties, for identifying the EIV system. Besides, one may also get insight into important issues like how different user choices of the algorithm can influence the accuracy.

By extending the approach used for the second-order based methods Söderström [2005], Hong et al. [2007], the aim is to derive an expression for the asymptotic covariance

matrix of the estimated parameters for the HOS-based estimator:

$$\boldsymbol{P} = \lim_{N\to\infty} NE\left\{ \left(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o\right)\left(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o\right)^T \right\}$$

Let:

$$\hat{\boldsymbol{C}} = \boldsymbol{C} + \tilde{\boldsymbol{C}} \quad \text{and} \quad \hat{\boldsymbol{\Phi}} = \boldsymbol{\Phi} + \tilde{\boldsymbol{\Phi}} \qquad (8)$$

be the corresponding estimates of $\boldsymbol{C}, \boldsymbol{\Phi}$, while $\tilde{\boldsymbol{C}}, \tilde{\boldsymbol{\Phi}}$ denote the relevant estimation errors. Assuming that the true process belongs to the model set (that is $G_o \in \mathcal{G}$), and utilizing (8) and (7) leads to:

$$\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o = \left(\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Phi}}^T\right)^{-1}\left(\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{C}}\right) - \boldsymbol{\theta}_o$$

$$= \left(\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Phi}}^T\right)^{-1}\left(\hat{\boldsymbol{\Phi}}\left(\hat{\boldsymbol{C}} - \hat{\boldsymbol{\Phi}}^T\boldsymbol{\theta}_o\right)\right)$$

$$= \left(\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Phi}}^T\right)^{-1}\left(\boldsymbol{\Phi}\left(\tilde{\boldsymbol{C}} - \tilde{\boldsymbol{\Phi}}^T\boldsymbol{\theta}_o\right) + \tilde{\boldsymbol{\Phi}}\left(\tilde{\boldsymbol{C}} - \tilde{\boldsymbol{\Phi}}^T\boldsymbol{\theta}_o\right)\right) \qquad (9)$$

Assume that N is large. Since the estimates of the third-order cumulants are consistent, we have $\|\tilde{\boldsymbol{C}}\| \ll \|\boldsymbol{C}\|$ and $\|\tilde{\boldsymbol{\Phi}}\| \ll \|\boldsymbol{\Phi}\|$. The following lemma is now needed.

Lemma 1. Let $\boldsymbol{R}, \tilde{\boldsymbol{R}}$ be two non-singular square matrices such that $\|\tilde{\boldsymbol{R}}\| \ll \|\boldsymbol{R}\|$. Then:

$$\left(\boldsymbol{R} + \tilde{\boldsymbol{R}}\right)^{-1} \approx \boldsymbol{R}^{-1} - \boldsymbol{R}^{-1}\tilde{\boldsymbol{R}}\boldsymbol{R}^{-1} + O(\|\tilde{\boldsymbol{R}}\|^2)$$

Using this lemma, we get:

$$\left(\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Phi}}^T\right)^{-1} = \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T + \boldsymbol{\Phi}\tilde{\boldsymbol{\Phi}}^T + \tilde{\boldsymbol{\Phi}}\boldsymbol{\Phi}^T + \tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Phi}}^T\right)^{-1}$$

$$\approx \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1} - \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1}\left(\boldsymbol{\Phi}\tilde{\boldsymbol{\Phi}}^T + \tilde{\boldsymbol{\Phi}}\boldsymbol{\Phi}^T + \tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Phi}}^T\right)\left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1}$$

Injecting that expression in (9) and neglecting second-order terms yields:

$$\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o \approx \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1}\boldsymbol{\Phi}\left(\tilde{\boldsymbol{C}} - \tilde{\boldsymbol{\Phi}}^T\boldsymbol{\theta}_o\right) \qquad (10)$$

$$\approx \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1}\boldsymbol{\Phi}\left(\hat{\boldsymbol{C}} - \hat{\boldsymbol{\Phi}}^T\boldsymbol{\theta}_o\right) \qquad (11)$$

Let $\boldsymbol{w} = \hat{\boldsymbol{C}} - \hat{\boldsymbol{\Phi}}^T\boldsymbol{\theta}_o \in \mathbb{R}^M$. Note that $\boldsymbol{\Phi}$ is a deterministic matrix, while $\boldsymbol{w}$ is a random term which has zero mean and a variance that should decrease when N increases.

Due to Assumptions A1 and A5, it follows that the third-order (cross-)cumulants $C_{uyu}(\tau_1, 0)$ and $C_{uuu}(\tau_1, 0)$ have the property of exponential forgetting, *i.e.* decay exponentially with τ_1 increasing. Under the assumptions given in Section 1 and using Theorem 3 and 4 in Ljung [1977], the following Theorem can be stated:

Theorem 2. The estimated parameter $\hat{\boldsymbol{\theta}}_{\text{tocls}}$ is asymptotically Gaussian distributed:

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o) \xrightarrow{dist} \mathcal{N}(0, \boldsymbol{P}) \qquad (12)$$

where:

$$\boldsymbol{P} = \lim_{N\to\infty} NE\left\{\left(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o\right)\left(\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o\right)^T\right\}$$

$$= \boldsymbol{Q} \lim_{N\to\infty} NE\{\boldsymbol{w}\boldsymbol{w}^T\} \boldsymbol{Q}^T \qquad (13)$$

and $\boldsymbol{Q} = (\boldsymbol{\Phi}\boldsymbol{\Phi}^T)^{-1}\boldsymbol{\Phi}$.

Proof. From (11) it can be seen that $\hat{\boldsymbol{\theta}}_{\text{tocls}} - \boldsymbol{\theta}_o \approx \boldsymbol{Q}\boldsymbol{w}$. Then, using Lemma 9A.1. in [Ljung, 1987, p. 266], it can be shown that $\boldsymbol{Q}\boldsymbol{w}$ is asymptotically Gaussian distributed with covariance $\boldsymbol{P}$ given by (13).

Let us introduce a convenient notation for the moments used in the sequel. The moment of order r of a stationary signal x will be noted as:

$$\mu_{x^r}(\boldsymbol{\tau}) = E\left\{x(t_k)x(t_{k+\tau_1})\ldots x(t_{k+\tau_{r-1}})\right\}$$

where the size $r-1$ of $\boldsymbol{\tau}$ is determined by the order of the moment.

The following lemma gives some insights into the term $\lim_{N\to\infty} NE\{\boldsymbol{w}\boldsymbol{w}^T\}$ appearing in (13).

Lemma 3. It holds that:

$$\lim_{N\to\infty} NE\{\boldsymbol{w}_{j_1+1}\boldsymbol{w}_{j_2+1}\} = \sum_{\tau=-\infty}^{\infty}\Bigg\{$$

$$\sum_{l_1=0}^{n_a}\sum_{l_2=0}^{n_a} a_{l_1}a_{l_2}\mu_{\tilde{y}^2}(\tau + j_1 - j_2 - l_1 + l_2)$$

$$\Big[\mu_{u_0^4}(0,\tau,\tau) + 4\mu_{u_0^2}(\tau)\mu_{\tilde{u}^2}(\tau)$$

$$+ 2\mu_{u_0^2}(0)\mu_{\tilde{u}^2}(0) + \mu_{\tilde{u}^4}(0,\tau,\tau)\Big]$$

$$+ \sum_{l_1=0}^{n_b}\sum_{l_2=0}^{n_b} b_{l_1}b_{l_2}\Big[\mu_{\tilde{u}^6}(0,\tau,\tau,\tau+j_1-l_1, j_2-l_2)$$

$$+ 4\mu_{u_0^2}(\tau)\mu_{\tilde{u}^4}(\tau,\tau+j_1-l_1, j_2-l_2)$$

$$+ \mu_{u_0^2}(0)\Big(\mu_{\tilde{u}^4}(0,\tau+j_1-l_1, j_2-l_2)$$

$$+ \mu_{\tilde{u}^4}(0, j_1-l_1, j_2-l_2-\tau)\Big)$$

$$+ \mu_{u_0^4}(0,\tau,\tau)\mu_{\tilde{u}^2}(\tau+j_1-l_1-j_2+l_2)\Big]\Bigg\} \qquad (14)$$

Proof. See Appendix A.

To get the asymptotic covariance matrix $\boldsymbol{P}$ from (13), the problem left now is to calculate the fourth- and sixth-order moments needed in (14). This can be done by using the following lemma.

Lemma 4. Assume that $\tilde{u}$ is given by a filtered white noise $e_{\tilde{u}}$ with variance $\mu_{e_{\tilde{u}}^2}$ and fourth moment $\mu_{e_{\tilde{u}}^4}$:

$$\tilde{u}(t_k) = H^{\tilde{u}}(q^{-1})e_{\tilde{u}}(t_k) \qquad (15)$$

$$H^{\tilde{u}}(q^{-1}) = \sum_{j=0}^{\infty} h_j^{\tilde{u}}q^{-j} \qquad (16)$$

Then the general fourth- and a particular sixth-order moment of $\tilde{u}$ can be written as:

$$\mu_{\tilde{u}^4}(\boldsymbol{\tau}) = \mu_{e_{\tilde{u}}^2}^2 \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\left(h_{j_1}^{\tilde{u}}h_{j_1+\tau_1}^{\tilde{u}}h_{j_2}^{\tilde{u}}h_{j_2+\tau_2-\tau_3}^{\tilde{u}}\right.$$

$$+ h_{j_1}^{\tilde{u}}h_{j_1+\tau_2}^{\tilde{u}}h_{j_2}^{\tilde{u}}h_{j_2+\tau_1-\tau_3}^{\tilde{u}} + h_{j_1}^{\tilde{u}}h_{j_1+\tau_3}^{\tilde{u}}h_{j_2}^{\tilde{u}}h_{j_2+\tau_2-\tau_1}^{\tilde{u}}\Big)$$

$$+ \left(\mu_{e_{\tilde{u}}^4} - 3\mu_{e_{\tilde{u}}^2}^2\right)\sum_{j_1=0}^{\infty} h_{j_1}^{\tilde{u}}h_{j_1+\tau_1}^{\tilde{u}}h_{j_1+\tau_2}^{\tilde{u}}h_{j_1+\tau_3}^{\tilde{u}} \qquad (17)$$

and:

$$\mu_{\tilde{u}^6}(0, \tau, \tau, \tau_1, \tau_2) = \mu_{e_{\tilde{u}}^2}^3 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{j_3=0}^{\infty}$$

$$\sum_{\tau_a, \tau_b, \tau_c} h_{j_1}^{\tilde{u}} h_{j_1+\tau_a}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_b}^{\tilde{u}} h_{j_3}^{\tilde{u}} h_{j_3+\tau_c}^{\tilde{u}}$$

$$+ \left(\mu_{e_{\tilde{u}}^2} \mu_{e_{\tilde{u}}^4} - 3\mu_{e_{\tilde{u}}^2}^3 \right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty}$$

$$\sum_{\tau_d, \tau_e, \tau_f, \tau_g} h_{j_1}^{\tilde{u}} h_{j_1+\tau_d}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_e}^{\tilde{u}} h_{j_2+\tau_f}^{\tilde{u}} h_{j_2+\tau_g}^{\tilde{u}}$$

$$+ \left(\mu_{e_{\tilde{u}}^6} - 15\mu_{e_{\tilde{u}}^2} \mu_{e_{\tilde{u}}^4} + 30\mu_{e_{\tilde{u}}^2}^3 \right) \sum_{j_1=0}^{\infty}$$

$$(h_{j_1}^{\tilde{u}})^2 (h_{j_1+\tau}^{\tilde{u}})^2 h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} \quad (18)$$

where $\tau_a, \dots, \tau_g$ are given in Tables B.1 and B.2.

Proof. See Appendix B.

The moments of u_0 can be obtained in the same way, assuming that u_0 is also given by a filtered white noise. From (14) we see that the statistical properties of the input noise $\tilde{u}$ have a larger impact to the asymptotic estimation accuracy than those of the output noise. This is because of the use of (4). If another equation based on the third-order cumulants is used, as for example:

$$A(q^{-1})C_{yyy}(\tau_1, \tau_2) = B(q^{-1})C_{yuy}(\tau_1, \tau_2) \quad (19)$$

then the statistical properties of the output noise $\tilde{y}$ will have more effect. Theorem 1 remains true, the only change being that the third-cumulants appearing in $\boldsymbol{C}$ and $\boldsymbol{\Phi}$ are then $C_{yyy}(\cdot, \cdot)$ and $C_{yuy}(\cdot, \cdot)$. As a result, (14) becomes:

$$\lim_{N \to \infty} NE\{\boldsymbol{w}_{j_1+1}\boldsymbol{w}_{j_2+1}\} = \sum_{\tau=-\infty}^{\infty} \left\{ \vphantom{\sum_{l_1=0}^{n_a}} \right.$$

$$\sum_{l_1=0}^{n_a} \sum_{l_2=0}^{n_a} a_{l_1} a_{l_2} \Big[\mu_{y_0^4}(0, \tau, \tau)\mu_{\tilde{y}^2}(\tau+j_1-j_2-l_1+l_2)$$

$$+ 4\mu_{y_0^2}(\tau)\mu_{\tilde{y}^4}(\tau, \tau+j_1-l_1, j_2-l_2)$$

$$+ \mu_{y_0^2}(0)\big(\mu_{\tilde{y}^4}(0, \tau+j_1-l_1, j_2-l_2)$$

$$+ \mu_{\tilde{y}^4}(0, j_1-l_1, j_2-l_2-\tau)\big)$$

$$+ \mu_{\tilde{y}^6}(0, \tau, \tau, \tau+j_1-l_1, j_2-l_2) \Big]$$

$$+ \sum_{l_1=0}^{n_b} \sum_{l_2=0}^{n_b} b_{l_1} b_{l_2} \mu_{\tilde{u}^2}(\tau+j_1-j_2-l_1+l_2)$$

$$\Big[\mu_{y_0^4}(0, \tau, \tau) + 4\mu_{y_0^2}(\tau)\mu_{\tilde{y}^2}(\tau)$$

$$+ 2\mu_{y_0^2}(0)\mu_{\tilde{y}^2}(0) + \mu_{\tilde{y}^4}(0, \tau, \tau,) \Big] \Big\}$$

4. NUMERICAL SIMULATIONS

Consider a second-order system:

$$G_0(q) = \frac{1 - 0.5q^{-1}}{1 - 1.5q^{-1} + 0.7q^{-2}}$$

The noise-free input $u_0(t_k)$ is an ARMA process:

$$u_0(t_k) = \frac{1 + 0.7q^{-1}}{1 - 0.5q^{-1}} e_{u_0}(t_k)$$

where e_{u_0} is a zero-mean exponentially distributed white noise with unit variance. The third- and fourth-order

moment of e_{u_0} are thus $\mu_{e_{u_0}^3} = 2$ and $\mu_{e_{u_0}^4} = 9$.

The measurement noise at the input is an AR process:

$$\tilde{u}(t_k) = \frac{1}{1 - 0.7q^{-1}} e_{\tilde{u}}(t_k) \quad (20)$$

where $e_{\tilde{u}}$ is a zero-mean Gaussian white noise with unit variance, and the fourth- and sixth-order moments equal to 3 and 15, respectively.

The output measurement noise $\tilde{y}$ is an ARMA process:

$$\tilde{y}(t_k) = \frac{1 + 0.7q^{-1}}{1 - 0.7q^{-1}} e_{\tilde{y}}(t_k) \quad (21)$$

where $e_{\tilde{y}}$ is also a zero-mean Gaussian white noise with unit variance, and therefore its fourth-order moment equals 3.

The Monte Carlo simulations results (200 runs) are shown below for $N = 50000$. The parameter M is chosen equal to 5. Using the tocls algorithm the following results are obtained:

$$P_{\text{sim}} = \begin{pmatrix} 83.34 & & & \\ -60.16 & 44.37 & & \\ -14.92 & 9.10 & 21.32 & \\ 112.42 & -78.99 & -35.26 & 169.62 \end{pmatrix}$$

and the theoretical normalized asymptotic covariance matrix is:

$$P = \begin{pmatrix} 80.13 & & & \\ -58.14 & 43.51 & & \\ -12.84 & 6.56 & 20.62 & \\ 109.54 & -77.22 & -30.73 & 166.95 \end{pmatrix}$$

As can be seen, the agreement between theory and simulation is fairly good.

5. CONCLUSION

The asymptotic properties of a third-order cumulants based estimator for identifying discrete-time errors-in-variables models have been analyzed. An explicit expression for the asymptotic covariance matrix of the parameter vector estimates has been derived. Monte Carlo simulations support the theoretical results.

Appendix A. PROOF OF LEMMA 3

For jth element of the vector $\boldsymbol{w}$, $0 \leqslant j \leqslant M-1$, we have:

$$\boldsymbol{w}_{j+1} = \hat{C}_{uyu}(j, 0) - \hat{\boldsymbol{\varphi}}_{cum}^T(j, 0)\boldsymbol{\theta}_o$$

$$= \frac{1}{N} \sum_{t=1}^{N} u^2(t_k)\varepsilon(t_{k+j}) \quad (A.1)$$

where:

$$\varepsilon(t_{k+j}) = A(q^{-1})\tilde{y}(t_{k+j}) - B(q^{-1})\tilde{u}(t_{k+j})$$

According to (A.1) and Assumptions A3 and A4, we have:

$$E\{\boldsymbol{w}_{j_1+1}\boldsymbol{w}_{j_2+1}\}$$

$$= E\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}u^2(t_k)u^2(t_l)\varepsilon(t_{k+j_1})\varepsilon(t_{l+j_2})\right\}$$

$$= \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\Big\{E\{u_0^2(t_k)u_0^2(t_l)\}E\{\varepsilon(t_{k+j_1})\varepsilon(t_{l+j_2})\}$$

$$+ E\{u_0^2(t_k)\}E\{\tilde{u}^2(t_l)\varepsilon(t_{k+j_1})\varepsilon(t_{l+j_2})\}$$

$$+ 4E\{u_0(t_k)u_0(t_l)\}E\{\tilde{u}(t_k)\tilde{u}(t_l)\varepsilon(t_{k+j_1})\varepsilon(t_{l+j_2})\}$$

$$+ E\{u_0^2(t_l)\}E\{\tilde{u}^2(t_k)\varepsilon(t_{k+j_1})\varepsilon(t_{l+j_2})\}$$

$$+ E\{\tilde{u}^2(t_k)\tilde{u}^2(t_l)\varepsilon(t_{k+j_1})\varepsilon(t_{l+j_2})\}\Big\} \quad (A.2)$$

Set $\tau = k - l$. Then:

$$E\{\boldsymbol{w}_{j_1+1}\boldsymbol{w}_{j_2+1}\} = \sum_{\tau=-N}^{N}\frac{N-|\tau|}{N^2}\Big\{$$

$$\mu_{u_0^4}(0,\tau,\tau)E\{\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}$$

$$+ \mu_{u_0^2}(0)E\{\tilde{u}^2(t_k)\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}$$

$$+ 4\mu_{u_0^2}(\tau)E\{\tilde{u}(t_{k+\tau})\tilde{u}(t_k)\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}$$

$$+ \mu_{u_0^2}(0)E\{\tilde{u}^2(t_{k+\tau})\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}$$

$$+ E\{\tilde{u}^2(t_{k+\tau})\tilde{u}^2(t_k)\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}\Big\} \quad (A.3)$$

Using Assumption A4, the first term in (A.3) becomes:

$$E\{\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}$$

$$= \sum_{l_1=0}^{n_a}\sum_{l_2=0}^{n_a}a_{l_1}a_{l_2}E\{\tilde{y}(t_{k+\tau+j_1-l_1})\tilde{y}(t_{k+j_2-l_2})\}$$

$$+ \sum_{l_1=0}^{n_b}\sum_{l_2=0}^{n_b}b_{l_1}b_{l_2}E\{\tilde{u}(t_{k+\tau+j_1-l_1})\tilde{u}(t_{k+j_2-l_2})\} \quad (A.4)$$

Therefore:

$$E\{\varepsilon(t_{k+\tau+j_1})\varepsilon(t_{k+j_2})\}$$

$$= \sum_{l_1=0}^{n_a}\sum_{l_2=0}^{n_a}a_{l_1}a_{l_2}\mu_{\tilde{y}^2}(\tau+j_1-l_1-j_2+l_2)$$

$$+ \sum_{l_1=0}^{n_b}\sum_{l_2=0}^{n_b}b_{l_1}b_{l_2}\mu_{\tilde{u}^2}(\tau+j_1-l_1-j_2+l_2) \quad (A.5)$$

Following the same steps the other expectations in (A.3) can be obtained, and thus:

$$E\{\boldsymbol{w}_{j_1+1}\boldsymbol{w}_{j_2+1}\} = \sum_{\tau=-N}^{N}\frac{N-|\tau|}{N^2}\Big\{$$

$$\sum_{l_1=0}^{n_a}\sum_{l_2=0}^{n_a}a_{l_1}a_{l_2}\Big[\mu_{u_0^4}(0,\tau,\tau)\mu_{\tilde{y}^2}(\tau+j_1-l_1-j_2+l_2)$$

$$+ 4\mu_{u_0^2}(\tau)\mu_{\tilde{u}^2\tilde{y}^2}(\tau,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{u_0^2}(0)\Big(\mu_{\tilde{u}^2\tilde{y}^2}(0,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{\tilde{u}^2\tilde{y}^2}(0,j_1-l_1,j_2-l_2-\tau)\Big)$$

$$+ \mu_{\tilde{u}^4\tilde{y}^2}(0,\tau,\tau,\tau+j_1-l_1,j_2-l_2)\Big]$$

$$+ \sum_{l_1=0}^{n_b}\sum_{l_2=0}^{n_b}b_{l_1}b_{l_2}\Big[\mu_{u_0^4}(0,\tau,\tau)\mu_{\tilde{u}^2}(\tau+j_1-l_1-j_2+l_2)$$

$$+ 4\mu_{u_0^2}(\tau)\mu_{\tilde{u}^4}(\tau,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{u_0^2}(0)\Big(\mu_{\tilde{u}^4}(0,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{\tilde{u}^4}(0,j_1-l_1,j_2-l_2-\tau)\Big)$$

$$+ \mu_{\tilde{u}^6}(0,\tau,\tau,\tau+j_1-l_1,j_2-l_2)\Big]\Big\} \quad (A.6)$$

Consequently, by Assumption A5, it follows that:

$$\lim_{N\to\infty}NE\{\boldsymbol{w}_{j_1+1}\boldsymbol{w}_{j_2+1}\} = \sum_{\tau=-\infty}^{\infty}\Big\{$$

$$\sum_{l_1=0}^{n_a}\sum_{l_2=0}^{n_a}a_{l_1}a_{l_2}\Big[\mu_{u_0^4}(0,\tau,\tau)\mu_{\tilde{y}^2}(\tau+j_1-j_2-l_1+l_2)$$

$$+ 4\mu_{u_0^2}(\tau)\mu_{\tilde{u}^2\tilde{y}^2}(\tau,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{u_0^2}(0)\Big(\mu_{\tilde{u}^2\tilde{y}^2}(0,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{\tilde{u}^2\tilde{y}^2}(0,j_1-l_1,j_2-l_2-\tau)\Big)$$

$$+ \mu_{\tilde{u}^4\tilde{y}^2}(0,\tau,\tau,\tau+j_1-l_1,j_2-l_2)\Big]$$

$$+ \sum_{l_1=0}^{n_b}\sum_{l_2=0}^{n_b}b_{l_1}b_{l_2}\Big[\mu_{u_0^4}(0,\tau,\tau)\mu_{\tilde{u}^2}(\tau+j_1-j_2-l_1+l_2)$$

$$+ 4\mu_{u_0^2}(\tau)\mu_{\tilde{u}^4}(\tau,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{u_0^2}(0)\Big(\mu_{\tilde{u}^4}(0,\tau+j_1-l_1,j_2-l_2)$$

$$+ \mu_{\tilde{u}^4}(0,j_1-l_1,j_2-l_2-\tau)\Big)$$

$$+ \mu_{\tilde{u}^6}(0,\tau,\tau,\tau+j_1-l_1,j_2-l_2)\Big]\Big\} \quad (A.7)$$

which gives (14).

Appendix B. PROOF OF LEMMA 4

By virtue of the assumptions in Lemma 3, it holds that:

$$\mu_{\tilde{u}^4}(\boldsymbol{\tau}) = E\{\tilde{u}(t_k)\tilde{u}(t_{k+\tau_1})\tilde{u}(t_{k+\tau_2})\tilde{u}(t_{k+\tau_3})\}$$

$$= \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\sum_{j_3=0}^{\infty}\sum_{j_4=0}^{\infty}h_{j_1}^{\tilde{u}}h_{j_2}^{\tilde{u}}h_{j_3}^{\tilde{u}}h_{j_4}^{\tilde{u}}E\Big\{e_{\tilde{u}}(t_{k-j_1})$$

$$e_{\tilde{u}}(t_{k+\tau_1-j_2})e_{\tilde{u}}(t_{k+\tau_2-j_3})e_{\tilde{u}}(t_{k+\tau_3-j_4})\Big\} \quad (B.1)$$

As $e_{\tilde{u}}(t_k)$ is a white noise, the expectation in (B.1) is nonzero only when the time arguments are pairwise equal or all equal. Therefore:

$$E\{e_{\tilde{u}}(t_{k-j_1})e_{\tilde{u}}(t_{k+\tau_1-j_2})e_{\tilde{u}}(t_{k+\tau_2-j_3})e_{\tilde{u}}(t_{k+\tau_3-j_4})\}$$

$$= \mu_{e_{\tilde{u}}^2}^2\big(\delta_{j_1,j_2-\tau_1}\delta_{j_3-\tau_2,j_4-\tau_3} + \delta_{j_1,j_3-\tau_2}\delta_{j_2-\tau_1,j_4-\tau_3}$$

$$+ \delta_{j_1,j_4-\tau_3}\delta_{j_2-\tau_1,j_3-\tau_2}\big)$$

$$+ \Big(\mu_{e_{\tilde{u}}^4} - 3\mu_{e_{\tilde{u}}^2}^2\Big)\delta_{j_1,j_2-\tau_1}\delta_{j_2-\tau_1,j_3-\tau_2}\delta_{j_3-\tau_2,j_4-\tau_3}$$

where $\delta_{\cdot,\cdot}$ is the Kronecker delta. Inserting this expression in (B.1) we obtain:

$$\mu_{\tilde{u}^4}(\boldsymbol{\tau}) = \mu_{e_{\tilde{u}}^2}^2 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left(h_{j_1}^{\tilde{u}} h_{j_1+\tau_1}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_2-\tau_3}^{\tilde{u}} \right.$$

$$\left. + h_{j_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_1-\tau_3}^{\tilde{u}} + h_{j_1}^{\tilde{u}} h_{j_1+\tau_3}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_2-\tau_1}^{\tilde{u}} \right)$$

$$+ \left(\mu_{e_{\tilde{u}}^4} - 3\mu_{e_{\tilde{u}}^2}^2 \right) \sum_{j_1=0}^{\infty} h_{j_1}^{\tilde{u}} h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} h_{j_1+\tau_3}^{\tilde{u}} \quad \text{(B.2)}$$

For the sixth-order moment of $\tilde{u}$ the result becomes slightly more complicated:

$$\mu_{\tilde{u}^6}(0, \tau, \tau, \tau_1, \tau_2)$$
$$= E\left\{ \tilde{u}(t_k)^2 \tilde{u}(t_{k+\tau})^2 \tilde{u}(t_{k+\tau_1}) \tilde{u}(t_{k+\tau_2}) \right\}$$
$$= \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \sum_{i_3=0}^{\infty} \sum_{i_4=0}^{\infty} \sum_{i_5=0}^{\infty} \sum_{i_6=0}^{\infty} h_{i_1}^{\tilde{u}} h_{i_2}^{\tilde{u}} h_{i_3}^{\tilde{u}} h_{i_4}^{\tilde{u}} h_{i_5}^{\tilde{u}} h_{i_6}^{\tilde{u}}$$
$$E\left\{ e_{\tilde{u}}(t_{k-i_1}) e_{\tilde{u}}(t_{k-i_2}) e_{\tilde{u}}(t_{k+\tau-i_3}) e_{\tilde{u}}(t_{k+\tau-i_4}) \right.$$
$$\left. e_{\tilde{u}}(t_{k+\tau_1-i_5}) e_{\tilde{u}}(t_{k+\tau_2-i_6}) \right\} \quad \text{(B.3)}$$

The last expectation in (B.3) is non-zero in the following cases:

(1) there are three groups of equal time lags, each group composed of two time lags (fifteen combinations – see Table B.1);
(2) there are two groups of equal time lags, one group composed of two time lags, the other of four time lags (fifteen combinations – see Table B.2);
(3) all the time lags are equal (one combination).

The sixth-order moment of $\tilde{u}$ can thus be written as:

$$\mu_{\tilde{u}^6}(0, \tau, \tau, \tau_1, \tau_2) = \mu_{e_{\tilde{u}}^2}^3 \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{j_3=0}^{\infty}$$
$$\sum_{\tau_a, \tau_b, \tau_c} h_{j_1}^{\tilde{u}} h_{j_1+\tau_a}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_b}^{\tilde{u}} h_{j_3}^{\tilde{u}} h_{j_3+\tau_c}^{\tilde{u}}$$
$$+ \left(\mu_{e_{\tilde{u}}^2} \mu_{e_{\tilde{u}}^4} - 3\mu_{e_{\tilde{u}}^2}^3 \right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty}$$
$$\sum_{\tau_d, \tau_e, \tau_f, \tau_g} h_{j_1}^{\tilde{u}} h_{j_1+\tau_d}^{\tilde{u}} h_{j_2}^{\tilde{u}} h_{j_2+\tau_e}^{\tilde{u}} h_{j_2+\tau_f}^{\tilde{u}} h_{j_2+\tau_g}^{\tilde{u}}$$
$$+ \left(\mu_{e_{\tilde{u}}^6} - 15\mu_{e_{\tilde{u}}^2} \mu_{e_{\tilde{u}}^4} + 30\mu_{e_{\tilde{u}}^2}^3 \right) \sum_{j_1=0}^{\infty}$$
$$(h_{j_1}^{\tilde{u}})^2 (h_{j_1+\tau}^{\tilde{u}})^2 h_{j_1+\tau_1}^{\tilde{u}} h_{j_1+\tau_2}^{\tilde{u}} \quad \text{(B.4)}$$

where $\tau_a, \ldots, \tau_g$ have different combinations leading to multiplicative coefficients as shown in Tables B.1 and B.2.

Coeff.	τ_a	τ_b	τ_c
1	0	0	$\tau_1 - \tau_2$
2	0	$\tau - \tau_1$	$\tau - \tau_2$
2	τ	τ	$\tau_1 - \tau_2$
2	τ	τ_1	$\tau - \tau_2$
2	τ	τ_2	$\tau - \tau_1$
2	τ_1	τ	$\tau - \tau_2$
2	τ_2	τ	$\tau - \tau_1$
2	τ_1	τ_2	0

Table B.1. Values of τ_a, τ_b and τ_c for the term in (B.4). 'Coeff.' denotes the leading coefficient of the term.

Coeff.	τ_d	τ_e	τ_f	τ_g
1	0	0	$\tau - \tau_1$	$\tau - \tau_2$
2	τ	τ	τ_1	τ_2
2	τ_1	τ	τ	τ_2
2	τ_2	τ	τ	τ_1
2	τ	τ	τ_1	τ_2
1	0	0	τ_1	τ_2
1	τ_1	0	τ	τ_2
1	τ_2	0	τ	τ_1
1	$\tau - \tau_1$	0	τ	τ_2
1	$\tau - \tau_2$	0	τ	τ_1
1	$\tau_1 - \tau_2$	0	τ	τ

Table B.2. Values of τ_d, τ_e, τ_f and τ_g for the term in (B.4). 'Coeff.' denotes the leading coefficient of the term.

REFERENCES

B.D.O. Anderson and M. Deistler. Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, 5(1):1–13, 1984.

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

A. Delopoulos and G.B. Giannakis. Consistent identification of stochastic linear systems with noisy input-output data. *Automatica*, 30(8):1271–1294, August 1994.

B. Friedlander and B. Porat. Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments. *IEEE Transactions on Automatic Control*, 35(1):27–35, January 1990.

M. Hong, T. Söderström, and W.X. Zheng. Accuracy analysis of bias-eliminating least squares estimates for errors-in-variables identification. *Automatica*, 43(9): 1590–1596, September 2007.

J.-L. Lacoume, P.-O. Amblard, and P. Comon. *Statistiques d'ordre Supérieur pour le Traitement du Signal.* Masson, Paris, 1997.

L. Ljung. Some limit results for functionals of stochastic processes. Report LiTH-ISY-I-0167. Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1977.

L. Ljung. *System Identification: Theory for the User.* Prentice Hall, Englewood Cliffs, 1987.

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.

W. Scherrer and M. Deistler. A structure theory for linear errors-in-variables models. *SIAM Journal on Control and Optimization*, 36(6):2148–2175, November 1998.

T. Söderström. Accuracy analysis of the Frisch estimates for identifying errors-in-variables systems. In *44th IEEE Conference on Decision and Control*, Seville, Spain, December 2005.

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 2008.

J.K. Tugnait. Stochastic system identification with noisy input using cumulant statistics. *IEEE Transactions on Automatic Control*, 37(4):476–485, April 1992.

J.K. Tugnait and Y. Ye. Stochastic system identification with noisy input-output measurements using polyspectra. *IEEE Transactions on Automatic Control*, 40(4): 670–683, April 1995.

425