



HAL
open science

A cumulant statistics-based method for continuous-time errors-in-variables model identification

Stéphane Thil, Hugues Garnier, Marion Gilson

► **To cite this version:**

Stéphane Thil, Hugues Garnier, Marion Gilson. A cumulant statistics-based method for continuous-time errors-in-variables model identification. 4th International Workshop on Total Least Squares and Errors-in-variables Modeling, Aug 2006, Leuven, Belgium. hal-00121389

HAL Id: hal-00121389

<https://hal.science/hal-00121389>

Submitted on 20 Dec 2006

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.

A cumulant statistic–based method for continuous–time errors–in–variables model identification

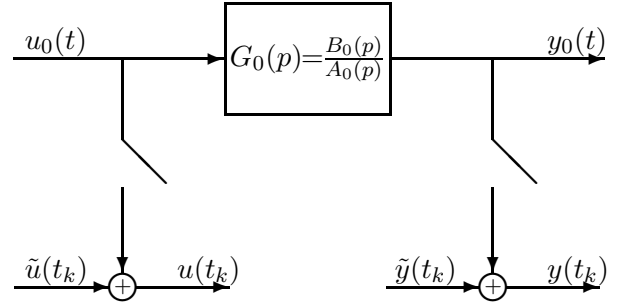
Stéphane Thil, Hugues Garnier, Marion Gilson

Centre de Recherche en Automatique de Nancy – CRAN UMR 7039 CNRS-UHP-INPL,
 Université Henri Poincaré, Nancy 1, BP 239 – 54506 Vandœuvre-lès-Nancy Cedex – France
 firstname.lastname@cran.uhp-nancy.fr

Continuous-time errors-in-variables system identification

In this paper the continuous-time errors-in-variables model depicted in the opposite Figure is considered.

Errors-in-variables (EIV) models, where uncertainties or measurement noises are present on both input and output observations, play 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 behavior. Numerous scientific disciplines use such EIV models, including time series modeling, array signal processing for direction-of-arrival estimation, blind channel equalization, multivariate calibration in analytical chemistry, image processing, or environmental modeling [9].



Furthermore, in many areas of science and engineering, the identified dynamic models should be physically meaningful. As a result, there is a need for modeling approaches that are able to yield directly from the sampled data efficiently parameterized (parsimonious) continuous-time models that have clear physical interpretations. The attention in the system identification community was almost completely focused on the discrete-time model identification techniques until recently. The last decade has indeed witnessed considerable development in continuous-time approaches to system identification from sampled data (see [5] and [8, 3] for more recent references).

The goal of this paper is to present an approach for continuous-time modeling that can take into account colored measurement noise in both input and output observations. Many methods have been proposed to solve the EIV problem in discrete-time, whereas in continuous-time it is relatively unexplored. An overview of the main discrete-time methods can be found in [6]. Regarding the continuous-time, an approach has been recently proposed in [4], assuming the noises contaminating the data to be white.

Unless we impose certain assumptions on the signal and noise models, it is well-known that the general EIV model is not uniquely identifiable from second order statistics [1]. Although that problem can be overcome by adding supplementary conditions, EIV models suffer from this lack of identifiability. This motivates the approaches based on higher-order statistics.

Higher-order statistics

The proposed methods are based upon the third-order cumulants; their main properties are quickly recalled. Some statistical assumptions on the noise-free input signal and on the noises are necessary: the probability density function of the input signal is assumed to be non-symmetric, whereas the noises are assumed to be symmetrically distributed. The differential equation of the system is then satisfied by the third-order cumulants [2]

$$C_{uyu}(\tau_1, \tau_2) = G(p, \boldsymbol{\theta}) C_{uuu}(\tau_1, \tau_2) = \frac{B(p)}{A(p)} C_{uuu}(\tau_1, \tau_2) \quad (1)$$

where C_{uuu} , C_{uyu} are the third-order (cross-)cumulants and $G(p, \boldsymbol{\theta})$ is the parametrization of the real system. The noise-cancellation property of the third-order cumulants implies that equation (1) is (asymptotically) noise-free, consequently the simple least-squares method gives consistent estimates. However, when only a

finite data record is available, errors appear in both left- and right-hand side of equation (1). To obtain estimates of the parameter vector, two possibilities are then considered.

Linear regression. To estimate the parameter vector θ , the linear regression theory can be applied to equation (1). Minimizing the following equation error

$$e_1(\tau_1, \tau_2) = A(p)C_{uyu}(\tau_1, \tau_2) - B(p)C_{uuu}(\tau_1, \tau_2) \quad (2)$$

two criterion-based estimators are derived: the simple LS estimator and the TLS estimator.

Non-linear regression: the Steiglitz-McBride algorithm. From equation (1), θ can also be derived by minimizing the following output error

$$e_2(\tau_1, \tau_2) = C_{uyu}(\tau_1, \tau_2) - \frac{B(p)}{A(p)}C_{uuu}(\tau_1, \tau_2) \quad (3)$$

This output error is non-linear in the parameters. To avoid the recourse to non-linear optimization, following the work of J.M.M. Anderson in discrete-time [2], the Steiglitz-McBride [7] algorithm is used. An equation error is consequently defined, converging towards the output error (3) in an iterative fashion. Another criterion-based estimator is then defined.

The state variable filter. One of the key points in continuous-time system identification is how to handle time-derivation. Here the cumulants time-derivatives are needed and to estimate them the state variable filter [10] is utilized: in a first step the derivatives of the input/output signals are estimated, then the cumulants derivatives are computed from these estimates.

References

- [1] B.D.O. Anderson and M. Deistler. Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, 5(1):1–13, 1984.
- [2] J.M.M. Anderson and G.B. Giannakis. Noisy input/output system identification using cumulants and the Steiglitz-McBride algorithm. *IEEE Transactions on Signal Processing*, 44(2):407–409, May 1996.
- [3] H. Garnier, M. Mensler, and A. Richard. Continuous-time model identification from sampled data. Implementation issues and performance evaluation. *International Journal of Control*, 76(13):1337–1357, 2003.
- [4] K. Mahata and H. Garnier. Direct identification of continuous-time errors-in-variables models. In *16th World Congress of the IFAC*, Prague, Czech Republic, July 2005.
- [5] N.K. Sinha and G.P. Rao. *Identification of continuous-time systems. Methodology and computer implementation*. Kluwer Academic Publishers, Dordrecht, 1991.
- [6] T. Söderström, U. Soverini, and K. Mahata. Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing*, 82:1139–1154, 2002.
- [7] K. Steiglitz and L.E. McBride. A technique for the identification of linear systems. *IEEE Transactions on Signal Processing*, 43(10):2386–2397, October 1965.
- [8] H. Unbehauen and G.P. Rao. A review of identification in continuous-time systems. *Annual reviews in Control*, 22:145–171, 1998.
- [9] S. Van Huffel. *Total least squares and errors-in-variables modeling*. Kluwer Academic Publishers, 2002.
- [10] P.C. Young. Parameter estimation for continuous-time models - a survey. *Automatica*, 17(1):23–39, 1981.

About the author Thil Stéphane received his M.S. degree in Mathematics from Metz University (France) in 2003, and his Master of Research degree in Automatic Control and Signal Processing from Henri Poincaré University (France) in 2004. He is currently a Ph.D. student in Henri Poincaré University. His research interests focus on system identification (errors-in-variables, discrete- and continuous-time, Bayesian methods, closed-loop).