ESAIM: COCV 23 (2017) 427–442
DOI: 10.1051/cocv/2015052

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

INSTRUMENTAL VARIABLES AND LSM IN CONTINUOUS-TIME PARAMETER ESTIMATION *

Jesica Escobar[1] and Martin Enqvist[2]

Abstract. In this paper the main goal is to compare the instrumental variables and the least squares methods applied to parameter estimation in continuous-time systems, avoiding any preliminary discretization of the process, and to analyse which method is more suitable for estimation in continuous-time under stochastic perturbations. A numerical example illustrates the effectiveness of the algorithms.

Mathematics Subject Classification. 93E03, 60H10, 93E10.

Received July 8, 2013. Revised April 21, 2015. Accepted October 27, 2015.

1. Introduction

System identification is the process of developing or improving a mathematical representation of a system using experimental data [11]. It consists of three basic steps, which are interrelated: the design of an experiment; the construction of a model, black box or from physical laws; and the estimation of the model parameters from measurements [15]. In control problems, the final goal is to design control strategies for a particular structural system. The identification task is to identify a model which will adequately describe the input and output map.

Some models do not show a deterministic behaviour and present stochastic perturbations. These systems are usually consisting of a deterministic part, captured by the system model; and a stochastic part, modelled as a noise distortion. During the past years, the interest in the study of stochastic models has increased dramatically. Intensified research activity in this area has been stimulated by the need to take into account random effects in complicated physical systems [8]. The tools required to deal with these kinds of systems are fairly recent developments [4].

The filtering and prediction theory developed by Wiener and Kolmogorov is one of the cornerstones in stochastic control theory. This theory makes it possible to extract a signal from observations of signal and disturbances.

For system identification in discrete time many techniques have been developed; this problem is often addressed by the least squares method (LSM) (see [10, 13]). The method of instrumental variables (IV) has

Keywords and phrases. Parameter estimation, continuous-time, stochastic systems, instrumental variable.

* *The authors thank the Mexican Council of Science and Technology (CONACYT) for the support given to this project.*

[1] School of Mechanical and Electrical Engineering, National Polytechnic Institute, Department of Control Automatics, Av. IPN Col. Lindavista 07738 Mexico City, Mexico. jeazesme@gmail.com

[2] Division of Automatic Control,University of Linköping, 581 83 Linköping, Sweden. maren@isy.liu.se

traditionally been viewed as a response to a common problem in regression contexts, *i.e.*, where one or more of the regressors on the right-hand side of the proposed equation are correlated with the equation disturbance [1].

In [9], various instrumental variable-based methods are proposed for estimating continuous-time models of systems operating in closed-loop, where the accuracy of these methods is also investigated leading to the definition of the optimal IV estimator which gives minimum variance. A computationally efficient estimator of continuous-time autoregressive (AR) process parameters from irregularly sampled data affected by discrete-time white noise is presented in [14]. It is described how an instrumental variable approach can be used for estimating the AR process parameters with high accuracy. The refined instrumental variable method for continuous-time (RIVC) has been well accepted in many disciplines. In [12], it is shown a study that fills up a theoretical gap by proving the convergence property of the RIVC method.

The use of a sliding mode (SM) control technique is beneficial in parameter estimation of deterministic dynamic models, since it can provide a global convergence of the estimation error to zero with immeasurable deterministic bounded noise [5,18]. In some mechanical systems the, so-called, supertwisting second-order sliding mode algorithms can be applied in order to design an observer (see [3]). These observers also permit the identification of disturbances using the equivalent output injection which, under some circumstances, also provides the estimation of the system parameters via a continuous version of the least squares method (see for example [17]). In [7], parameter estimation of continuous-time systems under colored perturbations using SM and LSM is presented, where the sliding modes are used as a step to construct an auxiliary process that is implemented in parallel with the LSM. This method, even when does not cancel the noise, permits to introduce an additional term in the estimation algorithms, that will improve the performance and speed.

In this paper we deal with the problem of parameter estimation in continuous-time stochastic systems using the instrumental variable method. The main idea is to compare the effectiveness of the IV method with the least squares method (LSM) in continuous-time systems under white noise perturbations. The IV method combined with the use of an observer based on the sliding mode (SM) technique, known as the "equivalent control", is performed too in order to analyze the benefits of variable structure techniques in parameter estimation. These methods are very suitable for analogous, hybrid electronic devices and integral schemes since it does not require an Analog-Digital converter that produces a loss of information and introduces additional uncertainties. The continuous-time estimation algorithms also have shown to be useful in neuro-fuzz applications (see [2]).

2. Problem formulation

Consider the stochastic differential equation in continuous time

$$\mathrm{d}x_t = A_t\zeta(x_t)\mathrm{d}t + f_t\mathrm{d}t + \sigma_t\mathrm{d}W_t \tag{2.1}$$

where A_t is the matrix to estimate, and is assumed to be time varying and bounded ($\|A_t\| \le A^+$), f_t is a bounded excitation input ($\|f_t\| \le f^+$), $\zeta(x_t)$ is a measurable nonlinear Lipschitz vector function or, a "regressor", σ_t is a bounded matrix too, and W_t is a standard vector Wiener process. The following assumptions will be used in the following sections

A1. For all x, the regressor $\zeta(x)$ is quasi-linear, *i.e.*, it satisfies the global Lipschitz condition

$$\|\zeta(x)\| \le L\,\|x\|$$

for some constant L.

A2. The parametric matrix A_t is differentiable and $\dot{A}_t$ is bounded for all $t \ge 0$

$$\left\|\dot{A}_t\right\| \le \delta_A.$$

A3. The noise power σ_t is bounded, *i.e.*, for any $t \ge 0$

$$\mathrm{tr}\left\{\sigma_t\sigma_t^\top\right\} \le D_\sigma < \infty.$$

This is required in order to guarantee the existence of any moment more than 2.

A4. The plant is considered to be L_8-stable, *i.e.*, the 8th moments exist and are bounded

$$\sup_{t>0} E\left\{\|x_t\|^8\right\} \leq X_8^+ < \infty$$

independently of the initial condition.

The main problem is to design an estimator of A_t using the instrumental variable method, compare the performance of this with the LSM algorithm, and analyze which one is more suitable for estimation in continuous-time.

First let us rewrite the system in a regressive form (since the parametric matrix to estimate is time varying we cannot take it outside of the integral, and apply least squares method and instrumental variables directly but we can change the integration interval making it equal to a window of width h, where $(t \geq h > 0)$); from this point, in order to simplify notation, we will define $\zeta(x_t) = x_t$. Then the system is given by

$$x_t - x_{t-h} - \int_{t-h}^{t} f_{s'}\mathrm{d}s' = \int_{t-h}^{t} A_{s'}x_{s'}\mathrm{d}s' + \int_{t-h}^{t} \sigma_{s'}\mathrm{d}W_{s'} \tag{2.2}$$

and

$$F_{t,t-h} = A_t X_{t,t-h} + \xi_{t,t-h} \tag{2.3}$$

where

$$F_{t,t-h} := x_t - x_{t-h} - \int_{t-h}^{t} f_{s'}\mathrm{d}s', \tag{2.4}$$

$$X_{t,t-h} := \int_{s'=t-h}^{t} x_{s'}\mathrm{d}s', \tag{2.5}$$

and

$$\xi_{t,t-h} := \int\limits_{t-h}^{t} \sigma_{s'}\mathrm{d}W_{s'} + \int\limits_{t-h}^{t} (A_{s'} - A_t)x_{s'}\mathrm{d}s'. \tag{2.6}$$

Let us define an auxiliary system for the instrumental variable method as

$$\mathrm{d}z_t = \tilde{A}_t z_t \mathrm{d}t + f_t \mathrm{d}t$$

and let

$$Z_{t,t-h} := \int_{t-h}^{t} z_{s'}\mathrm{d}s' \tag{2.7}$$

be the chosen instrument for the IV method with

$$\sup_{t>0} E\left\{\|z_t\|^8\right\} \leq Z_8^+ < \infty \tag{2.8}$$

as in assumption A4. Now integrating (2.4) from $t-h$ to t and multiplying by $Z_{t,t-h}^\top$ in the right hand side yields

$$\begin{aligned}\int\limits_{t-h}^{t} F_{\tau,\tau-h}Z_{\tau,\tau-h}^\top \mathrm{d}\tau &= \int\limits_{t-h}^{t} A_\tau X_{\tau,\tau-h}Z_{\tau,\tau-h}^\top \mathrm{d}\tau + \int\limits_{t-h}^{t} \xi_{\tau,\tau-h}Z_{\tau,\tau-h}^\top \mathrm{d}\tau \\ &= A_t \int\limits_{t-h}^{t} X_{\tau,\tau-h}Z_{\tau,\tau-h}^\top \mathrm{d}\tau + \bar{\xi}_{t,t-h} \end{aligned} \tag{2.9}$$

where

$$\bar{\xi}_{t,t-h} = \int_{t-h}^{t} \left(\left[A_\tau - A_t \right] X_{\tau,\tau-h} + \xi_{\tau,\tau-h} \right) Z_{\tau,\tau-h}^\top \mathrm{d}\tau. \tag{2.10}$$

Equation (2.9), can be used as a starting point for the IV algorithm development.

3. Algorithm development

Using a similar approach as in [7], an estimator of A_t can be defined as

$$\hat{A}_t := \left[\int_{t-h}^{t} F_{\tau,\tau-h} Z_{\tau,\tau-h}^\top \mathrm{d}\tau \right] \mathit{\Gamma}_t \tag{3.1}$$

where

$$\mathit{\Gamma}_t := \left[\int_{t-h}^{t} X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top \mathrm{d}\tau \right]^{-1} \tag{3.2}$$

or expressed alternatively as

$$\hat{A}_t := \left[\int_{0}^{t} F_{\tau,\tau-h} Z_{\tau,\tau-h}^\top \chi \left(\tau \geq t-h \right) \mathrm{d}\tau \right] \mathit{\Gamma}_t \tag{3.3}$$

$$\mathit{\Gamma}_t := \left[\int_{0}^{t} X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top \chi \left(\tau \geq t-h \right) \mathrm{d}\tau \right]^{-1}. \tag{3.4}$$

Here $\chi \left(\tau \geq t-h \right)$ is the characteristic function defined by

$$\chi \left(\tau \geq t-h \right) := \begin{cases} 1 \text{ if } \tau \geq t-h \\ 0 \text{ if } \tau < t-h \end{cases}. \tag{3.5}$$

This function characterizes the window $[t-h, t]$. Now, instead of $\chi \left(\tau \geq t-h \right)$, a different class of window can be used, for example, an "extended window" corresponding to the forgetting factor that provides the following estimation

$$\hat{A}_t = Y_t \mathit{\Gamma}_t, \ t \geq h \tag{3.6}$$

with

$$Y_t = \int_{0}^{t} F_{\tau,\tau-h} Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \tag{3.7}$$

and

$$\mathit{\Gamma}_t^{-1} := \int_{0}^{t} X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \tag{3.8}$$

where r is the scalar forgetting factor $0 < r < 1$.

3.1. Differential form of the estimation algorithm

To be able to implement this estimation algorithm, it should be written in a differential form. Now, differentiating (3.6)–(3.8) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}_t = Y_t\dot{\Gamma}_t + \dot{Y}_t\Gamma_t \tag{3.9}$$

where

$$\dot{Y}_t = F_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h} + \int_0^t F_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h}\frac{\mathrm{d}}{\mathrm{d}t}r^{t-\tau}\mathrm{d}\tau. \tag{3.10}$$

In this case $\frac{\mathrm{d}}{\mathrm{d}t}r^{t-\tau} = r^{t-\tau}\ln r$. Then, the equation can be rewritten as follows:

$$\begin{aligned}\dot{Y}_t &= F_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h} + \int_0^t F_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h}\frac{\mathrm{d}}{\mathrm{d}t}r^{t-\tau}\mathrm{d}\tau \\ &= F_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h} + Y_t\ln r\end{aligned} \tag{3.11}$$

In order to calculate $\dot{\Gamma}_t$, the identity $\Gamma_t\Gamma_t^{-1} = I$ will be used. By differentiation we get the following equations

$$\dot{\Gamma}_t\Gamma_t^{-1} + \Gamma_t\frac{\mathrm{d}}{\mathrm{d}t}\left(\Gamma_t^{-1}\right) = 0 \tag{3.12}$$

and

$$\dot{\Gamma}_t = -\Gamma_t\frac{\mathrm{d}}{\mathrm{d}t}\left(\Gamma_t^{-1}\right)\Gamma_t. \tag{3.13}$$

Differentiation yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Gamma_t^{-1} &= X_{t,t-h}Z^{\top}_{t,t-h} + \int_0^t X_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h}\frac{\mathrm{d}}{\mathrm{d}t}r^{t-\tau}\mathrm{d}\tau \\ &= X_{t,t-h}Z^{\top}_{t,t-h} + \Gamma_t^{-1}\ln r.\end{aligned} \tag{3.14}$$

Inserting $\frac{\mathrm{d}}{\mathrm{d}t}\Gamma_t^{-1}$ in (3.13) we get

$$\dot{\Gamma}_t = -\Gamma_t X_{t,t-h}Z^{\top}_{t,t-h}\Gamma_t - \ln r\,\Gamma_t. \tag{3.15}$$

Now, replacing (3.11) and (3.15) in (3.9) gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}_t &= -\hat{A}_t X_{t,t-h}Z^{\top}_{t,t-h}\Gamma_t + F_{t,t-h}Z^{\top}_{t,t-h}\Gamma_t \\ &= (-\hat{A}_t X_{t,t-h} + F_{t,t-h})Z^{\top}_{t,t-h}\Gamma_t.\end{aligned} \tag{3.16}$$

Finally, equations (3.15) and (3.16) will form the estimation algorithm:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}_t &= (-\hat{A}_t X_{t,t-h} + F_{t,t-h})Z^{\top}_{t,t-h}\Gamma_t \\ \dot{\Gamma}_t &= -\Gamma_t X_{t,t-h}Z^{\top}_{t,t-h}\Gamma_t - (\ln r)\,\Gamma_t \\ t \geq t_0 &:= \inf_t\left\{t\geq 0 : \det\Gamma_t^{-1} = \det\left(\int_0^t X_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h}r^{t-\tau}\mathrm{d}\tau\right) > 0\right\} \\ \Gamma_{t_0} &= \left[\int_0^{t_0} X_{\tau,\tau-h}Z^{\top}_{\tau,\tau-h}r^{t_0-\tau}\mathrm{d}\tau\right]^{-1},\ \hat{A}_{t_0} = Y_{t_0}\Gamma_{t_0}.\end{aligned} \tag{3.17}$$

In fact, t_0 is any time just after the moment when the matrix Γ_t^{-1} is non-singular.

3.2. Error estimation analysis

An upper bound for the estimation error $\Delta A_t = \hat{A}_t - A_t$ is presented in the following theorem:

Theorem 3.1. *Under assumptions* A1–A4 *the estimation algorithm* (3.17) *provides the following upper bound for the estimation error*

$$E\left\{\operatorname{tr}\left(\Delta A_t^\top \Gamma_t^{-2} \Delta A_t\right)\right\} \leq \left(\sqrt[8]{X_8^+ Z_8^+ h^2}\delta_A \sqrt{\frac{1}{|2\ln r|}\left[t^2 r^{2t} - \frac{t}{\ln r} r^{2t} - \frac{1-r^{2t}}{2\left(\ln r\right)^2}\right]} + \left[\sqrt[4]{Z_4^+ h}\sqrt{D_\sigma h \frac{1-r^{2t}}{|2\ln r|}} + \sqrt[4]{\delta_a Z_4^+ h}\sqrt{\frac{1-r^{2t}}{|2\ln r|}}\right]\right)^2 \tag{3.18}$$

Proof. See Appendix. □

Let us now compare this bound with the one presented in [6] for LSM given by:

$$E\left\{\operatorname{tr}\left(\Delta A_t^\top \Gamma_t^{-2} \Delta A_t\right)\right\} \leq \left(\sqrt[4]{X_8^+ h}\delta_A \sqrt{\frac{1}{|2\ln r|}\left[t^2 r^{2t} - \frac{t}{\ln r} r^{2t} - \frac{1-r^{2t}}{2\left(\ln r\right)^2}\right]} + \left[\sqrt[8]{X_8^+ h}\sqrt{\frac{D_\sigma h(1-r^{2t})}{|2\ln r|}} + \sqrt[4]{\delta_a}\sqrt[8]{X_8^+ h}\sqrt{\frac{1-r^{2t}}{|2\ln r|}}\right]\right)^2. \tag{3.19}$$

The structure of both upper bounds is very similar but in (3.18) we have the terms Z_8^+ and Z_4^+ that correspond to the 8th and 4th bounded moments of the chosen instrument z_t. These terms, that are deterministic, replace the bounded 8th moments of x_t in the upper bound reducing the influence of the stochastic terms in the estimation error, and as it will be shown in the numerical examples, this reduce the bias and the noise effects on the parameter estimation. Using the IV method, the upper bound presents fewer variables related to the stochastic noise, *i.e.*, the upper bound is less noise dependent compared to the LSM, and so, this bound should have less sensitivity to the color or level of noise. The dependence of the error estimation bound on the noise level, as well as the case for stationary parameters and the sharpness of the upper bound, are presented in [6].

3.3. Numerical examples

Example A. The performance of the algorithm is shown in the following example where the system is defined as follows

$$\mathrm{d}x_t = (-20x_t + 5\sin(0.5t) + 12)\mathrm{d}t + 0.6\mathrm{d}W_t,\ x_t(0) = 2. \tag{3.20}$$

The instrumental variable Z_t is defined by

$$\mathrm{d}z_t = (-10z_t + 5\sin(0.5t) + 12)\mathrm{d}t. \tag{3.21}$$

The numerical results are shown in Figures 1 and 2.

Choosing an instrumental variable identical to the original system is not realistic. In this case the parameter differs from the original one and still the performance of the estimation algorithm is good enough for the parameter estimation. In this example the forgetting factor for both methods is $r = 0.85$. From the figures it is possible to appreciate that the IV method shows a better performance, since the bias for LSM is quite evident compared to the IV method.

Example B. Now let us show the algorithm performance in a time-varying system defined by

$$\mathrm{d}x_t = ((\sin(0.6\pi t) - 2)x_t + 5\sin(0.5)t + 12)\mathrm{d}t + 0.5\mathrm{d}W_t,\ x(0) = 1.2 \tag{3.22}$$

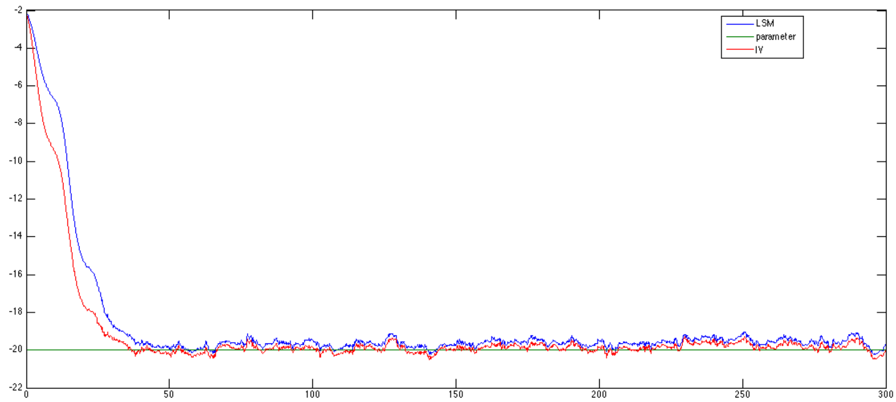

FIGURE 1. Parameter **a** and its estimates using LSM ($\hat{a}_{lsm}$) and IV ($\hat{a}_{iv}$).

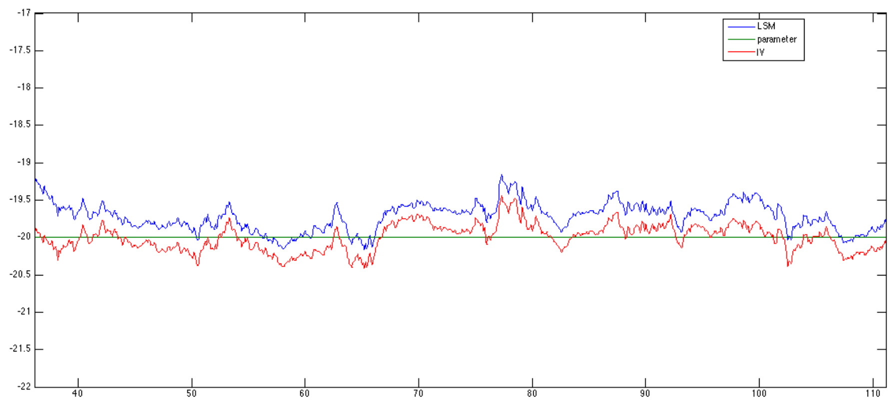

FIGURE 2. Zoomed picture showing the estimated parameters.

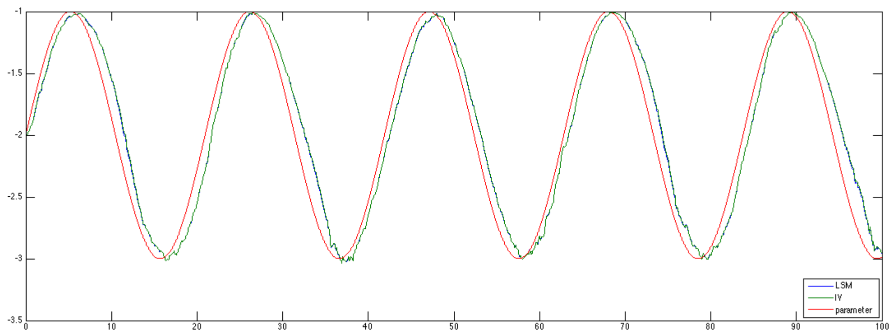

FIGURE 3. Parameter $\mathbf{a}_t$ and its estimates using LSM ($\hat{a}_{lsm}$) and IV ($\hat{a}_{iv}$).

where the instrument is defined as follows

$$\mathrm{d}z_t = ((\sin(0.6\pi t) - 1.7)z_t + 5\sin(0.5)t + 12)\mathrm{d}t. \tag{3.23}$$

Figures 3 and 4 show the numerical results for this example.

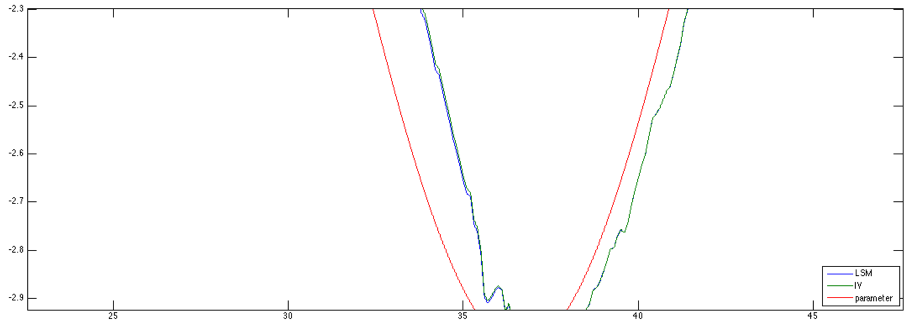

FIGURE 4. Zoomed picture showing the estimated parameters.

Here the forgetting factor for both algorithms is $r = 0.25$. Figures 3 and 4 show the performance of both algorithms, and for this example the performance is very similar and there is no a visible difference between this algorithms, like in the previous example, and both show a good performance. In the next section we will present some examples where the bias is more evident.

4. Sliding modes for estimation

In the continuous-time LSM algorithm the so-called "equivalent control", used in sliding modes techniques, has shown to be very useful for parameter estimation since we can introduce additional information into the algorithm that will improve its performance and speed.

The equivalent control method is a procedure suggested to obtain sliding equations along the intersection of a set of discontinuity surfaces. From the geometric point of view, this point implies a replacement of the undefined discontinued control on the discontinuity boundary with a continuous control which directs the velocity vector in the system state space along the discontinuity surfaces intersection (see [19]).

In this method first we introduce an auxiliary process $\hat{x}_t$ give by $\frac{\mathrm{d}}{\mathrm{d}t}\hat{x}_t = u_t$ or, in an equivalent form $\mathrm{d}\hat{x}_t$. For this first define the error vector $\mathit{\Delta}_t = x_t - \hat{x}_t$. This process u_t, provides an additional input to the parameter estimation algorithm that will improve the performance. Here the dynamic is covered by

$$\mathrm{d}\mathit{\Delta}_t = \mathrm{d}x_t - \mathrm{d}\hat{x}_t = (A_t x_t - u_t + f_t)\mathrm{d}t + \sigma_t \mathrm{d}W_t. \tag{4.1}$$

In order to design u_t the following Lyapunov-like function is suggested

$$V(\mathit{\Delta}, \widehat{x}) = \frac{1}{2}\|\mathit{\Delta}\|^2 + \beta\|\widehat{x}\|, \beta > 0. \tag{4.2}$$

By Ito's formula its differential is

$$\begin{aligned} \mathrm{d}V_t &= \left(\frac{\partial}{\partial \mathit{\Delta}}V(\cdot), \mathrm{d}\mathit{\Delta}_t\right) + \left(\frac{\partial}{\partial \hat{x}}V(\cdot), \mathrm{d}\hat{x}_t\right) + \frac{1}{2}\mathrm{tr}\left\{\sigma_t\sigma_t^\top \bigtriangledown^2 V(\cdot)\right\}\mathrm{d}t \\ &= \left\{\|\mathit{\Delta}_t\|\left(A^+\|x_t\| + f^+\right) + \left(\beta\frac{\hat{x}_t}{\|\hat{x}_t\|} - \mathit{\Delta}_t, u_t\right) + \frac{\mathrm{tr}\left\{\sigma_t\sigma_t^\top\right\}}{2}\right\}\mathrm{d}t + \left(\mathit{\Delta}_t, \sigma_t \mathrm{d}W_t\right). \end{aligned} \tag{4.3}$$

In order to compensate the terms on (4.3) the following equivalent control is selected:

$$u_t = F_t\frac{\mathit{\Delta}_t}{\|\mathit{\Delta}_t\|} + G_t\frac{\widehat{x}_t}{\|\widehat{x}_t\|}\cdot \tag{4.4}$$

By replacing (4.4) in (4.3) we get

$$\begin{aligned}\mathrm{d}V_t &\leq \left\{ \|\Delta\| \left(A^+ L \|x_t\| + f^+\right) + \left(-\Delta_t + \beta \frac{\hat{x}_t}{\|\hat{x}_t\|}, F_t \frac{\Delta_t}{\|\Delta_t\|} + G_t \frac{\hat{x}_t}{\|\hat{x}_t\|}\right) + \frac{\operatorname{tr}\left\{\sigma_t \sigma_t^\top\right\}}{2} \right\} \mathrm{d}t + (\Delta_t, \sigma_t \mathrm{d}W_t) \\ &\leq \left\{ \|\Delta\| \left(A^+ L \|x\| + f^+ - F_t\right) + \beta F_t \left(\frac{\hat{x}_t}{\|\hat{x}_t\|}, \frac{\Delta_t}{\|\Delta_t\|}\right) - G_t \left(\Delta, \frac{\hat{x}_t}{\|\hat{x}_t\|}\right) + \beta G_t + \frac{\operatorname{tr}\left\{\sigma_t \sigma_t^\top\right\}}{2} \right\} \mathrm{d}t \\ &\quad + (\Delta_t, \sigma_t \mathrm{d}W_t) := R_t \mathrm{d}t + (\Delta_t, \sigma_t \mathrm{d}W_t) \end{aligned} \tag{4.5}$$

Taking in (4.5)

$$F_t := A^+ L \|x\| + f^+ + \rho, \qquad \rho > 0 \tag{4.6}$$

one can obtain

$$R_t \leq \left\{ -\rho \|\Delta_t\| + \beta F_t \left(\frac{\hat{x}_t}{\|\hat{x}_t\|}, \frac{\Delta_t}{\|\Delta_t\|}\right) + G_t \left[\beta - \left(\Delta_t, \frac{\hat{x}_t}{\|\hat{x}_t\|}\right)\right] + \frac{\operatorname{tr}\left\{\sigma_t \sigma_t^\top\right\}}{2} \right\}. \tag{4.7}$$

Finally, G_t is selected for $t > 0$ as follows:

(a) if $\left|\left(\Delta_t, \frac{\hat{x}_t}{\|\hat{x}_t\|}\right) - \beta\right| \geq \delta > 0$

$$G_t := \frac{\beta F_t \left(\frac{\hat{x}_t}{\|\hat{x}_t\|}, \frac{\Delta_t}{\|\Delta_t\|}\right) + \frac{\operatorname{tr}\left\{\sigma_t \sigma_t^\top\right\}}{2} + \varepsilon}{\left[\left(\Delta_t, \frac{\hat{x}_t}{\|\hat{x}_t\|}\right) - \beta\right]}, \ \varepsilon > 0 \tag{4.8}$$

(b) if $\left|\left(\Delta_t, \frac{\hat{x}_t}{\|\hat{x}_t\|}\right) - \beta\right| < \delta$

$$G_t := \frac{\beta F_t \left(\frac{\hat{x}_t}{\|\hat{x}_t\|}, \frac{\Delta_t}{\|\Delta_t\|}\right) + \frac{\operatorname{tr}\left\{\sigma_t \sigma_t^\top\right\}}{2} + \varepsilon}{\delta}. \tag{4.9}$$

The small parameter δ plays the role of the, so-called, regularizing parameter which provides the possibility to avoid the problems with a singularity of the term G_t. For the estimation algorithm, considering the equivalent control $u_t = u_t^{eq}$, the next equality will be used

$$\mathrm{d}\Delta_t = (A_t x_t - u_t^{eq} + f_t)\mathrm{d}t + \sigma_t \mathrm{d}W_t \underbrace{=}_{a.s.} 0. \tag{4.10}$$

Following the same procedure used in the previous section this yields the following estimation algorithm

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\hat{A}_t &= (-\hat{A}_t x_t + f_t - u_t^{eq}) z_t^\top \Gamma_t \\ \dot{\Gamma}_t &= -\Gamma_t x_t z_t^\top \Gamma_t - (\ln r)\, \Gamma_t \\ t \geq t_0 &:= \inf_t \left\{ t \geq 0 : \det \Gamma_t^{-1} = \det \left(\int_0^t x_\tau z_\tau^\top r^{t-\tau} \mathrm{d}\tau \right) > 0 \right\} \\ \Gamma_{t_0} &= \left[\int_0^{t_0} x_\tau z_\tau^\top r^{t-\tau} \mathrm{d}\tau \right]^{-1}, \ \hat{A}_{t_0} = Y_{t_0} \Gamma_{t_0}. \end{aligned} \tag{4.11}$$

The analysis of the upper bound for the estimation error is similar to the one shown in the previous section, and is also presented in [16].

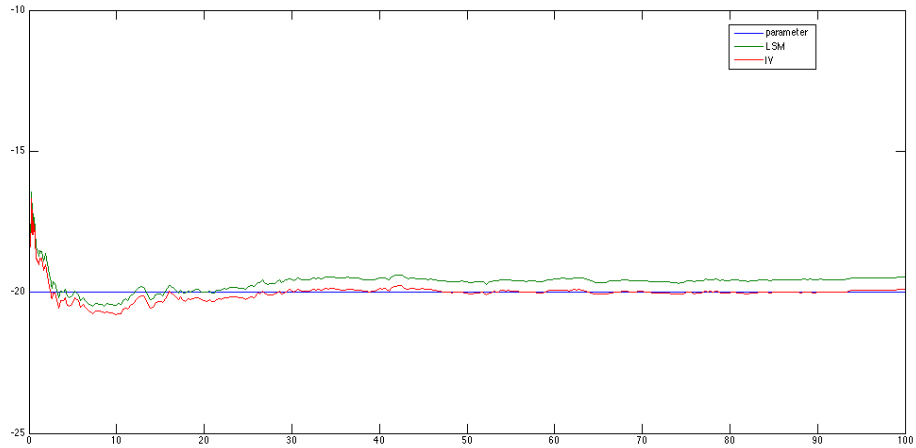

FIGURE 5. Parameter a and its estimated using LSM and IV.

4.1. Numerical examples

Example A. Here we will show the performance of the algorithm using the same system from the previous section but with $\sigma_t = 0.8$. The instrument is also the same used in the previous section and the numerical results are shown in Figure 5.

The forgetting factor for both methods is $r = 1$, this means that when the "equivalent control" technique is implemented we do not need to discard information while estimating constant parameters, whereas in the previous section even for this type of estimation was necessary to discard some data. Figure 5 also shows that the estimation improves significantly with the implementation of the variable structure technique. It is faster and the effect of the noise is not as evident as in Figures 1 and 2. For the LSM it is still possible to see some bias during the estimation while for IV this is not a problem.

Example B. The estimation algorithms were also tested on in a time-varying example given by

$$\mathrm{d}x_t = (\sin(0.5t) - 2)x_t\mathrm{d}t + 0.8\mathrm{d}W_t \tag{4.12}$$

with the instrument defined as follows

$$\mathrm{d}z_t = (\sin(0.45t) - 2)z_t\mathrm{d}t. \tag{4.13}$$

Figure 6 show the numerical results.

In this figure both algorithms have a forgetting factor $r = 0.15$, and while the IV is already performing the estimation, the LSM would still need to decrease the value of the forgetting factor in order to improve its performance. Here we can see that the "equivalent control" improves the performance of the estimation and that is not necessary to discard as much information as in the previous section.

5. CONCLUSION

In this paper, an estimation algorithm using instrumental variables was implemented in a continuous-time system under Gaussian perturbations, avoiding any preliminary discretization. The IV method was compared with the least squares method showing a better performance under higher levels of noise. The quality of the estimation is higher, and it is not necessary to discard much data with the forgetting factor. A second algorithm using IV with an "equivalent control" technique was presented too, showing that variable structure techniques can improve significantly the performance of a parameter estimation algorithm. With this method the LSM and IV algorithms have a similar performance, but LSM still shows some bias in the estimation.

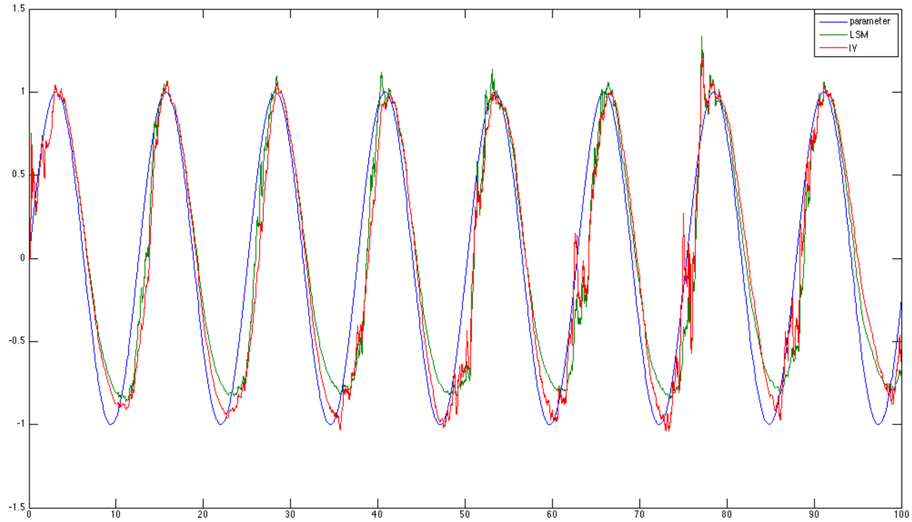

FIGURE 6. a_t and its estimated using LSM and IV.

APPENDIX A.

Proof of Theorem 3.1. Define first the error equation

$$\varDelta A_t = \hat{A}_t - A_t = Y_t \varGamma_t - A_t. \tag{A.1}$$

By replacing Y_t in (A.1) we get

$$\varDelta A_t = \left(\int_0^t F_{\tau,\tau-h} Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t - A_t \tag{A.2}$$

and by replacing equation (2.3) in the previous equation

$$\begin{aligned}
\varDelta A_t &= \left(\int_0^t \left(A_\tau X_{\tau,\tau-h} + \xi_{\tau,\tau-h} \right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t - A_t \\
&= \left(\int_0^t A_\tau X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t + \left(\int_0^t \left(\int_{\tau-h}^\tau \sigma_s \mathrm{d}W_s \right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t \\
&\quad + \left(\int_0^t \left(\int_{\tau-h}^\tau \left(A_{s'} - A_\tau \right) x_{s'} \mathrm{d}s' \right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t - A_t \\
&= \left(\int_0^t \left(A_\tau - A_t \right) X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t + \left(\int_0^t \left(\int_{\tau-h}^\tau \sigma_s \mathrm{d}W_s \right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t \\
&\quad + \left(\int_0^t \left(\int_{\tau-h}^\tau \left(A_{s'} - A_\tau \right) x_{s'} \mathrm{d}s' \right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \right) \varGamma_t
\end{aligned} \tag{A.3}$$

or in equivalent form

$$
\begin{aligned}
\Gamma_t^{-1}\Delta A_t &= P_1 + P_2' \\
P_2' &= P_2 + P_3 \\
P_1 &= -\int_0^t (A_t - A_\tau)\, X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \\
P_2 &= \int_0^t \left(\int_{\tau-h}^{\tau} \sigma_s \mathrm{d}W_s\right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau \\
P_3 &= \int_0^t \left(\int_{\tau-h}^{\tau} (A_{s'} - A_\tau)\, x_{s'} \mathrm{d}s'\right) Z_{\tau,\tau-h}^\top r^{t-\tau} \mathrm{d}\tau
\end{aligned}
$$

the next inequality it is applied to get the upper bound for the estimation error

$$
E\left\{\Delta A_t^\top \Gamma_t^{-2} \Delta A_t\right\} \le (1+\lambda)\, E\left\{\operatorname{tr}\left(P_1 P_1^\top\right)\right\} + \left(1+\lambda^{-1}\right) E\left\{\operatorname{tr}\left(P_2' P_2'^\top\right)\right\}
$$

where

$$
E\left\{\operatorname{tr}\left(P_2' P_2'^\top\right)\right\} \le (1+\lambda')\, E\left\{\operatorname{tr}\left(P_2 P_2^\top\right)\right\} + \left(1 + (\lambda')^{-1}\right) E\left\{\operatorname{tr}\left(P_3 P_3^\top\right)\right\}
$$

For the first term $E\left\{\operatorname{tr}\left(P_1 P_1^\top\right)\right\}$ we have

$$
\begin{aligned}
E\left\{\operatorname{tr}\left(P_1 P_1^\top\right)\right\} &= \int_0^t \int_0^t E\operatorname{tr}\left\{\left[(A_t\text{-}A_\tau)\right] X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top Z_{\tau',\tau'-h} X_{\tau',\tau'-h}^\top (A_t\text{-}A_{\tau'})^\top r^{t-\tau} r^{t-\tau'} \mathrm{d}\tau \mathrm{d}\tau'\right\} \\
&\le \int_0^t \int_0^t \sqrt[4]{E\left\{\|A_t\text{-}A_\tau\|^4 E\left\{\|A_t\text{-}A_{\tau'}\|^4\right\}\right\}} \\
&\quad \times \sqrt[4]{E\left\{\left\|X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top\right\|^4\right\} E\left\{\left\|X_{\tau',\tau'-h} Z_{\tau',\tau'-h}^\top\right\|^4\right\}} r^{t-\tau} r^{t-\tau'} \mathrm{d}\tau \mathrm{d}\tau' \\
&\le \left[\int_0^t \sqrt[4]{E\left\{\|A_t\text{-}A_\tau\|^4\right\}} \sqrt[4]{E\left\{\left\|X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top\right\|^4\right\}} r^{t-\tau} \mathrm{d}\tau\right]^2.
\end{aligned}
$$

Taking into account that $(A_t - A_\tau) = \int\limits_{x=\tau}^{t} \dot{A}_x \mathrm{d}x$, we get

$$
\|A_t - A_\tau\| \le \delta_A (t-\tau). \tag{A.4}
$$

This yields the following inequality

$$
E\left\{\|A_t\text{-}A_\tau\|^4\right\} \le \delta_A^4 (t-\tau)^4
$$

for the term

$$
E\left\{\left\|X_{\tau,\tau-h} Z_{\tau,\tau-h}^\top\right\|^4\right\} \le \sqrt{E\left\{\|X_{\tau,\tau-h}\|^8\right\}} \sqrt{E\left\{\|Z_{\tau,\tau-h}\|^8\right\}}
$$

and from equations (2.5) and (2.7) we get

$$
\begin{aligned}
E\left\{\|X_{\tau,\tau-h}\|^8\right\} &\le E\left\{\int_{t-h}^t \|x_{s'}\|^8 \mathrm{d}s'\right\} \\
E\left\{\|Z_{\tau,\tau-h}\|^8\right\} &\le E\left\{\int_{t-h}^t \|z_{s'}\|^8 \mathrm{d}s'\right\}.
\end{aligned}
$$

Considering assumption A4

$$\begin{aligned} E\left\{\|x_{s'}\|^8\right\} &\le X_8^+ \text{ and } E\left\{\|z_{s'}\|^8\right\} \le Z_8^+ \\ E\left\{\|X_{\tau,\tau-h}\|^8\right\} &\le X_8^+ h \\ E\left\{\|Z_{\tau,\tau-h}\|^8\right\} &\le Z_8^+ h \end{aligned} \tag{A.5}$$

and then, replacing (A.5) and (A.4) in $E\left\{\operatorname{tr}\left(P_1 P_1^\top\right)\right\}$

$$\begin{aligned} E\left\{\operatorname{tr}\left(P_1 P_1^\top\right)\right\} &\le \left[\int\limits_0^t \sqrt[4]{\delta_A^4 (t-\tau)^4}\sqrt[8]{X_8^+ h Z_8^+} r^{t-\tau} \mathrm{d}\tau\right]^2 \\ &= \sqrt[4]{X_8^+ h Z_8^8 h}\left[\int\limits_0^t \delta_A (t-\tau) r^{t-\tau}\mathrm{d}\tau\right]^2 \\ &\le \delta_A^2 \sqrt[4]{X_8^+ h Z_8^+ h}\int\limits_0^t x^2 r^{2x}\mathrm{d}x \end{aligned}$$

and since the integral in the previous equation can be expressed as follows $\int\limits_0^t x^2 r^{2x}\mathrm{d}x = \frac{1}{2\ln r}\left[t^2 r^{2t} - \frac{t}{\ln r} r^{2t} - \frac{1-r^{2t}}{2(\ln r)^2}\right]$ we get

$$E\left\{\operatorname{tr}\left(P_1 P_1^\top\right)\right\} \le \sqrt[4]{X_8^+ Z_8^+ h^2}\frac{\delta_A^2}{2\ln r}\left[t^2 r^{2t} - \frac{t}{\ln r} r^{2t} - \frac{1-r^{2t}}{2(\ln r)^2}\right].$$

The term $E\left\{\operatorname{tr}\left(P_2 P_2^\top\right)\right\}$ is estimated in the following way

$$\begin{aligned} E\left\{\operatorname{tr}\left(P_2 P_2^\top\right)\right\} &\le \int_0^t \int_0^t E\left\{\int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau} \sigma_{s'}\mathrm{d}W_{s'} Z_{\tau,\tau-h}^\top Z_{\tau,\tau-h}\mathrm{d}W_s^\top \sigma_s^\top\right\} r^{2t-\tau-\tau'}\mathrm{d}\tau\mathrm{d}\tau' \\ &\le \int\limits_0^t\int\limits_0^t E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^2\left\|\int\limits_{\tau-h}^{\tau}\sigma_{s'}\mathrm{d}W_{s'}\right\|^2\right\} r^{2t-\tau-\tau'}\mathrm{d}\tau\mathrm{d}\tau' \\ &\le \int\limits_0^t\int\limits_0^t \sqrt{E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\}}\sqrt{E\left\{\left\|\int\limits_{\tau-h}^{\tau}\sigma_{s'}\mathrm{d}W_{s'}\right\|^4\right\}} r^{2t-\tau-\tau'}\mathrm{d}\tau\mathrm{d}\tau'. \end{aligned}$$

To get the term $E\left\{\left\|\int_{\tau-h}^{\tau}\sigma_{s'}\mathrm{d}W_{s'}\right\|^4\right\}$ we will represent it in the following way

$$E\left\{\left\|\int_{\tau-h}^{\tau}\sigma_s\mathrm{d}W_s\right\|^4\right\}\le E\left\{\int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau}\left\|\sigma_s\sigma_{s'}^\top\right\|\left\|\sigma_{s''}\sigma_{s'''}^\top\right\|\mathrm{d}W_s\mathrm{d}W_{s'}^\top\mathrm{d}W_{s''}\mathrm{d}W_{s'''}^\top\right\}$$
$$=\left(\int_{\tau-h}^{\tau}\left\|\sigma_s\sigma_s^\top\right\|\mathrm{d}s\right)^2\le\left(\int_{\tau-h}^{\tau}\mathrm{tr}\left\{\sigma_s\sigma_s^\top\right\}\mathrm{d}s\right)^2\le\left(\int_{\tau-h}^{\tau}D_\sigma\mathrm{d}s\right)^2=D_\sigma^2h^2$$

where $D_\sigma\le\mathrm{tr}\left\{\sigma_s\sigma_s^\top\right\}$. Taking into account assumption A4 then $E\left\{\left\|Z_{\tau,\tau-h}\right\|^4\right\}\le Z_4^+h$, $E\left\{\mathrm{tr}\left(P_2P_2^\top\right)\right\}$ can be estimated as

$$E\left\{\mathrm{tr}\left(P_2P_2^\top\right)\right\}\le\sqrt{Z_4^+h}\sqrt{D_\sigma^2h^2}\int_0^t\int_0^t r^{2t-\tau-\tau'}\mathrm{d}\tau\mathrm{d}\tau'$$
$$=\sqrt{Z_4^+h}D_\sigma h\frac{\left(1-r^{2t}\right)}{|2\ln r|}.$$

The third term $E\left\{\mathrm{tr}\left(P_3P_3^\top\right)\right\}$ can be estimated in the following way

$$E\left\{\mathrm{tr}\left(P_3P_3^\top\right)\right\}\le E\left\{\left[\int_0^t\left(\int_{\tau-h}^{\tau}\left\|\left(A_{s'}\text{-}A_\tau\right)x_{s'}\right\|\mathrm{d}s'\right)\left\|Z_{\tau,\tau-h}^\top\right\|r^{t-\tau}\mathrm{d}\tau\right]^2\right\}$$
$$\le E\left\{\left(\int_0^t\phi_\tau\left\|Z_{\tau,\tau-h}^\top\right\|r^{t-\tau}\mathrm{d}\tau\right)^2\right\}$$

where

$$\phi_\tau:=\int_{\tau-h}^{\tau}\left\|\left(A_{s'}\text{-}A_\tau\right)x_{s'}\right\|\mathrm{d}s'.$$

This yields to the following inequality

$$E\left\{\mathrm{tr}\left(P_3P_3^\top\right)\right\}\le\int_0^t\int_0^t\sqrt{E\left\{\phi_\tau^2\left\|Z_{\tau,\tau-h}^\top\right\|^2\right\}}\sqrt{E\left\{\phi_{\tau'}^2\left\|Z_{\tau',\tau'-h}^\top\right\|^2\right\}}r^{t-\tau}r^{t-\tau'}\mathrm{d}\tau\mathrm{d}\tau'.$$

The mathematical expectation $\sqrt{E\left\{\phi_\tau^2\left\|Z_{\tau,\tau-h}^\top\right\|^2\right\}}$ can be rewritten as follows

$$E\left\{\phi_\tau^2\left\|Z_{\tau,\tau-h}^\top\right\|^2\right\}\le\sqrt{E\left\{\phi_\tau^4\right\}}\sqrt{E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\}}$$

and, as in the previous estimations, we can conclude that

$$E\left\{\phi_\tau^4\right\} = E\left\{\left(\int_{\tau-h}^{\tau} \|(A_{s'}\text{-}A_\tau)\, x_{s'}\|\, \mathrm{d}s'\right)^4\right\} \le \int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau}\int_{\tau-h}^{\tau} \Big(E\left\{\|x_s\|^8\right\} E\left\{\|(A_s\text{-}A_\tau)\|^8\right\} E\left\{\|x_{s\prime}\|^8\right\}$$

$$\times E\left\{\|(A_{s'}\text{-}A_\tau)\|^8\right\} E\left\{\|x_{s''}\|^8\right\} E\left\{\|(A_{s''}\text{-}A_\tau)\|^8\right\} E\left\{\|x_{s'''}\|^8\right\} E\left\{\|(A_{s'''}\text{-}A_\tau)\|^8\right\}\Big)^{1/8} \mathrm{d}s\mathrm{d}s'\mathrm{d}s''\mathrm{d}s'''$$

$$= \left[\int_{\tau-h}^{\tau} \left(E\left\{\|x_{s'}\|^8\right\} E\left\{\|(A_{s'}\text{-}A_\tau)\|^8\right\}\right)^{1/8} \mathrm{d}s'\right]^4. \tag{A.6}$$

From the equation (A.4) we can estimate $A_{s'}\text{-}A_\tau$

$$E\left\{\|(A_{s'}\text{-}A_\tau)\|^8\right\} \le \delta_A^8 \left(s' - \tau\right)^8$$

and for $E\left\{\|x_{s'}\|^8\right\}$ we get

$$E\left\{\|x_{s'}\|^8\right\} \le X_8^+.$$

From this point, we can rewrite (A.6) as follows

$$\delta_a := \left[\int_{\tau-h}^{\tau} \left(E\left\{\|x_{s'}\|^8\right\} E\left\{\|(A_{s'}\text{-}A_\tau)\|^8\right\}\right)^{1/8} \mathrm{d}s'\right]^4$$

$$\le \sqrt{X_8^+}\left(\int_{\tau-h}^{\tau} \delta_A \left(s' - \tau\right) \mathrm{d}s'\right)^4 = \sqrt[8]{X_8^+}\frac{h^8}{16}$$

taking into account

$$\int_0^t\int_0^t \sqrt[4]{E\left\{\phi_\tau^4\right\}}\sqrt[4]{E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\}}\sqrt[4]{E\left\{\phi_{\tau'}^4\right\}}\sqrt[4]{E\left\{\left\|Z_{\tau',\tau'-h}^\top\right\|^4\right\}} r^{t-\tau} r^{t-\tau'}\mathrm{d}\tau\mathrm{d}\tau'$$

$$= \left[\int_0^t \sqrt[4]{E\left\{\phi_\tau^4\right\}}\sqrt[4]{E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\}} r^{t-\tau}\mathrm{d}\tau\right]^2$$

$$\le \left[\sqrt[4]{\delta_a}\int_0^t \sqrt[4]{E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\}} r^{t-\tau}\mathrm{d}\tau\right]^2$$

and using the Lyapunov inequality

$$E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\} \le Z_4^+ h$$

we get

$$\left[\sqrt[4]{\delta_a}\int_0^t \sqrt[4]{E\left\{\left\|Z_{\tau,\tau-h}^\top\right\|^4\right\}} r^{t-\tau}\mathrm{d}\tau\right]^2 \le \sqrt{\delta_a}\sqrt{Z_4^+ h}\int_0^t r^{2(t-\tau)}\mathrm{d}\tau.$$

Then, the term $E\left\{\operatorname{tr}\left(P_3 P_3^\top\right)\right\}$ can be estimated as follows

$$E\left\{\operatorname{tr}\left(P_3 P_3^\top\right)\right\} \le \frac{\sqrt{\delta_a}\sqrt{Z_4^+ h}}{|2\ln r|}1 - r^{2t}.$$

Now using the next inequality $(1+\lambda)a + (1+\lambda^{-1}b) \ge (\sqrt{a}+\sqrt{b})^2$, $a, b > 0$ where $\lambda = \sqrt{b/a}$, we get

$$E\left\{\operatorname{tr}\left(\Delta A_t^\top \Gamma_t^{-2} \Delta A_t\right)\right\} \le \left(\sqrt[8]{X_8^+ Z_8^+ h^2}\,\delta_A \sqrt{\frac{1}{2\ln r}\left[t^2 r^{2t} - \frac{t}{\ln r} r^{2t} - \frac{1-r^{2t}}{2\left(\ln r\right)^2}\right]} + \left[\sqrt[4]{Z_4^+ h}\sqrt{D_\sigma h \frac{1-r^{2t}}{|2\ln r|}} + \sqrt[4]{\delta_a Z_4^+ h}\sqrt{\frac{1-r^{2t}}{|2\ln r|}}\right]\right)^2.$$

References

[1] R. Bowden and A. Darrel, *Instrumental Variables*. Cambridge University Press (1984).

[2] I. Chairez, Differential neuro-fuzzy controller for uncertain nonlinear systems. *IEEE Trans. Fuzzy Systems* **21** (2013) 369–384.

[3] J. Davila, L. Fridman and A. Levant, Second-Order Sliding-Mode Observers for Mechanical Systems. *IEEE Trans. Autom. Control* **50** (2005) 1785–1789.

[4] M.H.A. Davis, *Linear Estimation and Stochastic Control*. Chapman and Hall, London (1977).

[5] C. Edwards, S. Spurgeon, *Sliding Mode Control*. Taylor and Francis, London (1998).

[6] J. Escobar and A. Poznyak, Time-varying matrix estimation in stochastic continuous-time models under colored noise using LSM with forgetting factor. *Int. J. Syst. Sci.* **42** (2011) 2009–2020.

[7] J. Escobar and A. Poznyak, Robust Continuous-Time Matrix Estimation under Dependent Noise Perturbations: Sliding Modes Filtering and LSM with Forgetting. *Circuits Syst. Signal Process.* **28** (2009) 257–282.

[8] T. Gard, *Introduction to Stochastic Differential Equations*. Marcel Dekker, New York (1988).

[9] M. Gilson, H. Garnier and Van den P. Hof, Instrumental variable methods for continuous-time identification in closed-loop. *Proc. Amer. Control Conf.* **3** (2004) 2846–2851.

[10] P. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice Hall, Englewood Cliffs, NJ (1986).

[11] J.-N. Juan, *Applied System Identification*. Prentice Hall, Englewood Cliffs, New Jersey (1994).

[12] X. Liu, J. Wang and W.X Zheng, Convergence analysis of refined instrumental variable method for continuous-time system identification. *IET Control Theory Appl.* **5** (2011) 868–877.

[13] L. Ljung, *System Identification: Theory for the User*. Prentice Hall, Upper Saddle River, NJ (1999).

[14] M. Mossberg, High-Accuracy Instrumental Variable Identification of Continuous-Time Autoregressive Processes From Irregularly Sampled Noisy Data. *IEEE Trans. Signal Process.* **56** (2008) 4087–4091.

[15] R. Pintelon and J. Schoukens, *System Identification, A Frequency Domain Approach*. IEEE Press (2001).

[16] A. Poznyak, J. Escobar and Y. Shtessel, Sliding modes time varying matrix identification for stochastic systems. *Int. J. Syst. Sci.* **38** (2007) 847–859.

[17] D.I. Rosas Almeida, J. Alvarez and L. Fridman, Robust observation and Identification of nDOF Lagrangian systems. *Int. J. Robust Nonlin. Control* **17** (2007) 842–861.

[18] V. Utkin, *Sliding Modes Control and their Applications to Variable Structure Systems*. MIR (1978).

[19] V. Utkin, *Sliding Modes in Control and Optimization*. Springer-Verlag, Berlin, Heidelberg (1992).