

# Finite-time simultaneous parameter and state estimation using modulating functions

Jerome Jouffroy<sup>1</sup> and Johann Reger<sup>2</sup>

**Abstract**—This paper discusses the use of techniques related to the modulating function method for performing joint parameter and state estimation. After a short review of other methods related to modulating functions, we look at observability issues and the so-called Least-Squares observers to extend currently existing results on joint parameter and state estimation using modulating functions. An example is given as illustration.

**Index Terms**—modulating functions, parameter estimation, state estimation, system identification, least-squares observers.

## I. INTRODUCTION

Identifying the parameters of a dynamic system is one of the fundamental issues in control systems, and often a starting point in many control engineering processes. Among the many available techniques in identification, the so-called modulating function method, initially proposed by Shinbrot in 1957 [21], has several important advantages: it is defined in a deterministic and continuous-time framework, which makes it directly applicable to the usual models in control, mostly based on differential equations; it computes the estimate of the parameters through an integration process without resorting to the explicit differentiations of the system measurements; and its FIR-filter structure allows to obtain the estimate in finite-time.

More generally, continuous-time identification can also be related to the field of adaptive control where well-known techniques such as gradient descent, SPR-Lyapunov, etc. are typically used to estimate the parameters of a system (see [4] for an example). Among the topics of adaptive control, adaptive observers allow to jointly estimate the state of a system as well as its parameters, provided some conditions on the excitation of the system variables are satisfied, this using only the available output  $y$  (see e.g. [7] for an important reference). The convergence of the estimate toward the actual state and parameters in general is exponential or asymptotic.

This paper discusses several aspects related to simultaneous state and parameter estimation using the modulating function framework as defined originally by Shinbrot and then used in the subsequent developments of the method (see for example [1], [2], [11], [15], [16]). Surprisingly, there is actually little work done in this direction. The review paper by Rao and Unbehauen [18] hinted at this problem by using the related technique referred to as the Poisson Moment

Functional method to estimate the initial system state whose estimate would then be used in a simulation stage in order to obtain the current state. More recently, an interesting paper by Liu and Laleg-Kirati [10] considers parameter and state estimation directly on a class of linear systems.

After this introduction, Section II recalls a few basic facts about modulating functions and how they are generally used in parameter identification. In this section, we also review alternative but similar methods for estimating parameters of a system which can, in some sense, still be related to modulating functions. Then in Section III, we start with a short discussion on observability issues and the so-called Least-Squares observers [14]. Using these few observations, we extend some of the results proposed in [10] to a simple class of nonlinear systems in a way close in spirit to Least-Squares observers. A simple application example is treated in Section IV. Brief concluding remarks are given in Section V.

## II. MODULATING FUNCTIONS AND THEIR DEFINITIONS

For recalling some basic facts about modulating functions, let us first consider a linear system of order  $n$  described as

$$\begin{aligned} y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1\dot{y} + a_0y \\ = b_{n-1}u^{(n-1)} + \dots + b_1\dot{u} + b_0u \end{aligned} \quad (1)$$

where without loss of generality we assume that the system is devoid of a throughput. This may then be rephrased in the more compact form

$$y^{(n)} = \mathbf{Y}^T \boldsymbol{\theta} \quad (2)$$

where  $\mathbf{Y} = (-y, -\dot{y}, \dots, -y^{(n-1)}, u, \dot{u}, \dots, u^{(n-1)})^T$  contains the measured signals  $u$ ,  $y$  and its time derivatives, while  $\boldsymbol{\theta} = (a_0, a_1, \dots, a_{n-1}, b_0, b_1, \dots, b_{n-1})^T$  contains the constant parameters of the system to be identified.

In this section, we first assume that the input signal  $u(t)$ , but also  $\dot{u}(t)$ ,  $\ddot{u}(t)$ ,  $\dots$ , are known for any  $t$  (this assumption will be removed later in the next section). Obviously, using directly the derivatives of  $y$  to estimate  $\boldsymbol{\theta}$  should be avoided because of issues related to noise. The method of modulating functions is one way to circumvent this problem. A modulating function is usually (see for example [15], [16]) defined as follows.

**Definition 1** A function  $\varphi : [0, T] \rightarrow \mathbb{R}$  is called a modulating function (of order  $k$ ) if it is sufficiently smooth and if, for some fixed  $T$ , we have

$$\varphi^{(i)}(0) = \varphi^{(i)}(T) = 0 \quad (3)$$

for all  $i \in \{0, 1, \dots, k-1\}$ .  $\square$

<sup>1</sup>Jerome Jouffroy is with the Mads Clausen Institute, University of Southern Denmark, DK-6400 Alsion 2, Sønderborg, Denmark (e-mail: jerome@mci.sdu.dk).

<sup>2</sup>Johann Reger is director of the Institute for Automation and Systems Engineering and head of the Control Engineering Group, Technische Universität Ilmenau, P.O. Box 10 05 65, D-98684, Ilmenau, Germany (e-mail: reger@ieee.org)

Using the boundary conditions (3) and integration by parts, multiplication of a signal  $\xi$  by  $\varphi$  and integration yields

$$\int_0^T \varphi(\tau) \xi^{(i)}(\tau) d\tau = \int_0^T (-1)^i \varphi^{(i)}(\tau) \xi(\tau) d\tau. \quad (4)$$

Expression (4), most fundamental result of the modulating function method, allows both to avoid computing signal derivatives explicitly as well as to get rid of unknown initial and final conditions through conditions (3). In this case, system (2) is replaced with

$$z = \mathbf{w}^T \boldsymbol{\theta} \quad (5)$$

where

$$z = \int_0^T (-1)^n \varphi^{(n)}(\tau) y(\tau) d\tau \quad (6)$$

and the  $i$ -th term of vector  $\mathbf{w}$  is given by

$$w_i = \int_0^T (-1)^{i-1} \varphi^{(i-1)}(\tau) y(\tau) d\tau \quad \text{for } i = 1, \dots, n \quad (7)$$

and

$$w_i = \int_0^T (-1)^{i-n-1} \varphi^{(i-n-1)}(\tau) u(\tau) d\tau \quad \text{for } i = n+1, \dots, 2n. \quad (8)$$

In order to obtain an estimate  $\hat{\boldsymbol{\theta}}$  of  $\boldsymbol{\theta}$ , one usually gathers a collection of  $m \geq n$  equations (5), each of them using a different modulating function  $\varphi_k(t)$ . We then get

$$\mathbf{z} = \mathbf{W}^T \boldsymbol{\theta} \quad (9)$$

with  $\mathbf{z} = (z_1, z_2, \dots, z_m)^T$  and  $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m)$  where  $z_k$  and  $\mathbf{w}_k$  are defined similarly as in (6)–(8) and  $\varphi^{(\cdot)}(t)$  is replaced by  $\varphi_k^{(\cdot)}(t)$ . An estimate  $\hat{\boldsymbol{\theta}}$  is finally obtained by simple application of linear least squares:

$$\hat{\boldsymbol{\theta}} = (\mathbf{W}\mathbf{W}^T)^{-1} \mathbf{W}\mathbf{z}. \quad (10)$$

Obtaining the estimate as in (10) above is what is usually done in most of the literature when employing the modulating function method. However, as it is done for example in Byrski [1], one can also use a single modulating function by considering a receding-horizon version of (5). To this end write

$$z(t) = \mathbf{w}^T(t) \boldsymbol{\theta} \quad (11)$$

where

$$z(t) = \int_{t-T}^t (-1)^n \varphi^{(n)}(\tau - t + T) y(\tau) d\tau \quad (12)$$

and  $\mathbf{w}(t)$  is defined similarly. Then, one can get  $\hat{\boldsymbol{\theta}}$  by using the Gramian-based or  $L_2$ -norm-based estimator

$$\hat{\boldsymbol{\theta}}(t) = \left( \int_{t-T'}^t \mathbf{w}(\tau) \mathbf{w}^T(\tau) d\tau \right)^{-1} \int_{t-T'}^t \mathbf{w}(\tau) z(\tau) d\tau \quad (13)$$

where  $T'$  is the horizon length for the receding horizon estimation of parameter vector  $\boldsymbol{\theta}$  which can be chosen different from the horizon length  $T$  in expression (12).

Quite a few classes of modulating functions have been introduced over the years. For example, some consist of trigonometric functions

$$\varphi_k(t) = \left( \sin \frac{k\pi t}{T} \right)^k \quad (14)$$

like in the early work of Shinbrot [21], or polynomials

$$\varphi_k(t) = (T - t)^{k-1} \quad (15)$$

as in the 1965 article by Loeb and Cahen [11].

Interestingly, a number of methods aiming at avoiding the explicit differentiation of measured signals were developed somewhat in parallel to modulating functions. While they are not modulating functions per se, they can be shown to use advantage of integration by parts to avoid explicit derivative computation. This is the case for the so-called repeated integration method (see [12] for an important reference) where expression (4) is this time replaced with

$$\begin{aligned} & \int_0^t \int_0^{\tau_{k-1}} \dots \int_0^{\tau_1} \xi^{(i)}(\tau_1) d\tau_1 d\tau_2 \dots d\tau_{k-1} d\tau_k \\ &= \int_0^t \int_0^{\tau_{k-i-1}} \dots \int_0^{\tau_1} \xi(\tau_1) d\tau_1 d\tau_2 \dots d\tau_{k-1} d\tau_{k-i} \end{aligned} \quad (16)$$

for  $k > i$ . In a similar way, the so-called Poisson Moment Functional method [19] consists in a convolution of the considered signal with a kernel based on an exponential. Coming back to handling derivatives, this method gives

$$\begin{aligned} & \int_0^t \frac{(t-\tau)^k}{k!} e^{-\lambda(t-\tau)} \xi^{(i)}(\tau) d\tau \\ &= \int_0^t (-1)^i \frac{d^i}{d\tau^i} \left( \frac{(t-\tau)^k}{k!} e^{-\lambda(t-\tau)} \right) \xi(\tau) d\tau. \end{aligned} \quad (17)$$

As a difference with the technique initiated by Shinbrot and later re-invented independently by Loeb and Cahen, first note that expressions (16) and (17) are only valid if we assume that the initial conditions  $\xi(0), \dot{\xi}(0), \dots, \xi^{(i-1)}(0)$  be all equal to zero. An alternative to this consists in estimating these initial conditions as well if they are not (or wait for the initial conditions to be gradually forgotten thanks to the exponential decay within the Poisson Moment Functional method). Another difference compared to the modulating function technique is the fact that in both (16) and (17), the horizon of integration can be expanding while it is usually fixed (or receding) for conventional modulating functions.

However, and as already alluded to in [20], all the above is mostly motivated by the same basic principle, i.e. multiplying the signals by a kernel/function which is then integrated in time. As a further bridge between these methods, note that, using the formula for repeated integration, equation (16) also implies

$$\begin{aligned} & \int_0^t \int_0^{\tau_{k-1}} \dots \int_0^{\tau_1} \xi^{(i)}(\tau_1) d\tau_1 d\tau_2 \dots d\tau_{k-1} d\tau_k \\ &= \int_0^t \frac{(t-\tau)^{k-i-1}}{(k-i-1)!} \xi(\tau) d\tau. \end{aligned} \quad (18)$$

The repeated integration method can thereby be regarded as an ‘‘expanding’’ application of the modulating function

method. Hence, similarly to [9], let us relax Definition 1 with the following one.

**Definition 2** Consider a sufficiently smooth function  $\varphi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  and write one of its partial derivatives as

$$\varphi^{(i)}(t, t_1) := \left. \frac{\partial^i \varphi}{\partial \tau^i}(\tau, t_1) \right|_{\tau=t}. \quad (19)$$

Function  $\varphi(\cdot, \cdot)$  is called a modulating function (of order  $k$ ) if there exists  $t_0 < t_1$  such that

$$\varphi^{(i)}(t_0, t_1) \cdot \varphi^{(i)}(t_1, t_1) = 0 \quad (20)$$

for all  $i \in \{0, 1, \dots, k-1\}$ . A modulating function for which  $\varphi^{(i)}(t_0, t_1) = 0$  and  $\varphi^{(i)}(t_1, t_1) \neq 0$  is called a left modulating function, while a modulating function for which  $\varphi^{(i)}(t_0, t_1) \neq 0$  and  $\varphi^{(i)}(t_1, t_1) = 0$  is called a right modulating function. A modulating function whose boundaries verify  $\varphi^{(i)}(t_0, t_1) = \varphi^{(i)}(t_1, t_1) = 0$  is called total modulating function.  $\square$

From this definition, it is then possible to encompass the above-mentioned techniques in a quite simple manner. For example, it can be seen that the kernels in expression (17) and (18) are right modulating functions.

### III. SIMULTANEOUS PARAMETER AND STATE ESTIMATION IN FINITE TIME

Roughly speaking, the modulating function method and its relative are loosely based on the use of explicit integration in order to obtain an estimate of the parameters. When it comes to state estimation, reference [9] used a relaxation of the modulating function concept to obtain a state-estimate. However, as known from classical control theory, this explicit-integration perspective also takes the form of the Gramian-based state estimate (see [17])

$$\hat{\mathbf{x}}(t_1) = \mathbb{W}_r^{-1}(t_0, t_1) \int_{t_0}^{t_1} \Phi^T(\tau, t_1) \mathbf{c}^T(\tau) y(\tau) d\tau \quad (21)$$

for systems described by

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) \\ y(t) &= \mathbf{c}(t)\mathbf{x}(t) \end{aligned} \quad (22)$$

and where in (21) the matrix  $\mathbb{W}_r(t_0, t_1)$  is the so-called reconstructibility Gramian given by

$$\mathbb{W}_r(t_0, t_1) = \int_{t_0}^{t_1} \Phi^T(\tau, t_1) \mathbf{c}^T(\tau) \mathbf{c}(\tau) \Phi(\tau, t_1) d\tau \quad (23)$$

with  $\Phi(\tau, t_1)$  the state-transition matrix of system (22). In the receding horizon (i.e.  $t_0 = t - T$ ,  $t_1 = t$ ) and time-invariant case, Medvedev (see the interesting [14]) noted that state estimator (21), also called  $L_2$  estimator (see [22]) can actually be related to so-called least square observers (see Kailath [6, Section 6.2.1]). In their simplest form and for the linear time-invariant case, the state estimate of a least-square observer is given by

$$\hat{\mathbf{x}}(t) = (\mathbb{O}^T \mathbb{O})^{-1} \mathbb{O}^T \bar{\mathbf{y}} \quad (24)$$

where  $\bar{\mathbf{y}} = [y, \dot{y}, \ddot{y}, \dots, y^{(n-1)}]^T$  and  $\mathbb{O}$  is the well-known observability matrix of the Kalman criterion. In an earlier work on LS observers [13], Medvedev considers estimators which are less sensitive to noise than (24) with a more general setting involving pseudodifferential operators. In this context, the state estimate is given similarly as in (24) where matrix  $\mathbf{W}_\lambda$  replaces  $\mathbb{O}$  and vector  $\bar{\mathbf{Y}}_\lambda$  replaces  $\bar{\mathbf{y}}$ . Each element of  $\bar{\mathbf{Y}}_\lambda$  is an application of a pseudodifferential operator  $P$  on the measurement output  $y$ , and the elements of  $\mathbf{W}_\lambda$  include the symbol of operator  $P$  (for more details, see [13]). As we will see later, our results for state and parameter estimation show a strong resemblance with an LS observer as (24). Underlying these designs are mathematical expressions which can be regarded as different versions of an observability matrix  $\mathbb{O}$  combined with filtering of  $\mathbf{y}$ .

Thus, adopting this observability perspective, we recast system (1) in order to transform a state and estimation problem into a state one by writing the nonlinear system

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = -\bar{\mathbf{a}}^T \mathbf{x} + \bar{\mathbf{b}}^T \bar{\mathbf{u}}, \quad y = x_1 \\ \dot{\bar{\mathbf{a}}} = 0 \\ \dot{\bar{\mathbf{b}}} = 0 \end{cases} \quad (25)$$

where  $\bar{\mathbf{u}} = [u, \dot{u}, \ddot{u}, \dots, u^{(n-1)}]^T$ ,  $\bar{\mathbf{a}} = [a_0, a_1, \dots, a_{n-1}]^T$  and  $\bar{\mathbf{b}} = [b_0, b_1, \dots, b_{n-1}]^T$ . In terms of observability, which also has impact on state estimation, an important fact is that for a nonlinear system, as exemplified by (25), being observable might involve differentiating the output  $y$  a higher number of times than the dimension of the system (see e.g. [8], [3] for more on observability of nonlinear systems). This is illustrated by the following simple oscillator example:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -a_0 x_1, \quad y = x_1 \\ \dot{a}_0 = 0 \end{cases} \quad (26)$$

Differentiating  $y$  twice with respect to time gives the nonlinear observability mapping

$$\begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} h(x) \\ \mathcal{L}_f h(x) \\ \mathcal{L}_f^2 h(x) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ -a_0 x_1 \end{bmatrix} \quad (27)$$

which, if we replace  $x_1$  with  $y$  in the last line of (27), gives the expression

$$\begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -y \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ a_0 \end{bmatrix} \quad (28)$$

so that we could apply (24) in order to obtain an estimate  $\hat{\mathbf{x}}(t)$ . Note that matrix

$$\mathbb{O}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -y \end{bmatrix} \quad (29)$$

coincides with the Jacobian of the right hand side of (27) and

may be understood as a time-varying observability matrix. In this case, we would obviously obtain  $\hat{a}_0(t) = -\ddot{y}/y$ . However, this estimate, given  $y(t) = A \sin(\omega t + \phi)$  as a solution of (26) for  $a_0 \neq 0$ , will be singular for  $\omega t + \phi = k\pi$ ,  $k \in \mathbb{N}$ . This issue may be resolved by differentiating  $y$  one more time, yielding instead of (28) the linear expression

$$\begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \\ \ddot{\ddot{y}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -y \\ 0 & 0 & -\dot{y} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ a_0 \end{bmatrix} \quad (30)$$

which leads to the estimate

$$\hat{a}_0(t) = -\frac{-y\ddot{y} - \dot{y}\ddot{\ddot{y}}}{y^2 + \dot{y}^2}. \quad (31)$$

This can be proven to be never singular since  $y^2 + \dot{y}^2 = 1$  whenever  $y(t) = \sin(t)$ . Recall that one way to obtain  $\mathbb{O}(t)$  of appropriate dimension is to apply the differential operator  $P(D) = (1, D, D^2, \dots)^T$  with  $D = d/dt$  to the output equation in (25). However, and as discussed before, differentiation should be avoided.

Hence, following these few important considerations, we hereafter extend some of the results recently proposed in [10] by considering simultaneous state and parameter estimation for a class of nonlinear systems in a way close in spirit to what is done in the context of LS observers, as exemplified by expression (24). The class considered is given by the state-space representation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u + \sum_{k=2}^p (\alpha_k x_1^k + \beta_k u^k) + \sum_{k=1}^p \sum_{l=1}^p \gamma_{kl} x_1^k u^l \quad (32)$$

with

$$\mathbf{A} = \begin{bmatrix} -\mathbf{a} & \mathbf{I}_{n-1} \\ \mathbf{0}_{1 \times (n-1)} \end{bmatrix} \quad (33)$$

and where the output equation is simply

$$y = x_1. \quad (34)$$

The  $p(p+2)$   $n$ -dimensional parameter vectors  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\alpha_k$ ,  $\beta_k$  and  $\gamma_{kl}$  are assumed to be unknown, and only  $y(t)$  and  $u(t)$  are measured.

**Theorem 1** Consider a set of  $m_\phi + m_\varphi \geq 2np(p+2)$  modulating functions, of which  $m_\phi \geq np(p+2)$  are total modulating functions, while the  $m_\varphi \geq np(p+2)$  others are left modulating functions. Then, the state vector  $\mathbf{x}(t_1)$  of system (32)-(34) and parameter vectors  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\alpha_k$ ,  $\beta_k$  and  $\gamma_{kl}$  can be estimated in finite-time if the matrices  $\mathbf{W}$  and  $\mathbf{\Delta}$ , as given in (43) and (52), have full rank. In this case, the estimates  $\hat{\mathbf{a}}$ ,  $\hat{\mathbf{b}}$ ,  $\hat{\alpha}_k$ ,  $\hat{\beta}_k$ ,  $\hat{\gamma}_{kl}$  and  $\hat{\mathbf{x}}(t_1)$  are given by expression (10), with

$$\hat{\boldsymbol{\theta}}^T = [\hat{\mathbf{a}}^T, \hat{\mathbf{b}}^T, \hat{\alpha}_2^T, \dots, \hat{\alpha}_p^T, \hat{\beta}_2^T, \dots, \hat{\beta}_p^T, \hat{\gamma}_{11}^T, \hat{\gamma}_{12}^T, \dots, \hat{\gamma}_{pp}^T]^T \quad (35)$$

and

$$\hat{\mathbf{x}}(t_1) = \mathbf{P} \left( \mathbf{\Delta} \mathbf{\Delta}^T \right)^{-1} \mathbf{\Delta} \mathbf{q} \quad (36)$$

where  $n \times np(p+2)$  matrix  $\mathbf{P}$  and  $p(p+2)$  vector  $\mathbf{q}$  are obtained from (53)-(55) and (50).  $\square$

*Proof:* We start with the estimation of the unknown parameters taking the  $m_\phi$  total modulating functions  $\phi_q(\tau, t_1)$ ,  $q = 1, \dots, m_\phi$ . To do so, rewrite state-space representation (32)-(34) as

$$\begin{aligned} y^{(n)} = & \sum_{i=1}^n \left( -a(i)y^{(n-i)} + b(i)u^{(n-i)} \right) \\ & + \sum_{k=2}^p \sum_{i=1}^n \left( \alpha_k(i)(y^k)^{(n-i)} + \beta_k(i)(u^k)^{(n-i)} \right) \\ & + \sum_{k=1}^p \sum_{l=1}^p \sum_{i=1}^n \gamma_{kl}(i)(y^k u^l)^{(n-i)} \end{aligned} \quad (37)$$

where  $a(i)$  denotes the  $i$ -th component of parameter vector  $\mathbf{a}$  and similarly for the other vectors of (32)-(34). Then, we obtain expression  $y^{(n)} = \mathbf{Y}^T \boldsymbol{\theta}$  by defining  $\mathbf{Y}$  as

$$\mathbf{Y}^T = [-\mathbf{y}^T, \mathbf{u}^T, \mathbf{y}^{2T}, \dots, \mathbf{y}^{pT}, \mathbf{u}^{2T}, \dots, \mathbf{u}^{pT}, \mathbf{g}^{11T}, \mathbf{g}^{12T}, \dots, \mathbf{g}^{ppT}]^T \quad (38)$$

where here we defined  $\mathbf{y}$  by  $\mathbf{y} := [y^{(n-1)}, y^{(n-2)}, \dots, y^{(0)}]^T$  and similarly for other vectors, while the  $i$ -th component of vector  $\mathbf{g}^{kl}$  is  $g^{kl}(i) = (y^k u^l)^{(n-i)}$ . From there, we apply the standard modulating function method to obtain

$$z_q = \mathbf{w}_q^T \boldsymbol{\theta} \quad (39)$$

where

$$z_q = \int_{t_0}^{t_1} (-1)^n \phi_q^{(n)}(\tau, t_1) y(\tau) d\tau \quad (40)$$

and

$$\begin{aligned} \mathbf{w}_q^T = & [-\mathbf{w}_q^y, \mathbf{w}_q^u, \mathbf{w}_q^{y^2}, \dots, \mathbf{w}_q^{y^p}, \mathbf{w}_q^{u^2}, \dots, \mathbf{w}_q^{u^p}, \\ & \mathbf{w}_q^{g^{11}}, \mathbf{w}_q^{g^{12}}, \dots, \mathbf{w}_q^{g^{pp}}]^T \end{aligned} \quad (41)$$

where

$$w_q^y(i) = \int_{t_0}^{t_1} (-1)^{n-i} \phi_q^{(n-i)}(\tau, t_1) y(\tau) d\tau \quad (42)$$

and similarly for the other subvectors of  $\mathbf{w}_q$ . Using  $m_\phi$  modulating functions we then obtain expression (10), with

$$\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_q, \dots, \mathbf{w}_{m_\phi}]. \quad (43)$$

For the estimation of the state, multiply each term of (37) by a left modulating function  $\varphi_r(\tau, t_1)$  and integrate to obtain the following vectorial-form expression

$$\begin{aligned} q_r^n + \boldsymbol{\varphi}_r^T \mathbf{y}(t_1) = & -\mathbf{a}^T \mathbf{q}_r^y - \mathbf{a}^T \Gamma_r \mathbf{y}(t_1) + \mathbf{b}^T \mathbf{q}_r^u + \mathbf{b}^T \Gamma_r \mathbf{u}(t_1) \\ & + \sum_{k=2}^p \left( \alpha_k^T \mathbf{q}_{r,k}^y + \beta_k^T \mathbf{q}_{r,k}^u + \alpha_k^T \Gamma_r \mathbf{y}^k(t_1) + \beta_k^T \Gamma_r \mathbf{u}^k(t_1) \right) \\ & + \sum_{k=1}^p \sum_{l=1}^p \gamma_{kl}^T (\mathbf{q}_{r,kl}^g + \Gamma_r \mathbf{g}^{kl}(t_1)) \end{aligned} \quad (44)$$

where scalar term  $q_r^n$  is given by

$$q_r^n = \int_{t_0}^{t_1} (-1)^n \varphi_r^{(n)}(\tau, t_1) y(\tau) d\tau \quad (45)$$

and

$$\boldsymbol{\varphi}_r^T = \begin{bmatrix} \varphi_r(t_1, t_1), -\varphi_r^{(1)}(t_1, t_1), \dots \\ \dots, (-1)^{n-1} \varphi_r^{(n-1)}(t_1, t_1) \end{bmatrix}. \quad (46)$$

Vector  $\mathbf{q}_r^y$  is given by

$$\mathbf{q}_r^{yT} = \int_{t_0}^{t_1} \begin{bmatrix} (-1)^{n-1} \varphi_r^{(n-1)}(\tau, t_1), \dots \\ \dots, -\varphi_r^{(1)}(\tau, t_1), \varphi_r(\tau, t_1) \end{bmatrix} y(\tau) d\tau \quad (47)$$

and similarly for  $\mathbf{q}_r^u$ ,  $\mathbf{q}_{r,k}^y$ ,  $\mathbf{q}_{r,k}^u$  and  $\mathbf{q}_{r,kl}^g$ , while matrix

$$\boldsymbol{\Gamma}_r = \begin{bmatrix} 0 & \varphi_r(t_1, t_1) & \dots & (-1)^{n-2} \varphi_r^{(n-2)}(t_1, t_1) \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \varphi_r(t_1, t_1) \\ 0 & \dots & \dots & 0 \end{bmatrix}. \quad (48)$$

Isolating, in (44), the unknown dynamical variables (i.e.  $\mathbf{y}(t_1)$ ,  $\mathbf{u}(t_1)$ ,  $\mathbf{y}^k(t_1)$ ,  $\mathbf{u}^k(t_1)$  and  $\mathbf{g}^{kl}(t_1)$ ) and replacing the unknown parameters by their estimates  $\hat{\mathbf{a}}$ ,  $\hat{\mathbf{b}}$ ,  $\hat{\boldsymbol{\alpha}}_k$ ,  $\hat{\boldsymbol{\beta}}_k$  and  $\hat{\boldsymbol{\gamma}}_{kl}$ , we get

$$\boldsymbol{\delta}_r^T \mathbf{Y} = q_r \quad (49)$$

with

$$q_r = -q_r^n - \hat{\mathbf{a}}^T \mathbf{q}_r^y + \hat{\mathbf{b}}^T \mathbf{q}_r^u + \sum_{k=2}^p (\hat{\boldsymbol{\alpha}}_k^T \mathbf{q}_{r,k}^y + \hat{\boldsymbol{\beta}}_k^T \mathbf{q}_{r,k}^u) + \sum_{k=1}^p \sum_{l=1}^p \hat{\boldsymbol{\gamma}}_{kl}^T \mathbf{q}_{r,kl}^g \quad (50)$$

and

$$\boldsymbol{\delta}_r^T = - \left[ \boldsymbol{\varphi}_r^T + \hat{\mathbf{a}}^T \boldsymbol{\Gamma}_r, \hat{\mathbf{b}}^T \boldsymbol{\Gamma}_r, \hat{\boldsymbol{\alpha}}_2^T \boldsymbol{\Gamma}_r, \dots, \hat{\boldsymbol{\alpha}}_p^T \boldsymbol{\Gamma}_r, \hat{\boldsymbol{\beta}}_2^T \boldsymbol{\Gamma}_r, \dots, \hat{\boldsymbol{\beta}}_p^T \boldsymbol{\Gamma}_r, \hat{\boldsymbol{\gamma}}_{11}^T \boldsymbol{\Gamma}_r, \dots, \hat{\boldsymbol{\gamma}}_{pp}^T \boldsymbol{\Gamma}_r \right]^T. \quad (51)$$

Putting now the  $m_\varphi$  equations (49) together, we have

$$\boldsymbol{\Delta}^T \mathbf{Y} = \mathbf{q} \quad (52)$$

where  $\boldsymbol{\Delta} = [\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \dots, \boldsymbol{\delta}_{m_\varphi}]$  and  $\mathbf{q}^T = [q_1, q_2, \dots, q_{m_\varphi}]^T$ .

Finally, we obtain expression (36) for the state estimate  $\hat{\mathbf{x}}(t_1)$  after standard application of least-squares on (52) and by projecting estimate  $\hat{\mathbf{Y}}(t_1)$  using the matrix  $\mathbf{P}$  defined as

$$\mathbf{P} = [\mathbf{P}_a, \mathbf{P}_b, \mathbf{P}_\alpha, \mathbf{P}_\beta, \mathbf{P}_\gamma] \quad (53)$$

where submatrix  $\mathbf{P}_a$  is

$$\mathbf{P}_a = \begin{bmatrix} 0 & 0 & \dots & 0 & -1 \\ 0 & \ddots & \ddots & -1 & -\hat{a}(1) \\ \vdots & \ddots & \ddots & -\hat{a}(1) & \vdots \\ 0 & -1 & \ddots & \ddots & -\hat{a}(n-2) \\ -1 & -\hat{a}(1) & \dots & -\hat{a}(n-2) & -\hat{a}(n-1) \end{bmatrix} \quad (54)$$

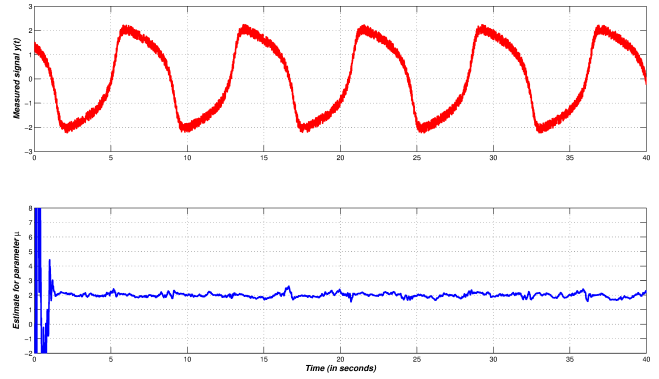


Fig. 1. Measured signal  $y(t)$  (top) and parameter estimate  $\hat{\mu}$  (bottom) for the unforced Van der Pol oscillator example.

and

$$\mathbf{P}_b = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & 0 & -\hat{b}(1) \\ \vdots & \ddots & \ddots & -\hat{b}(1) & \vdots \\ 0 & 0 & \ddots & \ddots & -\hat{b}(n-2) \\ 0 & -\hat{b}(1) & \dots & -\hat{b}(n-2) & -\hat{b}(n-1) \end{bmatrix} \quad (55)$$

while every  $n \times n$  matrix  $\mathbf{P}_{\alpha_k}$  in  $\mathbf{P}_\alpha = [\mathbf{P}_{\alpha_2}, \dots, \mathbf{P}_{\alpha_p}]$ , every  $\mathbf{P}_{\beta_k}$  in  $\mathbf{P}_\beta = [\mathbf{P}_{\beta_2}, \dots, \mathbf{P}_{\beta_p}]$ , and every  $\mathbf{P}_{\gamma_{kl}}$  in  $\mathbf{P}_\gamma = [\mathbf{P}_{\gamma_{11}}, \dots, \mathbf{P}_{\gamma_{pp}}]$  is defined similar to matrix  $\mathbf{P}_b$  in (55). This completes the proof of the theorem. ■

Note that the above is slightly more involved than the simple observability analysis of example (26), i.e. the linear form (28) is replaced with a bilinear one, mostly because of the integration by parts which is then solved by a two-stage least square procedure.

Additionally, we remark that in case one does not want to re-estimate signals which are already measured, it is quite simple, similar to the reduced-order observer context, to introduce further projection matrices and modify either of  $\mathbf{W}$ ,  $\mathbf{z}$ ,  $\boldsymbol{\Delta}$ ,  $\mathbf{q}$ , etc. accordingly.

#### IV. APPLICATION EXAMPLE

We have applied the results of Theorem 1 on a simple nonlinear example. Indeed, consider the standard form of the unforced Van der Pol oscillator represented by the following equation

$$\ddot{y} = \mu(1 - y^2)\dot{y} - y \quad (56)$$

where parameter  $\mu$  is an unknown constant. The state-space representation corresponding to (32) is

$$\begin{aligned} \dot{x}_1 &= \mu x_1 + x_2 - \frac{\mu}{3} x_1^3 \\ \dot{x}_2 &= -x_1 \end{aligned} \quad (57)$$

Expressions (32) through (55) for simultaneous parameter and state estimation were programmed on Matlab/Simulink in an S-function.

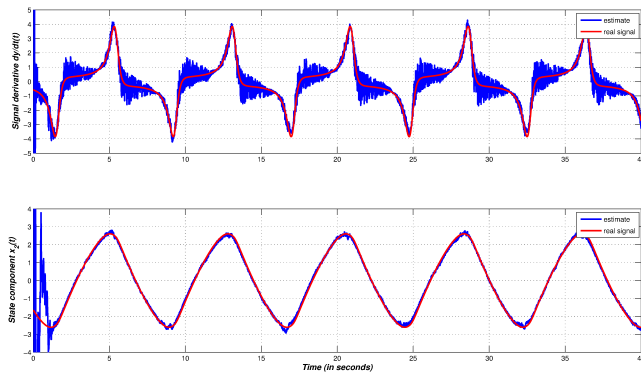


Fig. 2. Output derivative estimate (top) and  $x_2$  estimate (bottom) for the unforced Van der Pol oscillator example. Both estimates are shown against their “real” counterparts (note the noise on the derivative estimate due to a reduced-order version of the algorithm compared to the relatively good quality of  $\hat{x}_2$ ).

In order to push the algorithm to its limits, we have considered a reduced-order adaptive observer context. That is, the signals already available,  $y(t)$  and  $y^3(t)$ , are directly used to obtain the unmeasured states and not re-estimated.

The results of our simulations for  $\mu = 2$ , a SNR= 22dB (for the added measurement noise) and a sampling period of  $T_s = 0.01s$  are given in Figure 1 and Figure 2. In this case, we used 6 total modulating functions (Loeb and Cahen-like, see (15)) and 2 left modulating functions ( $t^2e^{-t}$  and  $t^3e^{-t}$ ) in the simulations.

The estimation process starts with an expanding horizon which is then switched at  $t = 2s$  to a receding-horizon of  $T = 2s$ . Note that the effect of the noise in the high amplitudes of  $y^3(t)$  can be seen on the estimate of the derivative of  $y(t)$  (see top plot of Figure 1), which still results in the quite good quality of the estimate of  $x_2(t)$  (bottom plot of Figure 2).

## V. CONCLUDING REMARKS

In this paper, we discussed a few aspects related to simultaneous parameter and state estimation using modulating functions. After a very short review, showing how other related methods may be re-attached to the modulating function method, we briefly discussed basic observability concepts, in particular those linked with the so-called Least-Squares observer approach. Based on this, we extended recently proposed results on modulating functions from [10] in a way closer to the LS observer concept.

Future work include applications in underwater navigation scenarios (see [5] for a related integral-based approach), as well as recursive aspects of the algorithms based on modulating functions and implementation issues in environments like Simulink.

## ACKNOWLEDGEMENTS

The authors are very grateful to an anonymous reviewer for providing references [9] and [10] which were not known to the authors at the time of submission.

## REFERENCES

- [1] W. Byrski, S. Fuksa and M. Nowak, “The quality of identification for different normalizations of continuous transfer functions” in *IASTED Int. Conf. on Modelling, Identification and Control*, Innsbruck, Austria, 2003.
- [2] T. B. Co and S. Ungarala, “Batch scheme recursive parameter estimation of continuous-time systems using the modulating functions method,” *Automatica*, vol. 33, no. 6, pp. 1185–1191, 1997.
- [3] E. W. Griffith and K. S. P. Kumar, “On the observability of nonlinear systems,” *Journal of Mathematical Analysis and Applications*, vol. 35, pp. 135–147, 1971.
- [4] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. ACM, 1996.
- [5] J. Jouffroy and J. Reger, “An algebraic perspective to single-transponder underwater navigation,” *IEEE Int. Conf. on Control Applications*, Munich, Germany, 2006.
- [6] T. Kailath, *Linear Systems*. Prentice-Hall, 1980.
- [7] G. Kreisselmeier, “Adaptive observers with exponential rate of convergence,” *IEEE Transactions on Automatic Control*, vol. 22, no. 1, pp. 2–8, 1977.
- [8] R. Hermann and A. J. Krener, “Nonlinear controllability and observability,” *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728–740, 1977.
- [9] D.Y. Liu and T.M. Laleg-Kirati and W. Perruquetti and O. Gibaru, “Non-asymptotic state estimation for a class of linear time-varying systems with unknown inputs,” *IFAC World Congress*, Cape Town, South Africa, 2014.
- [10] D.Y. Liu and T.M. Laleg-Kirati, “Robust fractional order differentiators using generalized modulating functions method,” *Signal Processing*, vol. 107, pp. 395–406, 2015.
- [11] J. M. Loeb and G. M. Cahen, “More about process identification,” *IEEE Transactions on Automatic Control*, vol. 10, no. 3, pp. 359–361, 1965.
- [12] A. V. Mathew and F. W. Fairman, “Identification in the presence of initial conditions,” *IEEE Transactions on Automatic Control*, vol. 17, no. 6, pp. 394–396, 1972.
- [13] A. Medvedev, “Continuous least-square observers with applications,” *IEEE Transactions on Automatic Control*, vol. 41, no. 10, pp. 1530–1537, 1996.
- [14] A. Medvedev and H. T. Toivonen, “Directional sensitivity of continuous least-squares state estimators,” *Systems & Control Letters*, vol. 59, pp. 571–577, 2010.
- [15] A. E. Pearson and F. C. Lee, “On the identification of polynomial input-output differential systems,” *IEEE Transactions on Automatic Control*, vol. 30, no. 8, pp. 778–782, 1985.
- [16] H. A. Preisig and H. W. T. Rippin, “Theory and application of the modulating function method–I. Review and theory of the method and theory of the spline-type modulating functions,” *Computers and Chemical Engineering*, vol. 17, no. 1, pp. 1–16, 1993.
- [17] J. Reger and J. Jouffroy, “On algebraic time-derivative estimation and deadbeat state reconstruction,” *IEEE Int. Conf. on Decision and Control*, Shanghai, China, 2009.
- [18] G. P. Rao and H. Unbehauen, “Identification of continuous-time systems,” *IEE Control Theory and Applications*, vol. 153, no. 2, pp. 185–220, 2006.
- [19] D. C. Saha and G. P. Rao, “Identification of lumped linear systems in the presence of unknown initial conditions via Poisson moment functionals,” *Int. Journal of Control*, vol. 31, no. 4, pp. 637–644, 1980.
- [20] D. C. Saha and G. P. Rao, “A general algorithm for parameter identification in lumped continuous systems – the Poisson Moment Functional approach,” *IEEE Transactions on Automatic Control*, vol. 27, no. 1, pp. 223–225, 1982.
- [21] M. Shinbrot, “On the analysis of linear and nonlinear systems,” *Trans. ASME*, vol. 79, no. 3, pp. 547–542, 1957.
- [22] E. D. Sontag, *Mathematical Control Theory (2nd ed.)*. Springer, 1998.