# Estimation problems for a class of impulsive systems

Lotfi Belkoura, Thierry Floquet, Kaouther Ibn Taarit, Wilfrid Perruquetti, Yuri Orlov

# Estimation problems for a class of impulsive systems

L. Belkoura[1,*], T. Floquet[2], K. Ibn Taarit [3],W. Perruquetti[2], Y. Orlov[4]

[1] *LAGIS (CNRS UMR 8146) & INRIA Alien Project, Université Lille 1, France,*
[2] *LAGIS (CNRS UMR 8146) & INRIA Alien Project, Ecole Centrale de Lille, France,*
[3] *ACS Research Center, Ecole nationale d'Ingénieur de Tunis, Tunisia,*
[4] *CICESE Research Center, Electronics and Telecom Dpt., San Diego, USA.*

## SUMMARY

This paper deals with on-line identification of continuous-time systems subject to impulsive terms. Using a distribution framework, a scheme is proposed in order to annihilate singular terms in differential equations representing a class of impulsive systems. As a result, an on-line estimation of unknown parameters is provided, regardless of the switching times nor of the impulse rules. Numerical simulations of physical processes with noisy data are illustrating our methodology and results. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS:   Impulsive systems, Hybrid systems, Identification.

## 1. INTRODUCTION

The primary concern of the paper is on-line identification of hybrid systems the underlying model of which is viewed as a continuous-time system subject to impulsive terms. A motivating example behind the present investigation comes from modeling a biped robot whose dynamics are governed by ordinary differential equations at any time instant but the ones when a leg of the biped robot hits the ground, thereby resulting in an instantaneous change of the leg velocity. As the ground surface does not permit a perfect measurement, on-line identification of parameters of such a system, including identification of impact time instants, represents a challenging problem. Many other examples of this particular class of hybrid system (impulsive) can be found in [5], and for most of them, observability, parameters and state estimation are challenging problems since active mode and continuous state have to be estimated simultaneously within a finite time interval. In recent works such as [1], [6], [13], the estimation of the switching times for some classes of hybrid systems (mainly for switched ones) has been related to observability issues, where the problem consists in recovering, from available measurements, the state of the system and/or the switching signal, and eventually the switching time. Usually, the hybrid observer consists of two parts: an index estimator of

---

*Correspondence to: Lotfi Belkoura, Université Lille1, UFR IEEA Bat P2, 59650 Villeneuve d'Ascq, France, lotfi.belkoura@univ-lille1.fr*

the current active sub-model and a continuous observer that estimates, asymptotically in most cases, the continuous state of the hybrid system.

Even though considerable efforts have been undertaken on various aspects of parameters estimation, there are still many open problems due to the specific structure of the underlying models. For hybrid discrete time systems, an overview of the existing approaches, mainly based on classification, bayesian or bounded error procedures, can be found in [10]. In the present paper, an annihilation procedure for the singular terms of the underlying continuous time models is proposed in a deterministic framework. By solving a generalized eigenvalue problem, this procedure leads to the on-line and non asymptotic estimation of the unknown parameters to be identified, and a simultaneous estimation of the switching times, when available, is also proposed. The approach considered here takes root in the work developed in [7] for parameter identification of continuous linear systems, using algebraic tools (differential algebra, module theory and operational calculus). It results in finite time estimates given by explicit algebraic formulae that can be implemented in a straightforward manner using standard tools from computational mathematics. Recently, those results have been also extended to the problems of state, parameter and unknown input (such as faults) for other classes of continuous time systems in [2], [4], [8], [9], [11], [12].

The paper is organized as follows. Due to the occurrence of nonsmooth dynamics, a distribution framework is adopted and the required notions are recalled in Section 2. The class of impulsive systems considered in this paper is given in Section 3, and Section 4 describes the proposed approaches for finite time estimation of parameters and switching times. The obtained results are illustrated in Section 5 by numerical simulations of a bouncing ball, a mechanical system with dry friction, and a thermostat with an incomplete description of the switching rule.

## 2. MATHEMATICAL TOOLS

We recall here some standard definitions and results from distribution theory [15, 17], and fix the notations to be used in the sequel. The space of $C^\infty$-functions having compact support in an open subset $\Omega$ of $\mathbb{R}$ is denoted by $\mathcal{D}(\Omega)$, and $\mathcal{D}'(\Omega)$ is the space of distributions on $\Omega$, i.e., the space of continuous linear functionals on $\mathcal{D}(\Omega)$. For $u \in \mathcal{D}'$, $\langle u, \varphi \rangle$ denotes the real number which linearly and continuously depends on $\varphi \in \mathcal{D}$. This number is defined as $\langle u, \varphi \rangle = \int_{-\infty}^{\infty} f.\varphi$ for a locally Lebesgue integrable function $u = f$. The support of a distribution $u$, denoted as supp $u$, is defined as the complement of the largest open subset of $\Omega$ in which the distribution $u$ vanishes.

For the Dirac distribution $u = \delta$ and its derivative $\dot{u} = \dot{\delta}$, the functional is defined as $\langle u, \varphi \rangle = \varphi(0)$ and $\langle \dot{u}, \varphi \rangle = -\dot{\varphi}(0)$ respectively. More generally, every distribution is indefinitely differentiable, by virtue of its definition:

$$\langle \dot{u}, \varphi \rangle = - \langle u, \dot{\varphi} \rangle, \forall \varphi \in \mathcal{D}(\Omega). \tag{1}$$

For notational convenience, we shall denote $\dot{u}$ or $u^{(1)}$ the distributional derivative of $u$, while for a function $u$, $\frac{du}{dt}$ will stand for the distribution stemming from the usual derivative (function) of

$u$ defined almost everywhere. Throughout this paper, functions (locally Lebesgue integrable) are considered through the distributions they define. Hence, if $u$ is a continuous function except at a point $a$ with a finite jump $\sigma_a$, one can easily show that its distributional derivative writes $\dot{u} = du/dt + \sigma_a \delta_a$, where $\delta_a$ is the Dirac distribution located at $\{a\}$. This result can be generalized to arbitrary derivation orders and discontinuity points as follows: let $\{t_\nu\}$ be an increasing sequence of points that are finite in number in every finite interval. Assume also that both left-hand and right hand derivatives $\frac{d^p u}{dt^p}(t_\nu)$ exist and denote the corresponding jump $\sigma_\nu^p = \frac{d^p u}{dt^p}(t_\nu+) - \frac{d^p u}{dt^p}(t_\nu-)$. Then one has:

$$u^{(p)} = \frac{d^p u}{dt^p} + \sum_\nu \sum_{k=0}^{p-1} \sigma_\nu^{p-1-k} \delta_{t_\nu}^{(k)}. \tag{2}$$

When rewritten in a distributional sense, the class of ordinary differential equations, we shall encounter in the sequel will always exhibit such singular terms. Inputs and outputs involved in such equations will be viewed as elements with support contained in $[0, \infty)$, and the singularities, stemming from the origin $t_0 = 0$, will be gathered into a single distribution denoted $\psi_0$, of order $n - 1$ and support $\{0\}$. With a slight abuse of language, the latter distribution $\psi_0$ will be referred to as the initial condition term.

Another useful result we shall exploit in the sequel is based on properties derived from the multiplication of distributions. Although this operation is not always defined for arbitrary distributions, it turns out that multiplication of two distributions (say $\alpha$ and $u$) is always well-defined when at least one of the two terms (say $\alpha$) is a smooth function. By definition:

$$\langle \alpha\, u, \varphi \rangle = \langle u, \alpha\varphi \rangle. \tag{3}$$

The previous definitions of derivation (1) and multiplication (3) also allow to transform terms of the form $\alpha\, u^{(n)}$ into linear combinations of derivatives of products $\alpha^{(k)}\, u$. Indeed, for any $\varphi \in \mathcal{D}(\Omega)$:

$$\langle \alpha\dot{u}, \varphi \rangle = \langle \dot{u}, \alpha\varphi \rangle = -\left\langle u, \dot{(\alpha\varphi)} \right\rangle = -\langle u, \dot{\alpha}\varphi + \alpha\dot{\varphi} \rangle = \left\langle -\dot{\alpha}u + \dot{(\alpha u)}, \varphi \right\rangle,$$

yelding

$$\alpha\dot{u} = -\dot{\alpha}u + \dot{(\alpha u)}.$$

This "reversed" Leibniz rule can be easily extended to any derivation order $n$ by:

$$\alpha\, u^{(n)} = \sum_{k=0}^{n} (-1)^{2n-k} C_n^k\, w_k^{(n-k)}, \quad w_k := \alpha^{(k)}\, u. \tag{4}$$

Note that successive integrations of the latter equation result in nothing but the integration by parts formula with available data $w_k = \alpha^{(k)}\, u$ when $u$ is obtained *via* sensor measurements. In a noisy context, such integrations can be however advantageously replaced by any appropriate filter of relative degree larger than $n$, also avoiding any measurement of the time derivatives.

The following Theorem is the key result from which the parameters (including the switching times) of the class of systems under study can be identified.

**Theorem 1.** *[15] If $u$ has a compact support $K$ and is of order $r$ (necessarily finite)[†], $\alpha\, u = 0$ whenever $\alpha$ and its derivatives of order $\leq r$ vanish on $K$.*

In this paper, we shall be mainly concerned with the annihilation of singular distributions, and as an illustration of the previous Theorem, one has for every smooth function $\alpha$:

$$\alpha\, \delta_\tau = \alpha(\tau)\, \delta_\tau, \tag{5}$$

$$\alpha\, \delta_\tau^{(r)} = 0 \qquad \forall \alpha \text{ s.t. } \alpha^{(k)}(\tau) = 0, \quad k = 0, \ldots, r. \tag{6}$$

We complete this introductory section with some well-known definitions and results from the convolution products, and as usual, denote $\mathcal{D}'_+$ the space of distributions with support contained in $[0, \infty)$. It is an algebra with respect to convolution with identity $\delta$. For $u, v \in \mathcal{D}'_+$, this product is defined as $\langle u * v, \varphi \rangle = \langle u(x).v(y), \varphi(x + y) \rangle$, and can be identified with the familiar convolution product $(u * v)(t) = \int_0^\infty u(\theta)v(t - \theta)d\theta$ in case of locally bounded functions $u$ and $v$. Derivation, integration and translation can also be defined from the convolutions $\dot{u} = \dot{\delta} * u$, $\int u = H * u$, $u(t - \tau) = \delta_\tau * u$, where $H$ is the familiar Heaviside step function. As for the supports, one has for $u, v \in \mathcal{D}'_+$:

$$\operatorname{supp} u * v \subset \operatorname{supp} u + \operatorname{supp} v, \tag{7}$$

where the sum in the righthand side is defined by $A + B = \{x + y \, ; \, x \in A, y \in B\}$. Finally, with no danger of confusion, we shall denote $u(s)$, $s \in \mathbb{C}$, the Laplace transform of $u$ (in a distribution setting).

## 3. IMPULSIVE MODELS

The general class of systems considered in this paper consists of $n$th order nonlinear ordinary differential equations subject to impulsive righthand sides, described in a distributional framework, with input $u$ and output $y$, both in $\mathcal{D}'_+$:

$$\sum_{j=0}^n a_j\, g_j^{(j)}(u, y) = \psi_0 + \sum_{j=1}^\infty b_j\, \delta(t - t_j). \tag{8}$$

Here, $a_j$ are constant and unknown coefficients (except for $a_n := 1$) and $g_j^{(j)}(u, y)$ are known and possibly nonlinear functions of the input and output measurements. The function $g_j(u, y)$ is assumed to be continuous, and the distribution $\psi_0$, of order $n - 1$ and support $\{0\}$, gathers all the singular terms derived from the initial conditions as described in the previous section. Throughout this paper, we shall assume the existence of a non zero smallest dwell time avoiding accumulation points and Zeno phenomena. Finally, the unknown coefficients $b_j$ are state dependent, since they are generally related to the possible state deviations at the impact or switching times $t_j$, i.e.:

$$b_j = b_j\left(t_j, u(t_j), \frac{du}{dt}(t_j\pm), \ldots, \frac{d^n u}{dt^n}(t_j\pm), y(t_j), \frac{dy}{dt}(t_j\pm), \ldots, \frac{d^n y}{dt^n}(t_j\pm)\right). \tag{9}$$

---

[†]The order of a distribution $u$ is the smallest integer $m$ such that: $\exists C > 0$ such that for any smooth function $\varphi$, one has $|\langle u, \varphi \rangle| \leq C \sup_{0 \leq i \leq m} \|\varphi^{(i)}\|_\infty$ (for example, $\delta^{(r)}$ is of order $r$).

It is worth noticing that the switching time sequence $\{t_j\}$ being *a priori* unknown, the proposed model may not be appropriate for some analysis issues (for example stability), nor for simulation purposes. However, the description of the right hand side of equation (8) as singular functions (or distributions) is of great interest since, on one hand, straightforward methods will be derived for the parameters (and possibly switching times) estimation, and on the other hand, the discrete transitions or switching rules need not to be known. The latter point is of particular interest when one knows that transition rules such as impacts of mechanical systems on irregular surfaces are not easy to model. The following example illustrates how, using an *a priori* unknown switching time sequence, an impulsive model can be obtained by simple differentiation of the underlying process, even in the presence of dynamic uncertainties.

### 3.1. Example: An impulsive model for a thermostat

Consider the academic control problem of heating a room using a thermostat [18]. In this example, we do not have an exact model of how the thermostat works but it is only known that the thermostat turns the radiator on when the temperature is between $\min_1$ and $\min_2$ and turns the radiator off when the temperature is between $\max_1$ and $\max_2$. This heating system can be described by the non deterministic hybrid automaton in Figure 1, where $y$ denotes the temperature and $a > 0$, $b > \max_1$.
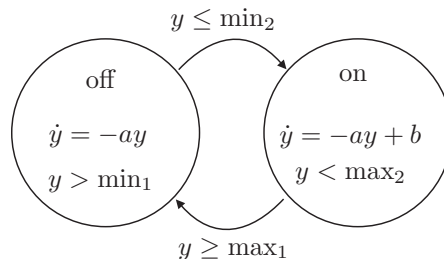


Figure 1. The thermostat as a hybrid system.

Due to uncertainties in the radiator dynamics, it is not possible to solve the system equations. However, given the *a priori* unknown switching sequence $\{t_j\}_{j=1}^{\infty}$ from one position to another, and assuming that the process is initially in configuration off with initial state $y_0$, the behavior of such a process can be described by:

$$\dot{y} = -ay + y_0\delta_0 + b\sum_{j=0}^{\infty}\chi_{[t_{2j+1},t_{2j+2}]}(t),\tag{10}$$

where $\chi_{[t_i,t_j]}(t)$ denote the characteristic function of the interval $[t_i,t_j]$, (i.e. $\chi_{[t_i,t_j]}(t) = 1$ if $t \in [t_i,t_j]$, and 0 elsewhere). Since, in a distribution setting, $\dot{\chi}_{[t_i,t_j]}(t) = \delta(t-t_i) - \delta(t-t_j)$, one obtains by derivation of (10):

$$\ddot{y} + a\dot{y} = y_0\dot{\delta}_0 + b\sum_{j=0}^{\infty}\delta(t-t_{2j+1}) - \delta(t-t_{2j+2}) = y_0\dot{\delta} - \sum_{j=1}^{\infty}(-1)^j b\,\delta(t-t_j),\tag{11}$$

which corresponds to the model description given in (8). Note that from this simple linear system, we can also derive an operational description based on the more familiar Laplace

transform, recalling the correspondence $\dot{y} \leftrightarrow sy(s) - y_0$, $\delta^{(r)} \leftrightarrow s^r$, and $\delta(t - t_i) \leftrightarrow e^{-t_i s}$. One therefore gets from (11):

$$y(s^2 + as) = y_0\, s - \sum_{j=1}^{\infty} (-1)^j b\, e^{-t_j s}. \tag{12}$$

The main motivation for using a distribution framework lies in the ability to consider the multiplication of equation (8) with a large class of candidate functions. In the Laplace setting, the equivalence with such a multiplication is not easy to handle unless restricted to multiplication with polynomial functions (for the linear identification problem considered in [7, 8, 9], this is done using the derivation with respect to $s$, by virtue of the property $dy(s)/ds \leftrightarrow -t\, y(t))$.

## 4. STRUCTURE OF THE ESTIMATION PROBLEMS

Two principal structures for estimation issues will be developed in this section, and both are derived from a straightforward application of the results of Section 2. They are mainly based on two steps consisting in: (i) a multiplication with smooth functions $\alpha_k$, $k = 1, 2$ and (ii) a convolution with an appropriate filter $h$. From the multiplication properties, given in (6), we first get that multiplying equation (8) with arbitrary smooth functions $\alpha_k$, $k = 1, 2$ such that $\alpha_k^{(i)}(0)$, $i = 0 \ldots, n - 1$, cancels the initial conditions, yielding:

$$\alpha_k \times \sum_{j=0}^{n} a_j\, g_j^{(j)}(u, y) = \sum_{j=1}^{\infty} b_j\, \alpha_k(t_j)\delta(t - t_j), \quad k = 1, 2. \tag{13}$$

Note that using this operation, the unknown switching times $t_j$ also appear as arguments of the known functions $\alpha_k$. As an equality of singular distributions, we are not able to recover these arguments from the previous equations. Hence, the next step consists in a convolution with a filter $h$, yielding an equality of functions that is more appropriate for measurements manipulations:

$$h * \left[ \alpha_k \times \sum_{j=0}^{n} a_j\, g_j^{(j)}(u, y) \right] = \sum_{j=1}^{\infty} b_j\, \alpha_k(t_j)h(t - t_j), \quad k = 1, 2. \tag{14}$$

From this formulation, causal relations (avoiding any measurement derivatives) will be obtained if the chosen function $h$ is of class $\mathcal{C}^n$, which is equivalent for a linear filter $h(s)$ to have a relative degree $\geq n$. Furthermore, in order to get a sequential estimation of the switching times, the main *a priori* assumption we shall need in the sequel is a lower bound estimate of the smallest dwell time. If the support of $h$ lies within the smallest dwell time, i.e. supp $h \subset (0, \min_j(t_{j+1} - t_j))$, then the main idea consists in getting the switching instants $t_j$ from the comparison of the left hand sides of (14) for $k = 1, 2$. For instance, if $\alpha_2 = t\,\alpha_1$, then the ratio of the left hand sides of (14), when defined, yields precisely, and in real time, the switching times $t_j$. More generally, let us introduce, with slight abuse of definition, the switching function $\zeta(t)$ as:

$$\zeta(t) = \sum_{j=1}^{\infty} \zeta_j\, \chi_{[t_j, t_{j+1}]}(t), \qquad \zeta_j = \alpha_1(t_j)/\alpha_2(t_j). \tag{15}$$

By virtue of the support of $h$, the local equality we have just obtained from such a comparison can be extended to the whole real line as:

$$\left( h * [\alpha_1 \times \sum_{j=0}^{n} a_j \, g_j^{(j)}(u,y)] \right) = \zeta(t) \left( h * [\alpha_2 \times \sum_{j=0}^{n} a_j \, g_j^{(j)}(u,y)] \right). \tag{16}$$

Such an extension has been performed only by letting $0 = \zeta(t)0$ outside the supports of $h(t - t_j)$. It should be pointed out that following the approaches developed in [7], algebraic (and hence non asymptotic) relations are obtained, unlike that of most estimation algorithms where asymptotic relations are only due. Moreover, the state dependent terms $b_j$ are canceled by the proposed method, avoiding any *a priori* knowledge of the switching rules.

### 4.1. Example of the thermostat (continued)

Figure 2 (top) shows a simulated trajectory of the thermostat described in Section 3.1, with the numerical values $(\min_1, \min_2) = (0, 1)$, $(\max_1, \max_2) = (4, 5)$, $a = -1.3$, $b = 5$, and an additive normally distributed noise of amplitude $3\% \max(y)$. The switching sequence $\{t_j\} = \{0, 2, 5, 8, 10, 13, 15, 17, 20\}$ has been fixed to satisfy the prescribed intervals. Based on the measurement $y$ and the known parameter $a$, Figure 2 (bottom) illustrates a realization of the left hand side of equation (14), which reduces in this case to:

$$h * [\alpha_k \times (\ddot{y} + a\dot{y})], \quad k = 1, 2. \tag{17}$$

The adopted filter and functions are given by $h = \left( \left( 1 - e^{-1.5\,s} \right)/s \right)^4$, $\alpha_1 = \sin(2\,t)$, and $\alpha_2 = \sin^2(2\,t)$. Hence, on each (unknown) interval $(t_j, t_{j+1})$, the formed functions in Figure 2 (bottom) are linked by the relation:

$$(h * [\alpha_2 \times (\ddot{y} + a\dot{y})])(t) = \sin(2\,t_j) \, (h * [\alpha_1 \times (\ddot{y} + a\dot{y})])(t), \quad t \in (t_j, t_{j+1}). \tag{18}$$

Finally, and for the sake of simplicity, the initial condition $y(0)$ has been set to 0 to avoid its cancellation by using the additional constraint $\ddot{\alpha}_1(0) = 0$.

### 4.2. Robustness and identifiability issues

The estimations algorithms we shall use in the next Sections 4.3 and 4.4 are both derived from equation (16), and it is worth noticing that the robustness of such procedures in a noisy context will highly depend on the filtering capabilities, within the reduced dwell time, of the function $h$, as well as on the choice of the candidate functions $\alpha_k$ for the multiplication. The combination of compact support filters and pre-multiplication by smooth functions is a non standard problem in signal processing, and selecting the optimal pair $\alpha_k$, $k = 1, 2$, and the filter $h$ is still an open problem. In the remaining part of this section, some comments and hints for the choice of the multiplicative functions is provided, and a discussion on the identifiability of the switching sequence is introduced.

First, getting the times impact $t_j$ from the coefficients $\zeta_j$ involved in (15) assumes a one to one map between $\zeta(t)$ and $t$, which implies that $\zeta(t)$ is a possibly unbounded function. In turn, this also means that the multiplicative functions $\alpha_1$ and/or $\alpha_2$ are also unbounded, yielding
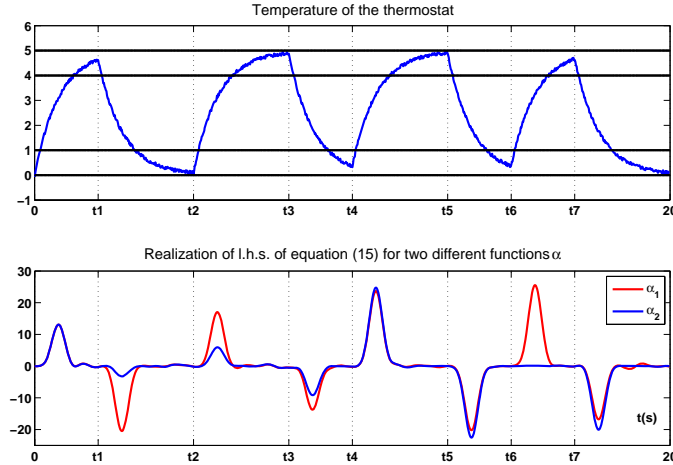
Figure 2. Illustration of equation (16) for the thermostat example.

non trivial robustness problems in case of noisy data. To illustrate this aspect, consider for instance a model description in (8) with $g(u, y) = y$, and assume we want an explicit time impact estimation by setting in (15) $t_j = \zeta_j$. By virtue of (15 second equation), this implies that $\alpha_1(t) = t\alpha_2(t)$. Hence, one has to consider from (16) (see also the detail algorithm of the next section) filtering by $h$ of terms of the form $\alpha_2(t)\,t\,y$, yielding non trivial robustness issues due to the multiplication by $t$ of the noisy data $y$. As illustrated in our examples, this unboundness issue can be however circumvented using for instance bounded and periodic functions $\alpha_i$, thus resulting in an estimation modulo a known fixed period: in that case, the real estimation of the the impact times can be easily deduced from the actual running time.

Next, from the properties of the support of a convolution product (7), or alternatively from the right hand side of (14), one can show that the estimation problem of the switching sequence is not consistent for all $t > 0$. In other words, and as illustrated in Figure 2 (bottom), both formed functions vanish locally, yielding $0 = \zeta(t)0$ outside the latter intervals and leading to a local loss of identifiability of $\zeta$. This drawback may require the use of *a priori* information (threshold) testing the consistency of (14). This issue is discussed in Section 4.3 where the estimation of both parameters and switching times is considered.

Equation (16) is obviously nonlinear with respect to both the switching function $\zeta$ and the unknown coefficients $a_i$, and is clearly not sufficient in case of a simultaneous estimation problem. However, additional relations can be obtained if the filter $h$ is replaced by a set of $h_i$ filters sharing the same previous condition on the support. Depending on the chosen factorization, two different structures of the estimation problem are derived; the first one, presented in Section 4.3 aims at providing a simultaneous estimation of the switching time function $\zeta$ and the unknown coefficients $a_i$. In the second approach detailed in Section 4.4, only parameters estimation is considered, which allows to avoid the identifiability problem of the switching function $\zeta$. The formulations, advantages and limitations of each approach are discussed in the rest of this section.

### 4.3. Simultaneous switching time and parameters estimation

Denoting $\Theta = (1, a_{n-1}, \ldots, a_0)^T$ the set of unknown parameters, (16) reads in a matrix form (where we dropped the arguments of $g$ for compactness):

$$\left[ h * (\alpha_1 \times g_n^{(n)}), \cdots, h * (\alpha_1 \times g_0^{(0)}) \right] \Theta = \zeta(t) \left[ h * (\alpha_2 \times g_n^{(n)}), \cdots, h * (\alpha_2 \times g_0^{(0)}) \right] \Theta. \quad (19)$$

Considering equation (19) for different filters $h_i$, $i = 0, \ldots, n$ lead to the following $(n+1 \times n+1)$ eigenproblem structure involving both unknown switching times and parameters:

$$[A_1(u,y) - \zeta(t) A_2(u,y)] \Theta = 0, \quad (20)$$

$$A_k(u,y) = h_i * [\alpha_k \times g_j^{(j)}(u,y)], \quad (21)$$

$$k = 1, 2; \quad i, j = 0, \ldots, n.$$

It should be stressed that the matrices $A_k(u,y)$, $k = 1, 2$, are only depending on inputs $u$ and outputs $y$ measurements, as detailed in the following algorithm describing the realization steps of their entries (the subscripts are omitted for the sake of readability):

1. From the inputs/outputs measurements, form the functions $g_j(u,y)$ according to the model (8).
2. For an arbitrary smooth function $\alpha$ such that $\alpha^{(i)}(0)$, $i = 0 \ldots, n-1$, form the functions $w_k = \alpha^{(k)} g$, and develop $\alpha g^{(n)}$ according to (4) to get:

$$\alpha \, g^{(n)} = \sum_{k=0}^{n} (-1)^{2n-k} C_n^k \, w_k^{(n-k)}. \quad (22)$$

3. Apply the filter $h$ to the previous equation. By virtue of the property $h * w^{(n-k)} = h^{(n-k)} * w$, one gets:

$$h * (\alpha \, g^{(n)}) = \sum_{k=0}^{n} (-1)^{2n-k} C_n^k \, h^{(n-k)} * w_k. \quad (23)$$

It is clear from the latter equation, whose Laplace transform reads as

$$\mathcal{L}(h * (\alpha \, g^{(n)})) = \sum_{k=0}^{n} (-1)^{2n-k} C_n^k \, s^{n-k} h(s) w_k(s), \quad (24)$$

that strictly causal relations are obtained if the relative degree of the filter $h(s)$ is larger than $n$, avoiding derivative measurements. Moreover, this allows us to consider real time implementations of the estimation algorithms.

The estimation problem has been therefore formulated in terms of a generalized eigenvalue problem where, at each time $t$, the switching function $\zeta(t)$ corresponds to one of the generalized eigenvalues of the problem (20), while the unknown coefficients are derived from the corresponding normalized eigenvector. The main interest of such a formulation lies in the simplicity of its numerical implementation, since many fast and efficient numerical algorithms are available for eigenvalue problems. As mentioned in the previous section, its principal limitation comes from the local loss of identifiability of $\zeta(t)$. This loss of consistency also

reads $0 = \zeta(t)\,0$ in (19), yielding a simultaneous rank deficiency of $A_1$ and $A_2$ in (20). A suggested method to avoid this drawback may be based on some informative function such as:

$$\gamma(t) = \sqrt{(\det A_1(t))^2 + (\det A_2(t))^2},\qquad(25)$$

who aims at selecting the appropriate structure of the estimation problem. More precisely, and given an *a priori* threshold $\varepsilon$, the eigenvalue problem will be considered as not consistent for $\gamma(t) \le \varepsilon$. Hence, one can only solve the linear structures $A_i\Theta = 0$ and estimate the coefficients $a_i$ (recall that the vector $\Theta$ is normalized). If $\gamma(t) \ge \varepsilon$, not both $A_1$ and $A_2$ are rank deficient and the eigenvalue problem $(A_1 - \zeta A_2)\Theta = 0$ can be solved to get both parameters and the switching time.

### 4.4. Parameter estimation regardless of the switches

In this section where only the parameters $a_i$ are to be estimated, the consistency issues of the previous section are avoided. An alternative representation of the estimation problem, involving matrices with only 2 columns, can be simply obtained using a different factorization of (16) as follows:

$$[M_n(u,y) + a_{n-1}M_{n-1}(u,y) + \ldots + a_0 M_0(u,y)]\,\xi(t) = 0,\qquad(26)$$

$$M_k(u,y) := h_i * [\alpha_j \times g_k^{(k)}(u,y)],\qquad(27)$$

$$i = 1,2\ldots,\, j = 1,2,\; k = 0,\ldots,n.$$

with $\xi(t) := (1, -\zeta(t))^T$. Although the eigenstructure is lost in case of more than one unknown coefficient $a_i$, its fundamental advantage lies in the fact that none of the matrices $M_k(u,y)$ defined in (27) is structurally zero on any interval. Note also that, as in the previous case, no measurement derivative has to be computed.

Therefore, one can consider the $a_i$'s estimation from (26) regardless of the switching time function $\xi(t)$ (i.e. even when $\zeta(t)$ is not identifiable), by considering for instance the set of nonlinear equations:

$$\det \sum_{i=0}^{n} a_i Q_{i,j}(u,y) = 0, \quad j = 1,2\ldots,\qquad(28)$$

where $Q_{i,j}$ are $2 \times 2$–matrices derived from the lines of $M_k(u,y)$ in (27). In the general case, this problem is however not trivial unless it is transformed into a least square problem with redundancy in the vector to be estimated, (which is given here by $\bar{\Theta} = (a_1, a_2, a_1^2, a_1 a_2, \ldots)^T$). Note also that in case of a single unknown parameter $a_0$, the problem (26) is still an eigenvalue problem and can be solved as such. To illustrate the capability of identifying the parameters regardless of the identifiability of $\zeta(t)$, let us recall that the generic eigenvalue problem $(\lambda I - A(t))v = 0$ is well posed even if $(\lambda I - A(t)) = 0$, and in such a case, any vector $v$ is an eigenvector.

To conclude this section, let us mention that while solving an eigenproblem, the eigenvalue corresponding to the actual unknown quantity has to be selected. The first approach would consist in the resolution of an augmented system of equations by adding more filters $h_i$ and hence forming a matrix pencil. However this situation "has the awkward feature that most matrices have no eigenvalues at all, whilst for those that do, an infinitesimal perturbation will

in general remove them" [14]. An alternative approach could be based on the pseudo-spectra analysis which consists in defining the $\epsilon-$pseudospectra as the set

$$\Lambda_\epsilon = \left\{ z \in \mathbb{C} : \|A + zB\|^+ > \epsilon^{-1} \right\}. \tag{29}$$

for rectangular matrices $A$ and $B$ (see e.g [14] and the references therein). However, this approach is not appropriate for on-line applications, and we adopt here a simpler technique, based on the *a priori* stationarity assumption of the unknown parameters: the selected parameters correspond to the eigenpair of the square pencil that minimizes the norm of the rectangular one.

## 5. APPLICATIONS

### 5.1. Impulsive mechanical systems: Bouncing ball and rocking block

The examples given here are (autonomous) models of a bouncing ball and a rocking block rotating around two pivots. The outputs correspond to the position of the ball and the angle that the block makes with the vertical, respectively [16]. The dynamical models of those systems are described by the following equations:

$$\frac{d^2y}{dt^2} = f(y) = \begin{cases} -m\,g & \text{(bouncing ball)} \\ \frac{1}{\alpha}\sin(\alpha\,(1 - y\,\text{sgn}(y))) & \text{(rocking block)} \end{cases} \tag{30}$$

$$\frac{dy}{dt}(t_k^+) = \gamma \frac{dy}{dt}(t_k^-), \quad k = 1, 2, \ldots, \tag{31}$$

where $\alpha$ is a constant coefficient depending on the size of the block, $\gamma$ is the coefficient of restitution of the velocity ($\gamma < 0$ for the bouncing ball and $\gamma > 0$ for the rocking block), and $t_k$ are the impact instants. Interpreting $y$ in the distribution sense with the formula of the jumps, and denoting $\sigma_k = \frac{dy}{dt}(t_k+) - \frac{dy}{dt}(t_k-)$ the jumps in the velocities, equation (30) reads as:

$$\ddot{y} - f(y) = \psi_0 + \sum_j \sigma_j \delta(t - t_j), \qquad \psi_0 = y_0 \dot{\delta} + \dot{y}_0 \delta, \tag{32}$$

which corresponds to the model described by (8) with $g_2(y, u) = y$, $a_1 = 0$, $g_1(y, u) = 0$, $a_0 = 1$, $g_0(y, u) = f(y)$, and $b_j = \sigma_j$ for the state dependent coefficient. It is worth noticing that, except if some additional external input is added, the above formulation in a distribution setting is only valid on a finite time interval since for very large time the impact times give rise to an accumulation point. The proposed procedure is hence applicable only at the beginning of the process evolution. Following the method of the previous section, multiplying this equation with functions $\alpha_k$, $k = 1, 2$ such that $\alpha_k(0) = \dot{\alpha}_k(0) = 0$ cancels the initial conditions, and a convolution with filters $h_i$ results in:

$$h_i * [\alpha_k \times (\ddot{y} - f(y)] = \sum \sigma_j \alpha_k(t_j) h_i(t - t_j), \quad k = 1, 2. \tag{33}$$

Let us focus in this section on the mass estimation for the bouncing ball case. Since only one parameter (the mass) is to be estimated, we adopt the eigenvalue formulation of Section 4.4 with:

$$M_1(y)_{i,j} = h_i * (\alpha_j \times \ddot{y}), \qquad M_0(y)_{i,j} = h_i * (\alpha_j \times g). \tag{34}$$

A simulated trajectory of the bouncing ball, with a moderate random white noise is shown in Figure 3-left, while Figure 3-right shows a generic realization of $M_1(y)$ avoiding any time derivative of the measured signals. A lower bound for the smallest dwell time has been fixed to $T = 2s$, and the adopted filters $h_i$ and multiplicative functions $\alpha_i$ are given by:

$$\alpha_1 = \sin^2(t), \quad \alpha_2 = \alpha_1 \sin(t), \quad h_1 = \left( \frac{1 - e^{-sT/3}}{s} \right)^3, \quad h_2 = \frac{1}{(1 + 0.25s)^4}, \quad (35)$$

and are depicted in Figure 4-left. The matrix pencil used for the eigenpair selection has been set using the additional filters $h_3(s) = sh_1(s)$ and $h_4(s) = sh_2(s)$. Note that due to their decay rate, the filter $h_2$ and its derivative $h_4$ have been also assimilated to bounded supported functions. Finally, Figure 4-right shows the on line bouncing ball's mass estimation, regardless of the switching rules (restitution law of the velocity) and impact instants.
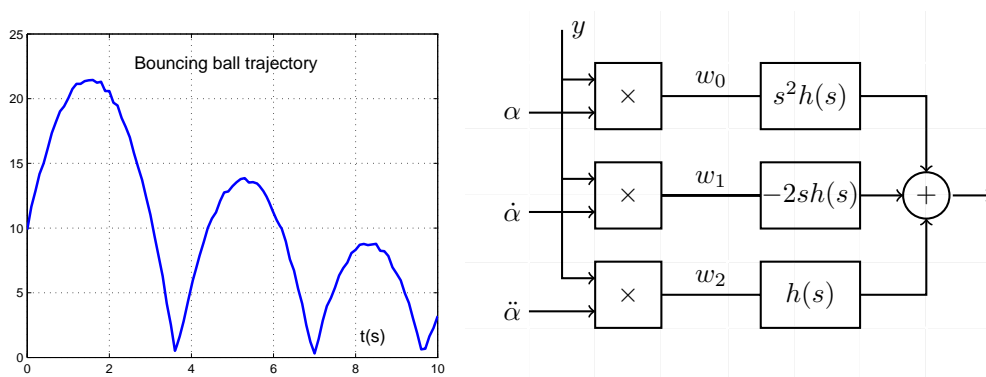


Figure 3. Bouncing ball trajectory (left) and realization scheme of $M_1(y)$ in (34) (right).
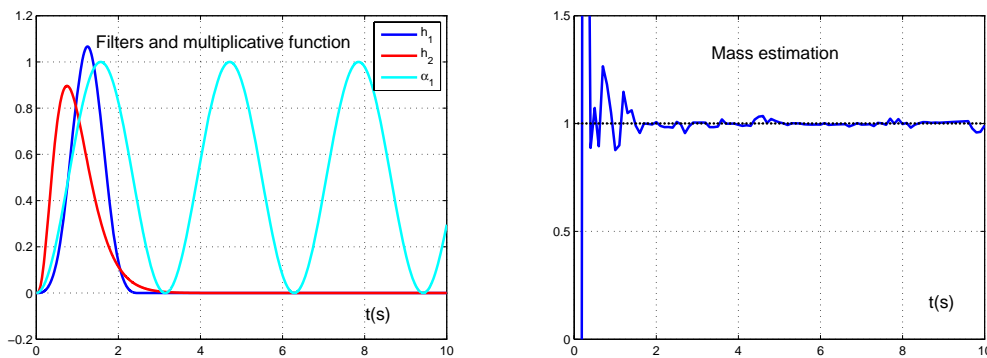


Figure 4. Adopted filters and multiplicative functions (left) and ball's mass estimation (right).

### 5.2. Mechanical systems with dry friction

The aim of this nonlinear example is to show that the impulsive terms in the proposed formulation (8) can be obtained by simple differentiation of the underlying process. Let $H(.)$ denote the Heaviside step function. Assuming zero initial conditions for simplicity, the simple pendulum subject to asymmetrical dry friction as well as a forcing term $u$ can be described by equation (36) leading to (37) by means of a first order differentiation:

$$\ddot{y} + a_0 \sin(y) = b_1\, H(\dot{y}) - b_2 H(-\dot{y}) + u, \tag{36}$$

$$y^{(3)} + (a_0\, \sin(y) - u)^{(1)} = \sum_j \sigma_j \delta(t - t_j), \tag{37}$$

where $t_j$ denotes the time sequence when the angular velocity is crossing zero (that is when $\text{sign}\left(\frac{dy}{dt}(t_j-)\right) \times \text{sign}\left(\frac{dy}{dt}(t_j+)\right) = -1$), and $\sigma_j = (-1)^{j+r}(b_2+b_1)$ the jumps in the dry friction effects ($r = 0$ or 1 depending on the initial state). Note that although the first order derivation term is affine and not linear with respect to the coefficient $a_0$, we can easily recover the general structure of (8) by considering $s^3(y - u/s^2)$ in the third order derivation. Using the singularities annihilation approach described in this paper, the coefficients $b_i$ and hence the jumps do not need to be known *a priori*, and in order to avoid an accumulation point phenomena, as encountered in the previous example, we shall assume a non trivial input $u$ excluding infinitely many switches between the upper and lower friction value for $\dot{y} = 0$. Here, we are still in the eigenvalue framework of Section 4.4. Figure 5-left shows the simulated trajectories of the pendulum subject to the input $u = 0.12 \sin(0.2\,t) \cos(0.1\,t))$, with an additive random white noise (magnitude of $\pm 4\% \max(y)$). The following filters and multiplicative functions have been selected:

$$\alpha_1(t) = 1(t), \quad \alpha_2(t) = \sin(0.3\,t), \quad h_1 = \left(\frac{1 - e^{-sT/5}}{s}\right)^5, \quad h_2 = e^{-sT/5}h_1, \quad T = 8\,s,$$

and as in the previous example, the matrix pencil is obtained using the successive derivatives of $h_1$ and $h_2$. Denoting $w_i := \alpha^{(i)}y$, $v_i := \alpha^{(i)} \sin(y)$ and $z_i := \alpha^{(i)}u$, the generic terms of the matrices of the eigenvalue problem are given by:

$$M_1(y,u)_{i,j} = s^3 h w_0 - 3s^2 h w_1 + 3shw_2 + shz_0 - hz_1, \tag{38}$$

$$M_0(y,u)_{i,j} = shv_0 - hv_1. \tag{39}$$

Figure 5-right shows the estimation of the parameter $a_0$, regardless of the switching time commutations and rules. For the sake of completeness, and in order to illustrate the previous comments on identifiability issues, Figure 6 shows, in a reduced noise context (magnitude of $\pm 1\% \max(y)$), the results of the $a_0$'s estimation algorithm, together with the corresponding normalized eigenvector corresponding to $\zeta(t)$. Figure 6-left clearly shows random and constant phases that illustrate the alternation of identifiability and non identifiability of the switching times function $\zeta(t)$.
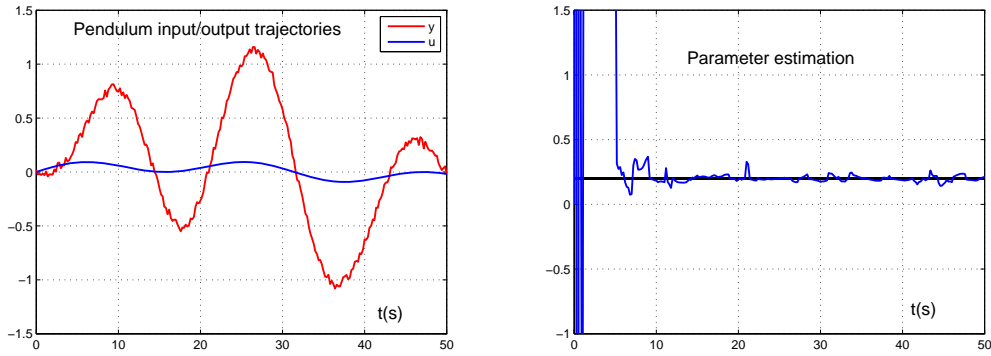
Figure 5. Pendulum trajectories (left) and $a_0$ parameter's estimation (right)
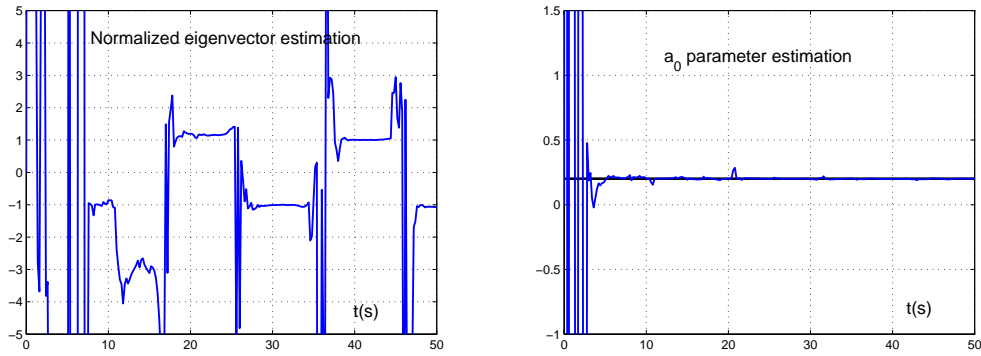


Figure 6. Normalized eigenvector (left) and $a_0$ parameter's estimation (right)

### 5.3. Example of the thermostat (continued)

In this example, we provide an estimation procedure for the mode $a$ involved in the dynamic of the thermostat, and adopt the factorization presented in Section 4.4:

$$
\begin{aligned}
&[M_1(y) + a\, M_0(y)]\, \xi(t) = 0,\\
&M_1(y)_{i,j} = h_i * (\alpha_j \times \ddot{y}),\\
&M_0(y)_{i,j} = h_i * (\alpha_j \times \dot{y}),
\end{aligned}
\tag{40}
$$

with $i = 1, 2$, $j = 1, 2$. The response $y(t)$ and the functions $h_1$, $\alpha_1$ and $\alpha_2$ are those described in Section 4.1, and the additional filters are given by $h_2(s) = h_1(s)\, e^{-0.3\, s}$. Rather than solving an eigenvalue problem, which turned out to be more noise sensitive in this application, we adopt the nonlinear formulation:

$$
\det\left[M_1(y) + a\, M_0(y)\right] = c_0(t) + a\, c_1(t) + a^2\, c_2(t) = 0,
\tag{41}
$$

and estimate, in a least square sense, the linear system derived from the discrete measurements (with sampling period $t_e = 0.02$ s.):

$$
\begin{pmatrix}
c_1(t_e) & c_1(t_e) \\
\vdots & \vdots \\
c_2(k\,t_e) & c_2(k\,t_e)
\end{pmatrix}
\begin{pmatrix}
a \\
a^2
\end{pmatrix}
=
\begin{pmatrix}
-c_0(t_e) \\
\vdots \\
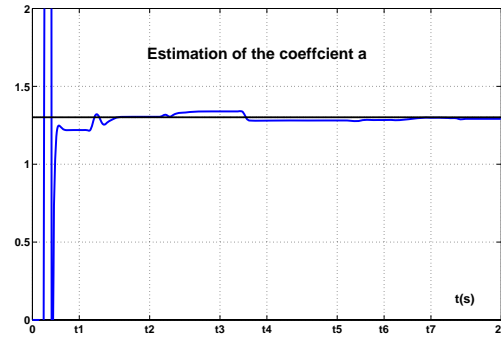-c_0(k\,t_e)
\end{pmatrix}
\tag{42}
$$



Figure 7. Estimate of the dynamic $a$ of the heating process

Figure 7 shows the estimate of the mode $a$, regardless of the unknown switching sequence. It should be stressed that, throughout this paper, an accurate quantification of the noise effects is not an easy task, with regard to the structure of the formed terms, and small bias in the presented identification technique may occur. Note finally that this last example also raises the question of choosing the best approach between a noise sensitive eigenvalue problem and a nonlinear and redundant formulation.

## 6. CONCLUSION

This paper has presented new tools for the on line parameters estimation of (possibly nonlinear) systems with impulsive terms. A unified method has been proposed using simple algebraic techniques which are based on distributions, annihilation of singular terms and integrations. The performance of the proposed procedure has been illustrated in simulation studies. Although these simulations have demonstrated favorable robustness properties against measurement noises, these properties highly depend on the filtering capability of the filters within the reduced dwell time, and on the choice of the structure of the estimation problem. Together with the selection of optimal multiplicative functions, such issues are under active investigations. According to these studies, the capability to estimate parameters and switching times, using non asymptotic methods, may provide new perspectives for real time control procedures.

## REFERENCES

1. M. Babaali and G.J. Pappas, Observability of switched linear systems in continuous time, *Lecture Notes in Computer Science (Hybrid Systems: Computation and Control)*, Vol. 3414, pp. 103-117, 2005.
2. J.-P. Barbot, M. Fliess, T. Floquet, "An algebraic framework for the design of nonlinear observers with unknown inputs", *IEEE Confonference on Decision and Control*, New-Orleans, USA, 2007.
3. A. Bemporad, G. Ferrari-Trecate, M. Morari, "Observability and controllability of piecewise affine and hybrid systems", *IEEE Transactions on Automatic Control*, Vol. 45, pp. 1864–1876, 2000.
4. L. Belkoura, J.-P. Richard, M. Fliess, "Parameters estimation of systems with delayed and structured entries", Automatica, Vol. 45, No. 5, pp. 1117–1125, 2009.
5. B. Brogliato, *Nonsmooth Mechanics*, Second edition, Series: Communications and Control Engineering, Springer London, 1999.
6. S. Chaib, D. Boutat, A. Benali, J.-P. Barbot, "Observability of the discrete state for dynamical piecewise hybrid systems" , *Nonlinear Analysis: Theory, Methods & Applications*, Vol. 63, pp. 423–438, 2005.
7. M. Fliess, H. Sira-Ramírez, "An algebraic framework for linear identification", *ESAIM Control Optim. Calc. Variat.*, Vol. 9, pp. 151–168, 2003.
8. M. Fliess, M., C. Join, H. Sira-Ramírez, "Robust residual generation for linear fault diagnosis: an algebraic setting with examples", *Int. J. of Control*, Vol. 77, pp. 1223–1242,2004.
9. M. Fliess and H. Síra-Ramírez, "Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques", *Continuous- Time Model Identification from Sampled Data*, Springer, 2008.
10. S. Paoletti, A. L. Juloski, G. Ferrari-Trecate, and R. Vidal, "Identification of hybrid systems: a tutorial", *European Journal of Control*, Vol. 13, pp. 242–260, 2007.
11. J. Rudolph, F. Woittennek, "An algebraic approach to parameter identification in linear infinite dimensional systems", *IEEE Mediterranean Conference on Control and Automation*, Ajaccio, France, 2008.
12. Y. Tian, T. Floquet, W. Perruquetti, "Fast state estimation in linear time-variant systems: an algebraic approach", *IEEE Conf. on Decision and Control*, Cancun, Mexico, 2008.
13. R. Vidal, A. Chiuso, S. Soatto and S. S. Sastry, "Observability of linear hybrid systems", in   *Hybrid Systems: Computation and Control, Lecture Notes in Computer Science*, Springer Verlag, Vol. 2623, pp. 526–539, 2003.
14. T.G. Wright, L.N. Trefethen, "Eigenvalues and pseudospectra of rectangular matrices" ,  *Tech. report. Oxford University, Numerical analysis group*, 2001.
15. L. Schwartz, "Theorie des distributions", *Paris, Hermann*, 1966.
16. J. Lygeros, K.H. Johansson, S.N. Simic, J. Zhang, and S. Sastry, "Dynamical Properties of Hybrid Automata" *IEEE Transactions on Automatic Control*, Vol. 48, No 1, 2003.
17. A.H. Zemanian, "Distribution theory and transform analysis: an introduction to generalized functions, with applications ", *New-York: McGraw-Hill, (International series in pure and applied mathematics)*, 1965.
18. J. Lygeros, "Lecture Notes on Hybrid Systems", Department of Electrical and Computer Engineering University of Patras, Rio GR-26500, Greece 2-6/2/2004.