Delay identification in time-delay systems using variable structure observers

Serguey Drakunov, Wilfrid Perruquetti, Jean-Pierre Richard, Lotfi Belkoura

Delay Identification in Time-Delay Systems using Variable Structure Observers

S. V. Drakunov[*], W. Perruquetti, J.-P. Richard [†], L. Belkoura[‡]

Abstract

In this paper we discuss delay estimation in time-delay systems. In the Introduction a short overview is given of some existing estimation techniques as well as identifiability studies. In the following sections we propose several algorithms for the delay identification based on variable structure observers.

1 Introduction

Numerous researches involve time-delay systems and their applications to modelling and control of concrete systems. To name a few, the two monographs [16, 23] give examples in biology, chemistry, economics, mechanics, viscoelasticity, physics, physiology, population dynamics, as well as in engineering sciences. In addition, actuators, sensors, field networks and wireless communications that are involved in feedback loops usually introduce such delays. As it was noted in the survey [27], delays are strongly involved in challenging areas of communication and information technologies: stability of networked controlled systems, quality of service in MPEG video transmission or high-speed communication networks, teleoperated systems, parallel computation, computing times in robotics... Finally, besides actual delays, time lags are frequently used to simplify very high order models.

For the purpose of stability analysis, it is known that necessary and sufficient conditions can be derived in the case of a known, constant delay h [14, 17]. If the value h is not available, then guaranteeing the robust stability for $h \in [h_m, h_M]$ is convenient but needs more constraining conditions, that can turn out to be only sufficient. From this point of view, the poorest information correspond to the most robust case: $h \geq 0$. Numerous authors (see references in [23], after

[*]Department of Electrical Engineering and Computer Science, Tulane University, New Orleans, 70118 LA, USA, e-mail: drakunov@tulane.edu

[†]INRIA ALIEN and LAGIS CNRS UMR8146, Ecole Centrale de Lille, BP 48, 59651 Villeneuve d'Ascq Cedex - FRANCE, email: wilfrid.perruquetti@ec-lille.fr, jean-pierre.richard@ec-lille.fr

[‡]INRIA ALIEN and LAGIS CNRS UMR8146, University of Lille 1, P2, 59655 Villeneuve dAscq Cedex, France, email: lotfi.belkoura@univ-lille1.fr

proposing "independent on delay" stabilization results (assumption $h \geq 0$), concentrated on "delay-dependent" ones: Mainly, with $h_m = 0$, but also with $h_m > 0$ [15].

Regarding to the delay knowledge, observers as well as predictors probably constitute the most demanding case of applications. Several authors proposed observers (or predictors) for linear systems with delays (for the state space approach, see references in [5] and [3], for the coprime factorization technique see [40]; for the discrete time, see [37]), as well as the linear stochastic [16] and nonlinear [13] cases. An overview is given in [30]. In most cases, the value of the delay (mainly constant) was involved in the realizations, which means that its measurement was assumed. Note that what is defined as "observers without internal delay" (see [5, 6, 10]) involves the output knowledge at the present and delayed instants. This means that the delay is known or, at least, is calculable. Similarly, [19] designed a finite-dimensional observer (then, without delay) since it was constructed just for the finite set of unstable or poorly damped modes of the delay system: however, the determination of these modes, here again, requires the delay knowledge.

In concrete applications, the delay invariance and delay knowledge remain assumptions coming more from the identification and analysis limits than from technical facts. So, the robustness with regard to the delay estimation (and variation) should receive additional interest.

Works on identification of time delay systems have shown the complexity of the question [34]. Identifying the delay is not an easy task for systems with both input and state delays, or when the delay is varying enough to require an adaptive identifier. Several authors use the relay-based approach initiated by Astrom and Hagglund [21, 31], which, however, is not a real-time procedure since it needs to close some switching feedback loop during a preliminary identification phase. The adaptive control of delay systems is not so much developed either [2, 12, 35] and the delay is generally assumed to be known. As noted in [7], the on-line delay estimation has a longstanding issue in signal processing: these applications [7, 32] however assume that both the present signal $u = u(t)$ and its delayed value $u(t-h)$ are known and their derivative to be bounded as follows: $0 < \alpha \leq \left|\dot{u}(t)\right| \leq \beta$.

Another approach to the delay identification [24] is based on a "multi-delay approach", which algorithm involves $N+1$ delays $h_i = \left[h_0 + \frac{i\Delta}{N},\right]$ for the identification of a unique delay $h \in [h_0, h_0 + \Delta]$.

Another adaptive delay identification scheme can be found in the book [16]: if h_0 is some approximation of the actual delay value $h(t) = h_0 + \Delta h(t)$ ($|\Delta h| < h_0$), then the algorithm requires the measurement of delayed variables $x(t-h_0)$ and $\dot{x}(t-h_0)$ ($x(t)$ denotes the instantaneous state). The results are local (*i.e.*, valid for $|\Delta h|$ small enough) and given without complete proof.

Lastly, let us mention that all these approaches suffer from a long computing time. The conclusion is that, despite the advantages one can expect from the delay knowledge, advanced *on-line* identification methods for delay estimation are still expected.

The sliding mode control methodology developed for years (see, for example,

[33], [26], [28] and references therein) allows to design sliding mode observers ([33], [8], [9]), and parameter identifiers including the once for distributed parameter systems. The developed control algorithms based on sliding modes have extremely robust behavior with respect to the heated material parameter variations and external disturbances.

1.1 Identifiability analysis

A standard approach for identification of systems implies that the structure of the system is known and the problem is in finding the values of parameters (including the delays) involved in the set of equations describing the process. The ability to ensure this objective is typically referred to as *parameter* identifiability. First results on identifiability for delay differential equations can be found in [29, 34, 20] and for recent results see also [**?**] and references therein. However, many of these results are limited to the homogeneous case (no forcing term) and use a spectral approach involving infinite dimensional spectrum. The approach used in [18, 1] extends the identifiability analysis to more general systems described by convolution equations of the form:

$$R * w = 0, \qquad R = [P, -Q]\,, \qquad w = \begin{bmatrix} y \\ u \end{bmatrix}, \tag{1}$$

where $P\,(n \times n)$ and $Q\,(n \times q))$ are matrices with entries in the space $\mathcal{E}'$ of distributions with compact support. Equation (1) correspond to a behavioral approach of systems described by convolutional equations (see [36]). Here, $R(s)$, the Laplace transform of R, provides a kernel representation of the behavior $\mathcal{B}$ which consists in the set $\tilde{w}$ of all admissible trajectories in the space of $C^\infty(\mathbb{R}, \mathbb{R})$ functions, and $\tilde{w} \in \mathcal{B} = \ker_{\mathcal{E}}\, R(s)$.

The concept of identifiability is based on the comparison of the original system and its associated reference model governed by (1) and in which R, P,Q and y are replaced by $\hat{R}$, $\hat{P}$, $\hat{Q}$ and $\hat{y}$ respectively. System (1) is therefore said to be identifiable if there exists a control u such that the output identity $\hat{y} = y$ results in $\hat{R} = R$, which means uniqueness of the matrix coefficients as well as that of the delays.

For most practical cases, and provided a sufficiently rich input signal, identifiability of (1) reduces to

1. rank $R(s) = n$, $s \in \mathbb{C}$,

2. $\operatorname{conv} \det P = n \operatorname{conv} R$.

where $\operatorname{conv} R$ denote the smallest closed interval that contains the support of R (i.e. the convex hull of $\operatorname{supp} R$), and $\det P$ is the determinant with respect to the convolution product. These conditions are closely linked to the property of approximate controllability in the sense that the reachable space is dense in the state space [39].

The following example [1] shows the applicability of the previous result to systems with distributed delays. Consider the multivariable delay system

$$\begin{array}{l} \dot{x}_1(t) = x_1(t) + \int_{-1}^{0} x_2(t+\tau)d\tau \\ \dot{x}_2(t) = x_1(t-1) + x_2(t) + \int_{-1}^{0} u(t+\tau) \end{array} \tag{2}$$

and denote $\pi(t) = H(t) - H(t-1)$, with H the Heaviside function. Here, $\operatorname{supp}\pi = [0,1]$ and some simple manipulations show that system (2) admits a kernel representation $R * \omega = 0$ with $\omega = (x_1, x_2, u)^T$ and

$$R = \left[\begin{array}{cc} P, & -Q \end{array} \right] = \left[\begin{array}{cc|c} \delta' - \delta & -\pi & 0 \\ -\delta_1 & \delta' - \delta & -\pi \end{array} \right] \tag{3}$$

Clearly, $\operatorname{conv} R = [0,1]$ while $\det P = \delta" - 2\delta' + \delta - \delta_1 * \pi$, from which one easily gets $\operatorname{conv}\det P = [0,2]$, so condition (2) of is satisfied. On the other hand,

$$R(s) = \left[\begin{array}{ccc} s-1 & -(e^{-s}-1)/s & 0 \\ -e^{-s} & s-1 & -(e^{-s}-1)/s \end{array} \right], \tag{4}$$

and the determinant formed by the second and third column of $R(s)$ is nonzero for $s \neq 0$, and for $s = 0$, the first and second column of $R(0)$ form a non singular matrix. Hence condition (1) is also satisfied and system (2) is identifiable.

In the case of distributed delays, the major limitation of the previous approach is the need of the largest delay involved in (1). In return for more restrictive models with lumped and commensurate delays of the form

$$\dot{x}(t) = \sum_{i=0}^{r} A_i \, x(t - i.h) + B_i \, u(t - i.h), \tag{5}$$

a simpler identifiability result which no longer requires the assumption of an a priori known memory length is obtained in [18]. It can be expressed in terms of weak controllability, concept introduced in [22] for systems over rings, through the rank condition (over the ring $R[\nabla]$):

$$\operatorname{rank}\left[B(\nabla), \ldots, A^{n-1}(\nabla)B(\nabla)\right] = n, \tag{6}$$

where

$$A(\nabla) = \sum_{i=0}^{r} A_i \, \nabla^i, \quad B(\nabla) = \sum_{i=0}^{r} B_i \, \nabla^i. \tag{7}$$

Note however that all the previous results are limited to linear and time invariant models. In case of nonlinear delay systems or time dependant delays, general identifiability results are still expected.

1.2 Sufficiently rich input

In identification procedures the design of a sufficiently rich input which enforces identifiability is also an important issue. Given a reference model associated to the process under study, one has to know whether equality of the outputs results in that of the transfer functions. Few results are dealing with such issue for time delay systems. In [24, 1] the input design is considered in the time domain rather than the frequency domain and the approaches are mainly based on the non smoothness of the input. More precisely, if

$$\Lambda_u = \{s_0, s_1, ..., s_L, ...\} \tag{8}$$

denote the singular support of u (i.e. the set of points in $\mathbb{R}$ having no open neighborhood to which the restriction of u is a C^∞ function), the input is required to be sufficiently "discontinuous" in the sense that

$$\text{rank}\,[U_0(D), ..., U_L(D)] = q \tag{9}$$

where the polynomial matrices $U_l(D)$ are formed with the (possible) jump of $u^{(k)}(t)$ for some $k \geq 0$ at $t = s_l$ by

$$U_l(D) = \sum_{i=0}^{k} [u^{(k-i)}(s_l + 0) - u^{(k-i)}(s_l - 0)]\, D^i \tag{10}$$

On the other hand, "the discontinuity points" $s_0, s_1, ..$ should be sufficiently spaced in the general case of distributed delays, although for lumped delays, this constraint (which may constitute a serious drawback in situations where on line procedures are used) can be relaxed using commensurability considerations.

The simplest example consists of a piece-wise constant $\mathbb{R}^q$-valued function with appropriate discontinuities, although inputs of class $\mathcal{C}^r$ for an arbitrary finite integer r can be formed.

1.3 Parameter identification

Papers dealing with parameter identification of systems with time-delay in the state variables are not so numerous. Let us mention [25] which concerns parameter identification of time-delay systems with *commensurate delays*, and [38, 24]. The following algorithm, first presented in [38], allows on-line identification of linear dynamic systems with finitely many lumped delays in the state vector and control input. Associated with the model

$$\dot{x}(t) = \sum_{i=0}^{r} A_i\, x(t - h_i) + B_i\, u(t - h_i) \tag{11}$$

with unknown matrices A_i and B_i, he considered the identifier system:

$$\dot{\hat{x}}(t) = \sum_{i=0}^{r} \hat{A}_i(t)\hat{x}(t - h_i) + \hat{B}_i(t)u(t - h_i) - G\Delta x(t), \tag{12}$$

where $\Delta x(t) = x(t) - \hat{x}(t)$ is the "state" error, $G \in \mathbb{R}^{n\times n}$ is a Hurwitz matrix, and time-varying matrices $\hat{A}_i(t)$, $\hat{B}_i(t)$ satisfying

$$\begin{aligned} \dot{\hat{A}}_i(t) &= F_i P \Delta x(t) x^T(t - h_i), \quad \hat{A}_i(0) = \hat{A}_i^0, \\ \dot{\hat{B}}_i(t) &= \Phi_i P \Delta x(t) u^T(t - h_i), \quad \hat{B}_i(0) = \hat{B}_i^0, \end{aligned} \tag{13}$$

with adaptation gain matrices F_i, Φ_i being positive definite and P the (positive definite) solution of the Lyapunov equation $G^T P + PG = -Q$ for a given positive definite matrix Q. Under stability and identifiability conditions, and using a sufficiently rich input signal, it is shown that the state error $\Delta x(t)$ converges asymptotically to 0, and the time-varying matrices $\hat{A}_i(t)$, $\hat{B}_i(t)$ converge towards the plant parameter matrices A_i, B_i.

1.4 Delay identification

When facing unknown delays, the previous approach [24] may also provide an *estimation* of them by considering an identifier with a large number m of fictitious delays

$$\dot{\hat{x}}(t) = \sum_{j=0}^{m} \hat{A}_j(t)\hat{x}(t-\tau_j) + \hat{B}_j(t)u(t-\tau_j) - G\Delta x(t), \tag{14}$$

and in which, by virtue of the identifiability property, the $\hat{A}_j(t)$ and $\hat{B}_j(t)$ coefficients tend to zero except for $h_i \simeq \tau_j$. However, the accuracy of this identification depends on the number m of implemented delays and the computational effort strongly increases with m, which might restrain the real-time identification possibilities.

Another (single) delay estimation technique can be found in [7] where the present signal $u = u(t)$ and its delayed value, denoted by $v(t) = u(t-h)$ are supposed to be known and their derivative to be bounded as follows: $0 < \alpha \leq \left|\dot{u}(t)\right| \leq \beta$. The following scheme is used:

$$\dot{\hat{h}} = -\frac{p(t)\dot{u}(t-\hat{h})}{1 + p(t)\dot{u}^2(t-\hat{h})}\left[u(t-h) - u(t-\hat{h})\right], \tag{15}$$

$$\dot{p}(t) = \frac{p^2(t)\dot{u}^2(t-\hat{h})}{1 + p(t)\dot{u}^2(t-\hat{h})},\ p(0) = p_0 > 0. \tag{16}$$

It was involved, with simulation, in the speed control system of a direct injected diesel engine, thus adjusting the gains of a simple PI controller. Note however that all parameters (but the delay) are supposed to be known.

Let us also mention the adaptive (single) delay identification scheme in the book [16]: if h_0 is some approximation of the actual (time-varying) delay value $h(t) = h_0 + \Delta h(t)$, then the algorithm requires the measurement of delayed variables: $x(t-h_0)$ and $\dot{x}(t-h_0)$, together with the assumption $|\Delta h| < h_0$. The results are local (i.e., valid for $|\Delta h|$ small enough) and, here also, the real-time possibilities are to be checked.

2 Class of systems under consideration

The present paper contributes to this problem *via* a variable structure identification algorithm. This means it combines differential equations with unknown delays together with a dynamic system with discontinuous functions in such a way that some variables of the combined system converge to the delays.

A particular attention is paid to the case of a discrete delay model, which appears in many applications:

Discrete Delay Linear Model

$$\dot{x}(t) = \sum_{i=0}^{r} \left[A_i x(t-h_i) + B_i u(t-h_i) \right], \tag{17}$$

where $0 = h_0 < h_1 < \ldots < h_r \le h$ are time delays, A_i and B_i are matrices of appropriate dimensions. As in the general case, we assume that there are two initial continuous functions $x_o(t)$, $u_o(t)$ and $x(t) \equiv x_o(t)$, $u(t) \equiv u_o(t)$ for $t \in [-h, 0]$.

The method, which we develop, can be also applied to systems described by very general equations with distributed delay:

Distributed Delay Linear Model

$$\dot{x}(t) = \int_0^h \left[d\mu(\xi) x(t-\xi) + d\nu(\xi) u(t-\xi) \right]. \tag{18}$$

Here $x \in \mathbb{R}^n$ is a system vector (this is not a state vector), and $u \in \mathbb{R}^p$ is a system input, $\mu(\xi)$ and $\nu(\xi)$ are matrix valued functions of bounded variation[1] ($Var(\mu) < C$, $Var(\nu) < C$) with corresponding dimensions, or, equivalently, matrix valued measures (not necessarily positive).

The following assumptions are made regarding $\mu(\xi)$ and $\nu(\xi)$: for any continuously differentiable functions with compact support $\phi(\xi)$, and $\psi(\xi)$ we assume that

$$\| \int_{h-s}^{h} d\mu(\xi)\phi(\xi) - A\phi(\xi) \| \le \gamma_a(s) sup_{h-s \le \xi \le h} \|\phi(\xi)\|, \tag{19}$$

and

$$\| \int_{h-s}^{h} d\nu(\xi)\psi(\xi) - B\psi(\xi) \| \le \gamma_b(s) sup_{h-s \le \xi \le h} \|\psi(\xi)\|, \tag{20}$$

where

$$A = \mu(h) - \mu(h^-),$$
$$B = \nu(h) - \nu(h^-),$$

and $\gamma_a(s)$,$\gamma_b(s)$, $(s \ge 0)$ are continuous functions such that $\gamma_a(0) = 0$, $\gamma_b(0) = 0$.

From (19), (20) it follows (see, for example, [14]) that a solution of (18) exists and is unique for $u(\cdot) \in C_{[0,\infty]}(\mathbb{R}^p)$ and any continuous initial condition $x_o(s)$ $(-h \le s \le 0)$, $x_o(\cdot) \in C_{[-h,0]}(\mathbb{R}^n)$.

---

[1] $Var(\mu) \triangleq \sup_{0 \le t_1 < \ldots < t_i < \ldots \le h} \sum_i \| \int_{t_i}^{t_{i+1}} d\mu(\xi) \|$, where sup is taken over all finite partitions $0 \le t_1 < \ldots < t_i < \ldots \le h$ of the interval $[0, h]$.

We develop a class of identification algorithms to estimate $\mu(\xi)$, and $\nu(\xi)$, which are assumed to be unknown[2], while $x(t)$, and $u(t)$ are known.

The model (17) is a particular case of 18 for δ-measures[3] :

$$d\mu(\xi) = \sum_{i=0}^{r} A_i \delta(\xi - h_i) d\xi,$$

$$d\nu(\xi) = \sum_{i=0}^{r} B_i \delta(\xi - h_i) d\xi.$$

Considering this model, we concentrate on the case when matrices A_i and B_i are known, and the main problem is to estimate the delays. Although, a general algorithm for estimation of the measures $\mu(\xi)$, and $\nu(\xi)$, allows simultaneous identification of the delays, and parameters.

Other important cases also include distributed delay model:

$$\dot{x}(t) = \int_0^h \left[A(\xi)x(t-\xi) + B(\xi)u(t-\xi)\right] d\xi, \tag{21}$$

here $A(\xi)$ and $B(\xi)$ are variable matrices.
This equation is a particular case of (18) for absolutely continuous measures with densities $A(\xi)$, and $B(\xi)$:

$$d\mu(\xi) = A(\xi)d\xi$$
$$d\nu(\xi) = B(\xi)d\xi.$$

Another possible generalization which will be considered is a model with discrete variable delays:

Discrete Variable Delay Linear Model

$$\dot{x}(t) = \sum_{i=0}^{r} \left[A_i x(t - h_i(t)) + B_i u(t - h_i(t))\right]. \tag{22}$$

In some cases we can even consider a very general nonlinear time-varying model.

---

[2]Since $\int_0^h d\mu(\xi)\phi(t-\xi) = \int_0^h d\widetilde{\mu}(\xi)\phi(t-\xi)$ for $\widetilde{\mu} - \mu = const$, in fact, we just find a representative in an equivalence class.

[3]We use a standard notation $\delta(\xi - h)d\xi$ for a discrete measure $d\mu(\xi)$ formally defined through the equality

$$\int \phi(\xi) d\mu(\xi) = \phi(h),$$

for any function $\phi(\xi)$ continuous at h.

Distributed Variable Delay Nonlinear Model

$$\dot{x}(t) = \int_0^h d\mu(\xi, t) F(x(t-\xi), u(t-\xi)). \tag{23}$$

The paper is organized as follows. After introducing some notations in section 3, section 4 recalls the main results concerning the general problem of identifiability. In section 5 to demonstrate our technique we provide three delay estimation algorithms for the model (17), and prove that they converge as soon as the initial values of the delay estimates are close enough to the real delays. In this section we also consider variable delay model (22). In the section (7) we consider the algorithms for the general distributed delay model (18). Finally, in section 8 we consider the examples.

3 Notations

The following notations are used:

1. $\mathrm{cl}(\mathcal{S}), \overline{\mathrm{conv}}(\mathcal{S})$ denotes respectively the closure, and the closure of the convex hull of the set $\mathcal{S}$,

2. for x in a normed vector space $\mathcal{X}$: $\mathcal{B}_\varepsilon(x)$ is the open ball centered at x of radius ε, this is $\mathcal{B}_\varepsilon(x) = \{x \in \mathcal{X} : \|x\| < \varepsilon\}$, for a set $\mathcal{A} \subset \mathcal{X} : \mathcal{B}_\varepsilon(\mathcal{A}) = \{y \in \mathcal{X} : \exists x \in \mathcal{A} \wedge y \in \mathcal{B}_\varepsilon(x)\} = \bigcup_{x \in \mathcal{A}} \mathcal{B}_\varepsilon(x)$,

3. Let $V : \mathbb{R}^{n+1} \to \mathbb{R}_+, (t, x) \mapsto V(t, x)$. When the classical gradient exists, it is denoted by $\nabla V(t, x)$ (gradient of V evaluated at (t, x)). Let Ω be the union of any set of zero measure with the set where V fails to be differentiable, then, if V is locally Lipschitz in (t, x), one can defined the generalized gradient (see [4]) by

$$\partial_C V(t, x) = \overline{\mathrm{conv}} \left\{ \lim_{(t_i, x_i) \notin \Omega \to (t, x)} \nabla V(t_i, x_i) \right\}, \tag{24}$$

also called Clarke's gradient for finite dimensional Banach space in short *generalized gradient*,

4. For $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, sign$(x)$ is a vector with components

$$\mathrm{sign}(x_i) = \begin{cases} 1, \text{ if } x_i > 0 \\ 0, \text{ if } x_i = 0 \\ -1, \text{ if } x_i < 0 \end{cases}. \tag{25}$$

5. A function st is such that

$$\mathrm{st}(h) = \begin{cases} 0, \text{ if } h \leq 0 \\ 1, \text{ if } h > 0 \end{cases}. \tag{26}$$

6. For $h \in \mathbb{R}$, a function $sp_{\varepsilon,h}(s)$ is such that

$$sp_{\varepsilon,h}(s) = \begin{cases} 0, \text{ if } s \le 0, \text{ or } s \ge h \\ 1, \text{ if } \varepsilon \le s \le h - \varepsilon \\ \text{arbitrary, if } 0 < s < \varepsilon, \text{ or } h - \varepsilon < s < h \end{cases}, \tag{27}$$

   where "arbitrary" means any function such that overall the function $sp_{\varepsilon,h}(s)$ is continuously differentiable everywhere, and bounded together with its derivative.

7. As $dsp_{\varepsilon,h}(s)$ we denote the derivative of $sp_{\varepsilon}(s)$.

8. For $v, w \in \mathbb{R}^n$, $\langle v, w \rangle$ will denote the scalar product of v and w.

4 Identifiability conditions

As was mentioned earlier, we develop algorithms of estimating μ, and ν by observation of $x(t)$, and $u(t)$.

Let us assume that μ_1, ν_1, and μ_2, ν_2 satisfy the equations:

$$\dot{x}(t) = \int_0^h \left[d\mu_1(\xi)x(t-\xi) + d\nu_1(\xi)u(t-\xi) \right],$$

$$\dot{x}(t) = \int_0^h \left[d\mu_2(\xi)x(t-\xi) + d\nu_2(\xi)u(t-\xi) \right],$$

with the same $x(t)$, and $u(t)$. Subtracting these equations we obtain

$$\int_0^h \left[d\bar{\mu}(\xi)x(t-\xi) + d\bar{\nu}(\xi)u(t-\xi) \right] = 0, \tag{28}$$

where $\bar{\mu} = \mu_1 - \mu_2$, $\bar{\nu} = \nu_1 - \nu_2$.

The question arises: *When the equation (28) implies that $d\bar{\mu} = 0$, $d\bar{\nu} = 0$ a.e. (or $\bar{\mu} \equiv const$, $\bar{\nu} \equiv const$ a.e.) ?* Note that, $x(t)$, and $u(t)$ in (28) are not arbitrary but they satisfy the equation (18). The general answer to this question is that the initial function $x_o(t)$, and/or $u(t)$ should be rich enough. Different definition of richness result in different conditions.

For a particular case of discrete delays (17) the following definitions and identifiability from [18] can be used.

We consider two different delay systems of the type (17) with respective states x and $\widehat{x}$. Does the equality $x \equiv \widehat{x}$ implies that the parameters, including the delays, are equal? More explicitly, we consider the system:

$$\dot{\widehat{x}}(t) = \sum_{i=0}^{m} \left[\widehat{A}_i \widehat{x}(t - \widehat{h}_i) + \widehat{B}_i u(t - \widehat{h}_i) \right], \tag{29}$$

$$\widehat{x} \in \mathbb{R}^n, \; 0 \le \widehat{h}_0 < \widehat{h}_1 < \ldots < \widehat{h}_m. \tag{30}$$

Definition 1 *System (17) is said to be identifiable under arbitrary initial conditions if there exists an input signal* $u(t)$ *such that the equality* $x(t) \equiv \widehat{x}(t)$, $t \geq 0$ *implies* $r = m$, *for all* $i = 0, ..., m$, $A_i = \hat{A}_i$, $B_i = \hat{B}_i$, $h_i = \widehat{h}_i$.

If we assume that A_i, B_i are known, a number of sufficient conditions can be obtained for identifiability with respect to the delay vector $\mathcal{H} = [h_1, ..., h_r]^T$ for the case of discrete-delay (17), or $\mathcal{H} = [\mu, \nu]$ for the case of distributed delay (18). Let us introduce the following function

$$\Delta(t, \mathcal{H}) = \dot{x}(t) - \sum_{i=0}^{r} \left[A_i x(t - h_i) + B_i u(t - h_i)\right],$$

for the model (17), or

$$\Delta(t, \mathcal{H}) = \dot{x}(t) - \int_0^h \left[d\mu(\xi) x(t - \xi) + d\nu(\xi) u(t - \xi)\right].$$

for the model (18).

If the delays $\mathcal{H}$ are are not time varying the local identifiability may be expressed as uniqueness of solution of the equation $\Delta(t, \mathcal{H}) = 0$ with respect to $\mathcal{H}$, for all $t \geq 0$.

One example of such sufficient conditions can be developed as follows: let us consider r moments of time $0 \leq t_1 < ... < t_r$. To study the identifiability we introduce the r-dimensional function $W(t_1, ..., t_r, \mathcal{H}) = [w(t_1, \mathcal{H}), ..., w(t_r, \mathcal{H})]^T$, whose components are $w(t, \mathcal{H}) = \|\Delta(t, \mathcal{H})\|^2$. As a norm we can consider the Euclidian norm: $\|\Delta(t, \mathcal{H})\| = \langle \Delta(t, \mathcal{H}), \Delta(t, \mathcal{H}) \rangle^{1/2}$. If a Jacobian $\frac{\partial}{\partial \mathcal{H}} W(t_1, ..., t_r, \mathcal{H}) \neq 0$ is nonsingular for all $\mathcal{H}$, the map $W(t_1, ..., t_r; \cdot) : \mathbb{R}^r \to \mathbb{R}^r$ is an injection. Therefore, there is only one $\mathcal{H}$ satisfying $W(t_1, ..., t_r, \mathcal{H}) = 0$.

Lemma 1 *If* $x(t)$, *and* $u(t)$ *are continuously differentiable, the system (17) is locally identifiable with respect to* $\mathcal{H}$ *if for any* $0 \leq t_1 < ... < t_r$ *the following matrix* $H^0 = H^0(t_1, ..., t_r)$ *is nonsingular* ($\det(H^0(t_1, ..., t_r)) \neq 0$)

$$\begin{aligned} H^0 &= \begin{pmatrix} H^0_{11}(t_1) & ... & H^0_{1r}(t_1) \\ \vdots & \ddots & \vdots \\ H^0_{r1}(t_r) & ... & H^0_{rr}(t_r) \end{pmatrix}, \\ H^0_{ij}(t_i) &= \langle \left[A_j x(t_i - h_j) + B_j u(t_i - h_j)\right], \left[A_j \dot{x}(t_i - h_j) + B_j \dot{u}(t_i - h_j)\right] \rangle \end{aligned}$$

Obviously, this condition depends on the state x and the input u, as well as their derivatives. Since the derivative of the state can be expressed using system equations we can obtain a form of the above condition which is a function only of the system vector $x(t)$, the control $u(t)$, and its derivative $\dot{u}(t)$. However, this condition is more difficult to verify.

We will be considering our systems on a finite time interval: $0 \leq t \leq t_1$. If $x(t)$ is twice continuously differentiable, and $u(t)$ is continuously differentiable, then for some $C > 0$:

$$\|\Delta(t, \mathcal{H}) - \Delta(t, \mathcal{H}')\| \leq C \|\mathcal{H} - \mathcal{H}'\|,$$

or if $\Delta(t, \mathcal{H}) = 0$

$$\|\Delta(t, \mathcal{H}')\| \leq C\|\mathcal{H} - \mathcal{H}'\|.$$

The inverse of this inequality represents strong identifiability condition which means not just uniqueness of a solution $\Delta(t, \mathcal{H}) = 0$, but also a continuity of a map inverse to $\Delta(t; \cdot) : \mathbb{R}^r \to \Delta(t, \mathbb{R}^r)$, (here $\Delta(t, \mathbb{R}^r)$ is a set of all functions $\{d(t) : \exists \mathcal{H} \in \mathbb{R}^r, d(t) = \Delta(t, \mathcal{H})\}$).

Every time we say that the corresponding system is "identifiable", the following local **Identifiability Assumptions** will be made in all further results:

- For the system (17): $\forall t : 0 \leq t \leq t_1\ \exists c_1 : c_1 > 0$ such that $\|\mathcal{H}' - \mathcal{H}\|_1 \leq c_1\|\Delta(t, \mathcal{H}')\|_1$ for all $\mathcal{H}' \in \mathcal{B}_\varepsilon(\mathcal{H})$ where $\mathcal{B}_\varepsilon(\mathcal{H})$ is some ball of the radius $\varepsilon > 0(\exists \varepsilon > 0)$, with the center in $\mathcal{H}$.

- For the system with variable delays (22): $\forall t : 0 \leq t \leq t_1\ \exists c_1 : c_1 > 0$ such that $\|\mathcal{H}'(t) - \mathcal{H}(t)\|_1 \leq c_1\|\Delta(t, \mathcal{H}'(t))\|_1$ for all $\mathcal{H}'(\cdot) \in \widetilde{\mathcal{B}}_\varepsilon(\mathcal{H}(\cdot))$ where $\widetilde{\mathcal{B}}_\varepsilon(\mathcal{H}(\cdot))$ is some ball in $C^1_{[-h,+\infty]}$ of the radius $\varepsilon > 0(\exists \varepsilon > 0)$, with the center in $\mathcal{H}(\cdot)$.

- For the system (17). If we denote as $\widetilde{\mu} = [\mu(\xi), \nu(\xi)]$ a $(n \times (n+m))$-matrix[4]. Then the identifiability condition is $\forall t : 0 \leq t \leq t_1\ \exists c_1 : c_1 > 0$ such that $\|\widehat{\widetilde{\mu}} - \widetilde{\mu}\|_1 \leq c_1\|\Delta(t, \widehat{\widetilde{\mu}})\|_1$ for all $\widehat{\widetilde{\mu}} \in \mathcal{B}_\varepsilon(\widetilde{\mu})$ where $\mathcal{B}_\varepsilon(\widetilde{\mu})$ is some ball of the radius $\varepsilon > 0(\exists \varepsilon > 0)$, with the center in $\widetilde{\mu}$.

5 Sliding-mode observers for delay estimation

As was mentioned in the Introduction, in order to demonstrate our technique, we, first, we consider here the algorithm[5] for delay estimation in system (17).

Let us assume at first that the parameters A_i, B_i are known and only the delays h_i need to be estimated. We also assume to know the upper bound h of the unknown delays

$$\max[h_1, ..., h_r] \leq h.$$

Introducing the delay vector $\mathcal{H} = [h_1, \ldots, h_r]^T$ we are going to design an algorithm for the estimate $\widehat{\mathcal{H}}(t) = \left[\widehat{h}_1(t), \ldots, \widehat{h}_r(t)\right]^T$. For this, let us introduce the

[4]We consider the set of all such $\widetilde{\mu}$ as a linear normed space with the norm defined as $|\ \widetilde{\mu}|\ = \sum_{i=1}^n \sum_{j=1}^n Var(\mu_{ij}(\xi)) + \sum_{i=1}^n \sum_{j=1}^m Var(\nu_{ij}(\xi))$. In fact, elements of this space are not individual $\widetilde{\mu}$, but equivalent classes of the matrix functions whose difference is a constant matrix, (since $Var(const) = 0$) (see the footnote on the previous pages).

[5]In our first algorithm, we assume that not only $x(t)$, but also $\dot{x}(t)$ is measured. To make the algorithm more robust in practice one can use a shifted value $\dot{x}(t-\tau)$ (with a known fixed value τ), and consider a system

$$\dot{x}(t-\tau) = \sum_{i=0}^{r} \left[A_i x(t - h_i - \tau) + B_i u(t - h_i - \tau)\right], \tag{31}$$

then the same approach can be worked out.

Figure 1: System structure.

5.1 Delay Estimation Algorithm A

We propose the following delay estimator for $i = 1, \ldots, r$:

[6] Here and further we denote as $\dot{\Delta} = \frac{d\Delta(t,\hat{\mathcal{H}}(t))}{dt}$.

$$\frac{d\widehat{h}_i(t)}{dt} = -L_i \left\langle g_i(t, \widehat{h}_i(t)), \mathrm{sign}(\Delta) \right\rangle \mathrm{st}(\widehat{h}_i(t)), \tag{36}$$

where the vector functions g_i are defined by (35), and initial condition $\widehat{h}_i(0) \geq 0$.

The following theorem states local convergence of the algorithm (36).

Theorem 2 *Let us assume that there exist a finite time $t_1 > 0$ such that the following holds for $0 \leq t \leq t_1$:*

H1: The input $u(t)$ is, at least once, continuously differentiable and have bounded derivative. The system vector $x(t)$ is, at least, twice continuously differentiable and have bounded derivatives.

H2: The system (17) is identifiable with respect to delay vector for the input $u(t), t \geq -h$, and the initial condition $x_o(t), -h \leq t \leq 0$.

H3: There exists $\delta > 0$, that $|\langle g_i(t, h_i), \beta \rangle| > \delta$ for any vector $\beta = [\beta_1, ..., \beta_n]^T \in \mathbb{R}^n$, such that $|\beta_j| = 1$, for all $j = 1, ..., n$.

Then, using the delay estimator (36) with initial conditions $\widehat{\mathcal{H}}(0) = \left[\widehat{h}_1(0), \ldots, \widehat{h}_r(0)\right]^T$ ($\widehat{h}_j(0) \geq 0, \forall j = 1, ..., r$), close enough to $\mathcal{H} = [h_1, \ldots, h_r]^T$ ($\exists \epsilon > 0, \widehat{\mathcal{H}}(0) \in \mathcal{B}_\epsilon(\mathcal{H})$) there exist finite gains L_i such that after a finite time $t_0, (0 \leqq t_0 \leq t_1)$ we have $\widehat{h}_i(t) \equiv h_i$ for all $i = 0, ..., r$, $t_0 \leq t \leq t_1$.

Proof (Algorithm Convergence): ▶ Using $\dot{\Delta}$ given in (34), (36), and substituting instead of $\ddot{x}$ the derivative of the right hand side of (17) we obtain

$$\begin{aligned} \dot{\Delta} &= \sum_{i=1}^{r} A_i \left(\dot{x}(t - h_i) - \dot{x}(t - \widehat{h}_i(t)) F_i \right) \\ &\quad + \sum_{i=1}^{r} B_i \left(\dot{u}(t - h_i) - \dot{u}(t - \widehat{h}_i(t)) F_i \right), \end{aligned}$$

where

$$F_i = 1 + L_i \left\langle g_i(t, \widehat{h}_i(t)), \mathrm{sign}(\Delta) \right\rangle \mathrm{st}(\widehat{h}_i(t)).$$

Due to presence of the "st" function in the right hand side of (36), and the assumption that $\widehat{h}_i(0) \geq 0$. the estimate $\widehat{h}(t)$ always stays nonnegative, so in the last expression $\mathrm{st}(\widehat{h}_i(t)) \equiv 1$.

Using that, and the definition of g_i, the expression for $\dot{\Delta}$ can be written as

$$\dot{\Delta} = \sum_{i=1}^{r} \left[g_i(t, h_i) - g_i(t, \widehat{h}_i(t)) - L_i \left\langle g_i(t, \widehat{h}_i(t)), \mathrm{sign}(\Delta) \right\rangle g_i(t, \widehat{h}_i(t)) \right] \tag{37}$$

Now, let us consider the following nonsmooth Lyapunov function (see, [11], [33], [**?**])

$$V(\Delta) = \|\Delta\|_1 = \langle \Delta, \mathrm{sign}(\Delta) \rangle \,, \tag{38}$$

Differentiating V in a non smooth context (this is using generalized gradient see [4])

$$\dot{V} \in \cup_{w \in \partial_C V} \left\{ \left\langle w, \dot{\Delta} \right\rangle \right\}, \tag{39}$$

this is

$$\begin{aligned} \dot{V}(\Delta) &= V_{\mathrm{dot}}(\Delta) \text{ if for } i = 1, ..., n \ \Delta_i \neq 0 \\ V_{\mathrm{dot}}(\Delta) &= \sum_{i=0}^{r} \left\langle g_i(t, h_i) - g_i(t, \widehat{h}_i(t)), \mathrm{sign}(\Delta) \right\rangle \\ &\quad - \sum_{i=0}^{r} L_i \left\langle g_i(t, \widehat{h}_i(t)), \mathrm{sign}(\Delta) \right\rangle^2, \end{aligned}$$

or using a notation $\alpha(t, h_i, \widehat{h}_i) = g_i(t, h_i) - g_i(t, \widehat{h}_i(t))$ we have

$$\begin{aligned} V_{\mathrm{dot}}(\Delta) &= \sum_{i=0}^{r} \left\langle \alpha(t, h_i, \widehat{h}_i), \mathrm{sign}(\Delta) \right\rangle && (40) \\ &\quad - \sum_{i=0}^{r} L_i \left\langle g_i(t, h_i(t)) - \alpha(t, h_i, \widehat{h}_i), \mathrm{sign}(\Delta) \right\rangle^2. && (41) \end{aligned}$$

Using the assumptions H1, H2 for some constant C we have the following estimate

$$\|\alpha(t, h_i, \widehat{h}_i)\| \leq C \|\Delta\|_1 = CV$$

which results in

$$V_{\mathrm{dot}}(\Delta) \leq C_1 V - \sum_{i=0}^{r} L_i \left\langle g_i(t, h_i(t)), \mathrm{sign}(\Delta) \right\rangle^2, (C_1 > 0). \tag{42}$$

Finally, using the assumption H3 we obtain that for some $C_2 > 0$

$$V_{\mathrm{dot}}(\Delta) \leq C_1 V - C_2. \tag{43}$$

This inequality guarantees that $V_{\mathrm{dot}}(\Delta)$ is negative for $V < C_2/C_1$. Since $V = \|\Delta\|_1$ due to H1, $V_{\mathrm{dot}}(\Delta)$ is negative for for some ball $\mathcal{B}_{\varepsilon_1}(\mathcal{H})$ with the center in $\mathcal{H}$.

According to the result which can be found in [11], [33] the negativeness of the nonsmooth Lyapunov function derivative on the set $\{\Delta_i \neq 0, i = 1, ..., n\}$ guarantees that the Filippov definition set in the right hand side of (39) is located on the negative part of the real axis. It implies the convergence $V \to 0$ everywhere, including along the sets $\{\Delta_i = 0\}$, when in sliding mode (see, again [11], [33]).

Moreover, from the previous computation if $V(0) \leq C_2/C_1 - \alpha$, where $\alpha > 0$ one gets

$$V_{\rm dot}(\Delta) \quad \leq \quad -\alpha.$$

which ensures that V decreases to zero in finite time. Finally, using the inequality from the assumption H2: $\|\widehat{\mathcal{H}} - \mathcal{H}\|_1 \leq c_1 \|\Delta(t, \widehat{\mathcal{H}})\|_1 = V$ we obtain the desired proof. ◀

5.2 Delay Estimation Algorithm B

Let us consider now a version of the Algorithm which requires more restrictive, but easier to check conditions to converge. In fact, we will consider (17) as a particular case of a more general situation where delays may be different in each equation of the system:

$$\dot{x_j}(t) = \sum_{i=0}^{r} \left[A_i x(t - h_{ij}) + B_i u(t - h_{ij})\right]_j, \tag{44}$$

where $j = 1, ..., n$, and we denote as $[...]_j$ the jth component of a vector.

We design an estimation algorithm for a $r \times n$ matrix of delays $\mathcal{H} = [h_{ij}]$, and prove its convergence in the assumption that the system is identifiable. Obviously, such algorithm should also converge if the delays h_{ij} do not depend on the component number j since this is just a particular case.

For this, let us introduce the following function $\Delta = [\Delta_1, ..., \Delta_n]^T$, where

$$\Delta_j \quad = \quad \dot{x}_j(t) - \sum_{i=0}^{r} f_{ij}(t, \widehat{h}_{ij}(t)), \tag{45}$$

$$f_{ij}(t, \widehat{h}_{ij}(t)) \quad = \quad \left[A_i x\left(t - \widehat{h}_{ij}(t)\right) + B_i u\left(t - \widehat{h}_{ij}(t)\right)\right]_j, \tag{46}$$

The time derivative of each Δ_j along the system (17) is

$$\dot{\Delta}_j \quad = \quad \ddot{x}_j(t) - \sum_{i=0}^{r} g_{ij}(t, \widehat{h}_{ij}(t)) \left(1 - \frac{d\widehat{h}_{ij}(t)}{dt}\right), \tag{47}$$

$$g_{ij}(t, \widehat{h}_{ij}(t)) \quad = \quad \left[A_i \dot{x}\left(t - \widehat{h}_{ij}(t)\right) + B_i \dot{u}\left(t - \widehat{h}_{ij}(t)\right)\right]_j. \tag{48}$$

The following delay estimator can be used for $i = 1, \ldots, r$:; $j = 1, \ldots, n$

$$\frac{d\widehat{h}_{ij}(t)}{dt} = -L_i g_{ij}(t, \widehat{h}_{ij}(t))\, \mathrm{sign}(\Delta_j)\, \mathrm{st}(\widehat{h}_{ij}(t)) \tag{49}$$

.

Theorem 3 *Let us assume that there exist a finite time $t_1 > 0$ such that the following holds for $0 \leq t \leq t_1$:*

H1: The input $u(t)$ is, at least once, continuously differentiable and have bounded derivative.The system vector $x(t)$ is, at least, twice continuously differentiable and have bounded derivatives.

H2: For given $x(t)$, and $u(t)$, the system (44) is identifiable with respect to unknown delay matrix $\mathcal{H} = [h_{ij}]$ for the input $u(t), t \geq -h$, and the initial condition $x_o(t), -h \leq t \leq 0$.

H3: There exists $\delta > 0$, that for any $j = 1, ..., n$, there exists at least one $i = 1, ..., r$ such that $|g_{ij}(t, h_{ij})| > \delta$.

Then, using the delay estimator (49) with initial conditions $\widehat{\mathcal{H}}(0) = \left[\widehat{h}_{ij}(0)\right]$ ($\widehat{h}_{ij}(0) \geq 0, \forall j = 1, ..., r; i = 1, ..., n$), close enough to $\mathcal{H}$ ($\exists \epsilon > 0, \widehat{\mathcal{H}}(0) \in \mathcal{B}_\epsilon(\mathcal{H})$) there exist finite gains L_i such that after a finite time $t_0, (0 \leqq t_0 \leq t_1)$ we have $\widehat{h}_{ij}(t) \equiv h_{ij}$ for all $i = 0, ..., r; i = 1, ..., n,$ $t_0 \leq t \leq t_1$.

Proof (Algorithm Convergence): ▶ We will just outline the proof since it is very similar to the proof of the Theorem 2. In this case we can prove separately sliding mode existence for every scalar sliding surface $\Delta_j = 0$. Using $\dot{\Delta}_j$ given in (47) we obtain

$$\begin{aligned} \dot{\Delta}_j &= \ddot{x}_j(t) - \sum_{i=0}^{r} g_{ij}(t, \widehat{h}_{ij}(t)) \left(1 - \frac{d\widehat{h}_{ij}(t)}{dt}\right), & (50) \\ &= -\sum_{i=0}^{r} g_{ij}(t, \widehat{h}_{ij}(t)) \frac{d\widehat{h}_{ij}(t)}{dt} + \Phi_j(\widehat{h}_{ij}(t)) & (51) \end{aligned}$$

where we denoted as $\Phi_j(\widehat{h}_{ij}(t))$ the sum of all terms which do not depend on $\frac{d\widehat{h}_{ij}(t)}{dt}$:

$$\begin{aligned} \Phi_j(\widehat{h}_{ij}(t)) &= \ddot{x}_j(t) - \sum_{i=0}^{r} g_{ij}(t, \widehat{h}_{ij}(t)), & (52) \\ &= \sum_{i=0}^{r} g_{ij}(t, h_{ij}) - \sum_{i=0}^{r} g_{ij}(t, \widehat{h}_{ij}(t)) & (53) \end{aligned}$$

As in the previous algorithm, the function "st" in the right hand side of (49), and the assumption that $\widehat{h}_{ij}(0) \geq 0$ guarantee that the estimate $\widehat{h}_{ij}(t)$ always stays nonnegative. So using the algorithm equation (49) for $\widehat{h}_{ij}(t) \geq 0$ we have

$$\dot{\Delta}_j = -\sum_{i=0}^{r} L_i \left[g_{ij}(t, \widehat{h}_{ij}(t))\right]^2 \operatorname{sign}(\Delta_j) + \Phi_j(\widehat{h}_{ij}(t)), \qquad (54)$$

or using a notation $\alpha_{ij}(t, h_{ij}, \widehat{h}_{ij}(t)) = g_{ij}(t, h_{ij}) - g_{ij}(t, \widehat{h}_{ij}(t))$

$$\dot{\Delta}_j = -\sum_{i=0}^{r} L_i \left[g_{ij}(t, h_{ij}) - \alpha_{ij}(t, h_{ij}, \widehat{h}_{ij}(t))\right]^2 \operatorname{sign}(\Delta_j) + \sum_{i=0}^{r} \alpha_{ij}(t, h_{ij}, \widehat{h}_{ij}(t)).$$

From assumption H1 it follows that $\left|\alpha_{ij}(t, h_{ij}, \widehat{h}_{ij}(t))\right| \le c_1 \left|h_{ij} - \widehat{h}_{ij}(t)\right|$, and from assumption H3: $|g_{ij}(t, h_{ij})| > \delta$, for at least one $i = 1, ..., r$.

So, similar as in the proof of the previous Theorem we obtain that sliding mode occurs on $\Delta_j(t) \equiv 0$ for initial conditions of $\widehat{h}_{ij}$ sufficiently close to h_{ij}. Using the assumption H2 it follows the convergence of the estimate. ◀

5.3 Delay Estimation Algorithm C

Algorithms A, and B require measurements of the derivatives which can be obtained via sliding modes, but another version of the algorithm does not require these derivatives. For this, let:

$$\widetilde{\Delta}(t) = x(t) - x(0) - \sum_{i=0}^{r} \int_0^t f_i(\xi, \widehat{h}_i(t)) d\xi \tag{55}$$

where the functions f_i are given by (33). We propose the following delay estimator for $i = 1, \ldots, r$:

$$\frac{d\widehat{h}_i(t)}{dt} = -L_i \left\langle \widetilde{f}_i\left(t, \widehat{h}_i(t)\right), \mathrm{sign}(\widetilde{\Delta}) \right\rangle \mathrm{st}(\widehat{h}_i(t))\,\mathrm{st}(h - \widehat{h}_i(t)), \tag{56}$$

where the functions $\widetilde{f}_i(t, \widehat{h}_i)$ are

$$\begin{aligned} \widetilde{f}_i(t, \widehat{h}_i) &= A_i\left[x\left(t - \widehat{h}_i\right) - x\left(-\widehat{h}_i\right)\right] + \\ B_i\left[u\left(t - \widehat{h}_i\right) - u\left(-\widehat{h}_i\right)\right] &= f_i(t, \widehat{h}_i) - f_i(0, \widehat{h}_i). \end{aligned} \tag{57}$$

Here $x(-\widehat{h}_i(t)) = x_o(-\widehat{h}_i(t))$, and $u(-\widehat{h}_i(t)) = u_o(-\widehat{h}_i(t))$ are using the values of the initial functions for x, u, and h is the upper limit of the delays $h_1, ..., h_r$. The initial conditions for the estimation algorithm (56) are assumed to be in the interval $0 \le h_i(0) \le h$.

Let's note here, that the function $\mathrm{st}(\widehat{h}_i(t))\,\mathrm{st}(h - \widehat{h}_i(t))$ in the right-hand side of (56) guarantees that $0 \le \widehat{h}_i(t) \le h$ for all $t \ge 0$, so the values $x(-\widehat{h}_i(t))$, and $u(-\widehat{h}_i(t))$ are well defined.

Differentiating $\widetilde{\Delta}$, and using notations f_i,g_i (see (33), (35)), we obtain

$$\dot{\widetilde{\Delta}}(t) = \dot{x}(t) - \sum_{i=0}^{r} f_i(t, \widehat{h}_i(t)) + \sum_{i=0}^{r} \int_0^t g_i(\xi, \widehat{h}_i(t)) d\xi \frac{d\widehat{h}_i(t)}{dt}. \tag{58}$$

Using the fact that

$$\int_0^t g_i(\xi, \widehat{h}_i(t)) d\xi = f_i(t, \widehat{h}_i(t)) - f_i(0, \widehat{h}_i(t)) = \widetilde{f}_i\left(t, \widehat{h}_i(t)\right),$$

similarly, to the case of the Algorithm A we obtain the convergence proof based on the use of the Lyapunov function $V(\widetilde{\Delta}) = \|\widetilde{\Delta}\|_1 = \left\langle \widetilde{\Delta}, \mathrm{sign}(\widetilde{\Delta}) \right\rangle$.

The following theorem states local convergence of the Algorithm C.

Theorem 4 *Let us assume that there exist a finite time $t_1 > 0$ such that the following holds for $0 \le t \le t_1$:*

H1: The input $u(t)$ is continuous and bounded. The system vector $x(t)$ is, at least, once continuously differentiable and have bounded derivative.

H2: The system (17) is identifiable with respect to delay vector for the input $u(t), t \ge -h$, and the initial condition $x_o(t), -h \le t \le 0$. In particular, $\forall t : 0 < t < t_1\ \exists c_1 : c_1 > 0$ such that $\|\mathcal{H}' - \mathcal{H}\|_1 \le c_1 \|\widetilde{\Delta}(t, \mathcal{H}')\|_1$ for all $\mathcal{H}' \in \mathcal{B}_\varepsilon(\mathcal{H})$ where $\mathcal{B}_\varepsilon(\mathcal{H})$ is some ball of the radius $\varepsilon > 0 (\exists \varepsilon > 0)$, with the center in $\mathcal{H}$.

H3: There exists $\delta > 0$, that $\left|\left\langle \widetilde{f}_i(t, h_i), \beta \right\rangle\right| > \delta$ for any vector $\beta = [\beta_1, ..., \beta_n]^T \in \mathbb{R}^n$, such that $|\beta_j| = 1$, for all $j = 1, ..., n$.

Then, using the delay estimator (36) with initial conditions $\widehat{\mathcal{H}}(0) = \left[\widehat{h}_1(0), \ldots, \widehat{h}_r(0)\right]^T$ ($\widehat{h}_j(0) \ge 0, \forall j = 1, ..., r$), close enough to $\mathcal{H} = [h_1, \ldots, h_r]^T$ ($\exists \epsilon > 0, \widehat{\mathcal{H}}(0) \in \mathcal{B}_\epsilon(\mathcal{H})$) there exist finite gains L_i such that after a finite time $t_0, (0 \leqq t_0 \le t_1)$ we have $\widehat{h}_i(t) \equiv h_i$ for all $i = 0, ..., r$, $t_0 \le t \le t_1$.

6 Variable delay estimation in linear systems

We'll show here that the same algorithms A,B, and C without any changes can be used for estimation of the variable delay in the system (22):

$$\dot{x}(t) = \sum_{i=0}^{r} \left[A_i x(t - h_i(t)) + B_i u(t - h_i(t))\right].$$

The only distinction between constant and variable delay cases are the conditions for the convergence. Namely, the following theorems are true:

Theorem 5 *Let us assume that there exist a finite time $t_1 > 0$ such that the following holds for $0 \le t \le t_1$:*

H1: The input $u(t)$ is continuously differentiable and has bounded derivative. The system vector $x(t)$ is twice continuously differentiable and have bounded derivatives. The variable delay vector function $\mathcal{H}(t) = [h_1(t), ..., h_n(t)]^T$ is bounded by some $h > 0$, i.e. $0 \le h_i(t) \le h$, for every $i = 1, ..., n$, and every $h_i(t)$ is continuously differentiable with bounded derivative: $0 \le \dot{h}_i(t) \le h'$.

H2: The system (22) is identifiable with respect to delay vector $\mathcal{H}(t)$ for the input $u(t), t \ge -h$, and the initial condition $x_o(t), -h \le t \le 0$. In particular, $\forall t : 0 \le t \le t_1\ \exists c_1 : c_1 > 0$ such that $\|\mathcal{H}'(t) - \mathcal{H}(t)\|_1 \le c_1 \|\Delta(t, \mathcal{H}'(t))\|_1$ for all $\mathcal{H}'(\cdot) \in \widetilde{\mathcal{B}}_\varepsilon(\mathcal{H}(\cdot))$ where $\widetilde{\mathcal{B}}_\varepsilon(\mathcal{H}(\cdot))$ is some ball in $C^1_{[-h,+\infty]}$ of the radius $\varepsilon > 0 (\exists \varepsilon > 0)$, with the center in $\mathcal{H}(\cdot)$.

H3: There exists $\delta > 0$*, that* $|\langle g_i(t, h_i(t)), \beta\rangle| > \delta$ *for any vector* $\beta = [\beta_1, ..., \beta_n]^T \in \mathbb{R}^n$*, such that* $|\beta_j| = 1$*, for all* $j = 1, ..., n$ *.*

Then, using the delay estimator (36) with initial conditions $\widehat{\mathcal{H}}(0) = \left[\widehat{h}_1(0), \ldots, \widehat{h}_r(0)\right]^T$ $(\widehat{h}_j(0) \geq 0, \forall j = 1, ..., r)$*, close enough to* $\mathcal{H}(0) = [h_1(0), \ldots, h_r(0)]^T$ $(\exists \epsilon > 0, \widehat{\mathcal{H}}(0) \in \mathcal{B}_\epsilon(\mathcal{H}(0)))$ *there exist finite gains* L_i *such that after a finite time* $t_0, (0 \leqq t_0 < t_1)$ *we have* $\widehat{h}_i(t) \equiv h_i(t)$ *for all* $i = 0, ..., r$*,* $t_0 \leq t \leq t_1$*.*

Proof (Algorithm Convergence): ► The proof repeats the same steps as the proof of the Theorem 2 with the following modifications:

- The expression (37) contains additional terms depending on $\dot{h}_i(t)$:

$$\begin{aligned}\dot{\Delta} &= \sum_{i=1}^{r} \Big[g_i(t, h_i(t)) - g_i(t, \widehat{h}_i(t)) + g_i(t, h_i(t))\dot{h}_i(t) \\ &- L_i \left\langle g_i(t, \widehat{h}_i(t)), \mathrm{sign}(\Delta)\right\rangle g_i(t, \widehat{h}_i(t))\Big].\end{aligned}$$

- As a result, the expression for $V_{\mathrm{dot}}(\Delta)$ (41) will include a bounded function $\gamma_i(t) = g_i(t, h_i(t))\dot{h}_i(t)$ (assumption H1)

$$\begin{aligned}V_{\mathrm{dot}}(\Delta) &= \sum_{i=0}^{r} \left\langle \alpha_i(t, h_i, \widehat{h}_i) + \gamma_i(t), \mathrm{sign}(\Delta)\right\rangle \\ &- \sum_{i=0}^{r} L_i \left\langle g_i(t, h_i(t)) - \alpha_i(t, h_i, \widehat{h}_i), \mathrm{sign}(\Delta)\right\rangle^2.\end{aligned}$$

- Using boundedness of $\gamma_i(t)$ (42) can be written with an additional constant $C_0 > 0$ in the right hand side as

$$V_{\mathrm{dot}}(\Delta) \leq C_0 + C_1 V - \sum_{i=0}^{r} L_i \langle g_i(t, h_i(t)), \mathrm{sign}(\Delta)\rangle^2, (C_1 > 0).$$

- Finally, using the assumption H3 we obtain that:

$$V_{\mathrm{dot}}(\Delta) \leq C_0 + C_1 V - C_2,$$

  where $C_2 > C_0$ for sufficiently large L_i.

The last inequality guarantees that $V_{\mathrm{dot}}(\Delta) < -\delta_1$ for $V < (C_2 - C_0 - \delta_1)/C_1$. Which as in the Theorem 2 completes the proof. ◄

Similar proof can be obtained for the variable delay version of the Theorem 3 :

Theorem 6 *Let us assume that for the system (44) with variable* $h_{ij}(t)$ *there exist a finite time* $t_1 > 0$ *such that the following holds for* $0 \leq t \leq t_1$*:*

H1: The input $u(t)$ is continuously differentiable and has bounded derivative. The system vector $x(t)$ is twice continuously differentiable and have bounded derivatives. Every entry of the variable delay matrix function $\mathcal{H}(t) = [h_{ij}(t))]$ is bounded by some $h > 0$, i.e. $0 \le h_{ij}(t) \le h$, for every $i = 1, ..., r$, $j = 1, ..., n$, and every $h_{ij}(t)$ is continuously differentiable with bounded derivative: $0 \le \dot{h}_{ij}(t) \le h'$.

H2: The system (44) with variable $h_{ij}(t)$ is identifiable with respect to delay matrix $\mathcal{H}(t)$ for the input $u(t), t \ge -h$, and the initial condition $x_o(t), -h \le t \le 0$. In particular, $\forall t : 0 \le t \le t_1$ $\exists c_1 : c_1 > 0$ such that $\|\mathcal{H}'(t) - \mathcal{H}(t)\|_1 \le c_1 \|\Delta(t, \mathcal{H}'(t))\|_1$ for all $\mathcal{H}'(\cdot) \in \widetilde{\mathcal{B}}_\varepsilon(\mathcal{H}(\cdot))$ where $\widetilde{\mathcal{B}}_\varepsilon(\mathcal{H}(\cdot))$ is some ball in $C^1_{[-h,+\infty]}$ of the radius $\varepsilon > 0$($\exists \varepsilon > 0$), with the center in $\mathcal{H}(\cdot)$.

H3: There exists $\delta > 0$, that for any $j = 1, ..., n$, there exists at least one $i = 1, ..., r$ such that $|g_{ij}(t, h_{ij}(t))| > \delta$ for all $0 \le t \le t_1$.

Then, using the delay estimator (49) with initial conditions $\widehat{\mathcal{H}}(0) = \left[\widehat{h}_{ij}(0)\right]$ ($\widehat{h}_{ij}(0) \ge 0, \forall j = 1, ..., r;\ i = 1, ..., n$), close enough to $\mathcal{H}(0)$ ($\exists \epsilon > 0, \widehat{\mathcal{H}}(0) \in \mathcal{B}_\epsilon(\mathcal{H}(0))$) there exist finite gains L_i such that after a finite time $t_0, (0 \leqq t_0 \le t_1)$ we have $\widehat{h}_{ij}(t) \equiv h_{ij}(t)$ for all $i = 0, ..., r;\ i = 1, ..., n,\ t_0 \le t \le t_1$.

The following theorem states local convergence of the Algorithm C for a time varying case .

Theorem 7 *Let us assume that for the system (22) there exist a finite time $t_1 > 0$ such that the following holds for $0 \le t \le t_1$:*

H1: The input $u(t)$ is continuous and bounded. The system vector $x(t)$ is, at least, once continuously differentiable and have bounded derivative. Every entry of the variable delay vector function $\mathcal{H}(t) = [h_1(t), ..., h_r(t))]^T$ is bounded by some $h > 0$, i.e. $0 \le h_i(t) \le h$, for every $i = 1, ..., r$, and every $h_i(t)$ is continuously differentiable with bounded derivative: $0 \le \dot{h}_i(t) \le h'$.

H2: The system (22) is identifiable with respect to delay vector $\mathcal{H}(t)$ for the input $u(t), t \ge -h$, and the initial condition $x_o(t), -h \le t \le 0$.

H3: There exists $\delta > 0$, that $\left|\left\langle \widetilde{f}_i(t, h_i(t)), \beta \right\rangle\right| > \delta$ for any vector $\beta = [\beta_1, ..., \beta_n]^T \in \mathbb{R}^n$, such that $|\beta_j| = 1$, for all $j = 1, ..., n$.

Then, using the delay estimator (56) with initial conditions $\widehat{\mathcal{H}}(0) = \left[\widehat{h}_1(0), \ldots, \widehat{h}_r(0)\right]^T$ ($\widehat{h}_j(0) \ge 0, \forall j = 1, ..., r$), close enough to $\mathcal{H}(0) = [h_1(0), \ldots, h_r(0)]^T$ ($\exists \epsilon > 0, \widehat{\mathcal{H}}(0) \in \mathcal{B}_\epsilon(\mathcal{H}(0))$) there exist finite gains L_i such that after a finite time $t_0, (0 \leqq t_0 \le t_1)$ we have $\widehat{h}_i(t) \equiv h_i(t)$ for all $i = 0, ..., r,\ t_0 \le t \le t_1$.

Remark 1 *All algorithms described above can be used for the linear systems of the type (17), (44), (21), (18), with additional known term in the right hand side, such as, for example $W(t)$ in the variable delay model:*

$$\dot{x}(t) = \sum_{i=0}^{r} [A_i x(t - h_i(t)) + B_i u(t - h_i(t))] + W(t),$$

The corresponding modification of Δ is

$$\Delta = \dot{x}(t) - \sum_{i=0}^{r} \left[A_i x(t - \widehat{h}_i(t)) + B_i u(t - \widehat{h}_i(t)) \right] - W(t).$$

7 Sliding-mode observers for distributed delay estimation in linear systems

Let us now consider the model (18)

$$\dot{x}(t) = \int_0^h [d\mu(\xi)x(t - \xi) + d\nu(\xi)u(t - \xi)],$$

where, as was mentioned earlier, $\mu(\xi)$ and $\nu(\xi)$ are matrix valued functions of bounded variation, or matrix valued measures (not necessarily positive).

An estimation algorithm, which is an analog of the algorithm (36) is described by a functional equation. Namely, for $\phi(\xi)$, and $\psi(\xi)$ arbitrary continuously differentiable functions equal to zero outside the interval $[0, h]$, the estimates $\hat{\mu}(\xi, t)$, and $\hat{\nu}(\xi, t)$ satisfy the following equation:

$$\begin{aligned} &\int_0^h [d\hat{\mu}'_t(\xi, t)\phi(\xi) + d\hat{\nu}'_t(\xi, t)\psi(\xi)] \\ &= L_1(t) \int_0^h [d\hat{\mu}(\xi, t)\phi'(\xi) + d\hat{\nu}(\xi, t)\psi'(\xi)] \\ &\times \int_0^h [\dot{x}^T(t - \xi)d\hat{\mu}^T(\xi, t) + \dot{u}^T(t - \xi)d\hat{\nu}^T(\xi, t)] sign[\Delta(t)] \\ &+ L_2(t) \int_0^h [d\hat{\mu}(\xi, t)\phi(\xi) + d\hat{\nu}(\xi, t)\psi(\xi)] \\ &\times \int_0^h [x^T(t - \xi)d\hat{\mu}^T(\xi, t) + u^T(t - \xi)d\hat{\nu}^T(\xi, t)] sign[\Delta(t)], \end{aligned} \quad (59)$$

where, as $\hat{\mu}'_t(\xi, t), \hat{\nu}'_t(\xi, t)$ we denote derivatives of time-varying measures $\hat{\mu}(\xi, t), \hat{\nu}(\xi, t)$ with respect to time.

$L_1(t) > 0$, and $L_2(t) > 0$ are the algorithm's gains.

The function $\Delta(t, \widehat{\widehat{\mu}}))$ which defines the sliding manifold is

$$\Delta = \dot{x}(t) - \int_0^h [d\hat{\mu}(\xi, t)x(t - \xi) + d\hat{\nu}(\xi, t)u(t - \xi)]. \quad (60)$$

Theorem 8 *Let us assume that there exist a finite time $t_1 > 0$ such that the following holds for $0 \leq t \leq t_1$:*

H1: The input $u(t)$ is, at least once, continuously differentiable and have bounded derivative. The system vector $x(t)$ is, at least, twice continuously differentiable and have bounded derivatives.

H2: The system (18) is identifiable with respect to $\widetilde{\mu}$ for the input $u(t), t \geq -h$, and the initial condition $x_o(t), -h \leq t \leq 0$.

H3 There exist $\delta > 0$ such that

$$\left| \left\langle \int_0^h [d\mu(\xi,t)\dot{x}(t-\xi) + d\nu(\xi,t)\dot{u}(t-\xi)], \beta \right\rangle \right| > \delta,$$

or

$$\left| \left\langle \int_0^h [d\mu(\xi,t)x(t-\xi) + d\nu(\xi,t)u(t-\xi)], \beta \right\rangle \right| > \delta,$$

for any vector $\beta = [\beta_1, ..., \beta_n]^T \in \mathbb{R}^n$, such that $|\beta_j| = 1$, for all $j = 1, ..., n$.

Then, using the delay estimator defined by the equation (59) with initial conditions $\hat{\mu}(\xi,0), \hat{\nu}(\xi,0)$ close enough to $\mu(\xi), \nu(\xi)$ there exist finite gains L_i such that after a finite time $t_0, (0 \leqq t_0 \leq t_1)$ we have $t \geq t_2$ we have $\hat{\mu}(\xi,t) \equiv \mu(\xi)$.

Proof (Algorithm (59) Convergence): ▶ Differentiating Δ given by (60) with respect to time we obtain

$$\begin{aligned} \dot{\Delta} &= \ddot{x}(t) - \int_0^h [d\hat{\mu}(\xi,t)\dot{x}(t-\xi) + d\hat{\nu}(\xi,t)\dot{u}(t-\xi)] \\ &\quad - \int_0^h \left[d\hat{\mu}_t'(\xi,t)x(t-\xi) + d\hat{\nu}_t'(\xi,t)u(t-\xi)\right] \\ &= - \int_0^h \left[d\hat{\mu}_t'(\xi,t)x(t-\xi) + d\hat{\nu}_t'(\xi,t)u(t-\xi)\right] + \Phi(t,\hat{\mu},\hat{\nu}), \end{aligned}$$

where we denoted as Φ the first two terms in the above expression:

$$\Phi(t,\hat{\mu},\hat{\nu}) = \ddot{x}(t) - \int_0^h [d\hat{\mu}(\xi,t)\dot{x}(t-\xi) + d\hat{\nu}(\xi,t)\dot{u}(t-\xi)].$$

$\Phi(t,\hat{\mu},\hat{\nu})$ depends on the measures estimates $\hat{\mu}, \hat{\nu}$, such that on the actual measures its value is zero:

$$\Phi(t,\mu,\nu) \equiv 0.$$

since this expression is just the time derivative of the equation (18).

Now, let us consider, as in the proofs of the above theorems the following nonsmooth Lyapunov function

$$V(\Delta) = \|\Delta\|_1 = \langle \mathrm{sign}(\Delta), \Delta \rangle \,, \tag{61}$$

Differentiating V in a non smooth context as before (this is using generalized gradient see [4])

$$\dot{V} \in \cup_{w \in \partial_C V} \left\{ \left\langle w, \dot{\Delta} \right\rangle \right\},$$

where

$$\begin{aligned}
\dot{V}(\Delta) &= V_{\mathrm{dot}}(\Delta) \text{ if } \Delta_i \neq 0, \text{ for } i = 1, ..., n \\
V_{\mathrm{dot}}(\Delta) &= -\left\langle \mathrm{sign}(\Delta), \int_0^h \left[d\hat{\mu}_t'(\xi,t)x(t-\xi) + d\hat{\nu}_t'(\xi,t)u(t-\xi) \right] \right\rangle \\
&\quad + \langle \mathrm{sign}(\Delta), \Phi(t, \hat{\mu}, \hat{\nu}) \rangle \,,
\end{aligned}$$

Let's fix t, and choose the functions $\phi(\xi)$, and $\psi(\xi)$ as

$$\phi(\xi) = x(t-\xi) sp_{\varepsilon,h}(\xi)$$

$$\psi(\xi) = u(t-\xi) sp_{\varepsilon,h}(\xi).$$

This functions are continuously differentiable, and they are identically zero outside $[0, h]$.

Since, obviously,

$$\int_0^h \left[d\hat{\mu}_t'(\xi,t)x(t-\xi) + d\hat{\nu}_t'(\xi,t)u(t-\xi) \right] = \int_0^h \left[d\hat{\mu}_t'(\xi,t)\phi(\xi) + d\hat{\nu}_t'(\xi,t)\psi(\xi) \right] + o(\varepsilon),$$

where $o(\varepsilon) \to 0$ uniformly with respect to t, when $\varepsilon \to 0$, and $\phi(\xi), \psi(\xi)$ satisfy conditions mentioned after the equation (59). Using this equation we obtain:

$$
\begin{aligned}
\dot V(\Delta) &= V_{\rm dot}(\Delta) \text{ if } \Delta_i \neq 0,\ \text{for } i=1,...,n\\
V_{\rm dot}(\Delta) &= -\left\langle \mathrm{sign}(\Delta), \int_0^h \left[d\hat\mu_t'(\xi,t)\phi(\xi) + d\hat\nu_t'(\xi,t)\psi(\xi)\right]\right\rangle\\
&+ \langle \mathrm{sign}(\Delta), \Phi(t,\hat\mu,\hat\nu)\rangle + o(\varepsilon)\\
&= -\left\langle \mathrm{sign}(\Delta), L_1(t)\int_0^h [d\hat\mu(\xi,t)\phi'(\xi) + d\hat\nu(\xi,t)\psi'(\xi)]\times\right.\\
&\times \left.\int_0^h [\dot x^T(t-\xi)d\hat\mu^T(\xi,t) + \dot u^T(t-\xi)d\hat\nu^T(\xi,t)]\,\mathrm{sign}(\Delta)\right\rangle\\
&- \left\langle \mathrm{sign}(\Delta), L_2(t)\int_0^h [d\hat\mu(\xi,t)\phi(\xi) + d\hat\nu(\xi,t)\psi(\xi)]\times\right.\\
&\times \left.\int_0^h [x^T(t-\xi)d\hat\mu^T(\xi,t) + u^T(t-\xi)d\hat\nu^T(\xi,t)]\,\mathrm{sign}(\Delta)\right\rangle\\
&\quad + \langle \mathrm{sign}(\Delta), \Phi(t,\hat\mu,\hat\nu)\rangle + o(\varepsilon)\\
&= -L_1(t)\left\langle \mathrm{sign}(\Delta), \int_0^h [d\hat\mu(\xi,t)\dot x(t-\xi) + d\hat\nu(\xi,t)\dot u(t-\xi)]\right\rangle^2\\
&- L_2(t)\left\langle \mathrm{sign}(\Delta), \int_0^h [d\hat\mu(\xi,t)x(t-\xi) + d\hat\nu(\xi,t)u(t-\xi)]\right\rangle^2\\
&\quad + \langle \mathrm{sign}(\Delta), \Phi(t,\hat\mu,\hat\nu)\rangle + o(\varepsilon)
\end{aligned}
$$

On the other hand, introducing

$$
\begin{aligned}
\alpha(t) &= \int_0^h [d\hat\mu(\xi,t)x(t-\xi) + d\hat\nu(\xi,t)u(t-\xi)]\\
&- \int_0^h [d\mu(\xi,t)x(t-\xi) + d\nu(\xi,t)u(t-\xi)]
\end{aligned}
$$

we obtain

$$
\begin{aligned}
V_{\rm dot}(\Delta) &= -L_1\left\langle \mathrm{sign}(\Delta), \int_0^h [d\mu(\xi,t)\dot x(t-\xi) + d\nu(\xi,t)\dot u(t-\xi)] + \Phi(t)\right\rangle^2\\
&- L_2\left\langle \mathrm{sign}(\Delta), \int_0^h [d\mu(\xi,t)x(t-\xi) + d\nu(\xi,t)u(t-\xi)] + \alpha(t)\right\rangle^2\\
&\quad + \langle \mathrm{sign}(\Delta), \Phi(t)\rangle + o(\varepsilon)
\end{aligned}
$$

Using the assumptions H1-H3 we obtain, similar, as in the proof of the theorem 2, the following estimate:

$$
V_{\rm dot}(\Delta) \le C_0 - C_1 V + o(\varepsilon)
$$

for the initial estimates sufficiently close to the actual values of μ, and ν in the sense of a norm described above.

This inequality implies local convergence of the proposed algorithm. ◀

8 Examples

Example 1

In order to illustrate our approach, as a first example, we consider the following classical first order system without inputs, and with unknown delay:

$$\dot{x}(t) = a_0 x(t) + a_1 x(t-h). \tag{62}$$

The delay estimation algorithm has the form:

$$\dot{\hat{h}}(t) = -La_1 \dot{x}(t-\hat{h}(t)) sign(\Delta(t)) sg(\hat{h}(t)), \tag{63}$$

where $\Delta = \dot{x}(t) - a_0 x(t) - a_1 x(t-\hat{h}(t))$. We used the following parameters: $a_0 = -0.5$, $a_1 = 1.0$, and $h = 1.0$.

The plots below show the simulation experiments with initial value $x(0) = 1$, and zero initial function of the delay element ($x(\tau) = 0$ for $\tau \in [-1, 0)$). The observer gain was $L = 0.3$. The Euler simulation algorithm was used, since, as it is well known, the simulation of sliding mode algorithms require special attention. Systems with sliding modes from numerical point of view are "rigid" and such methods as Runge-Kutta can run into numerical difficulties. Higher order Adams methods may be applied for better accuracy.

The plots of $\hat{h}$ shows that the observer for the delay converges for different initial conditions, when the initial delay estimates is above (Fig. 3) and below (Fig. 4) the actual delay. In Fig. 2 we show the plot of the switching function Δ for the first case.

Example 2

As the second example we consider a first order system with two unknown delays:

$$\dot{x}(t) = ax(t-h_1) + bu(t-h_2). \tag{64}$$

The following numerical values were used $a = -0.1, b = 1, h_1 = 2, h_2 = 3$. The simulations were carried out for the initial condition 1 for x on the interval $[-2, 0]$.

According to our notations the equation (64) should be written as

$$\dot{x}(t) = \sum_{i=1}^{2} A_i x(t-h_i) + B_i u(t-h_i), \tag{65}$$

Figure 2: Example 1. Delta.

where $A_1 = a$, $A_2 = 0$, $B_1 = 0$, and $B_2 = b$. Therefore, the Algorithm A has the form

$$\begin{aligned} \dot{\hat{h}}_1 &= -L_1 g_1 sign(\Delta) st(\hat{h}_1) \\ \dot{\hat{h}}_2 &= -L_2 g_2 sign(\Delta) st(\hat{h}_2), \end{aligned} \tag{66}$$

where $\Delta = \dot{x}(t) - ax(t - \hat{h}_1) - bu(t - \hat{h}_2)$, $g_1 = a\dot{x}(t - \hat{h}_1)$, and $g_2 = b\dot{u}(t - \hat{h}_2)$. As the input u we used differentiable randomly chosen input. The plots depicting the behavior of the delays estimates $\hat{h}_1$ and $\hat{h}_2$ and the switching function Δ are shown in Fig . 5-7.

Figure 3: Example 1. Plot of $\hat{h}$ for the initial estimate (0.8) below the actual delay.

Example 3

As the third example we'll consider the situation when it is required to estimate simultaneously parameters and delays of a system distributed over a network.

One of the practical setups can be described as following: we have a network with three nodes is shown in the Fig. 8 . There is an unknown delay in information exchange between nodes.

The systems at the second and third nodes (Subsystems II and III) are controlled by the controller located at the first node (Subsystem I). The subsystems II and III are described by the equations:

$$\dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t - h'_3) + B_1u_1(t - h_1) \quad (67)$$

$$\dot{x}_2(t) = A_{21}x_2(t) + A_{22}x_1(t - h_3) + B_2u_2(t - h_2). \quad (68)$$

At the first node we can measure the states of subsystems II and III with the delays

$$y_1(t) = x_1(t - h_1), \quad (69)$$

$$y_2(t) = x_2(t - h_2), \quad (70)$$

correspondingly. Some (or all) delays and subsystems' parameters are unknown.

Figure 4: Example 1. Plot of $\hat{h}$ for the initial estimate (1.2) above the actual delay.

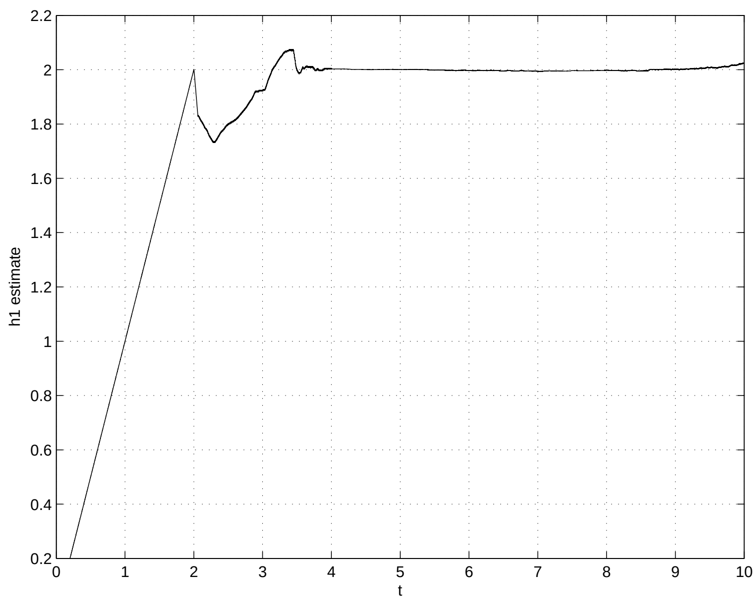

Figure 5: Estimation of delay h_1.

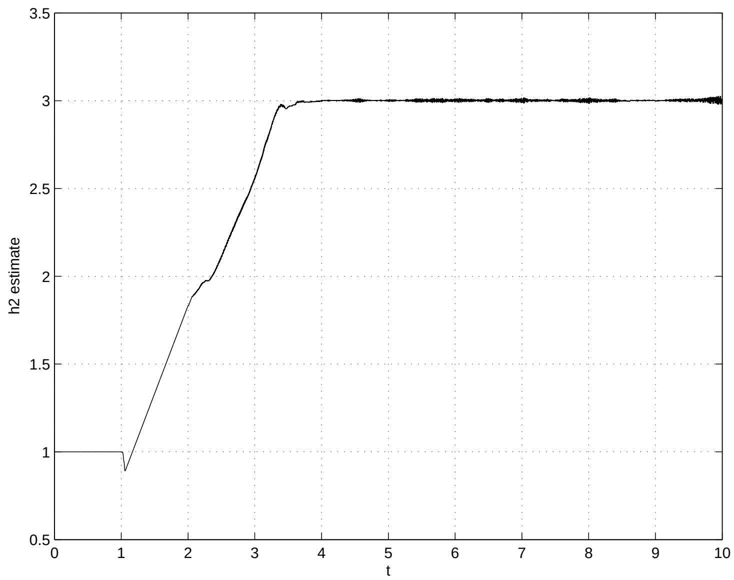

Figure 6: Estimation of delay h_2.

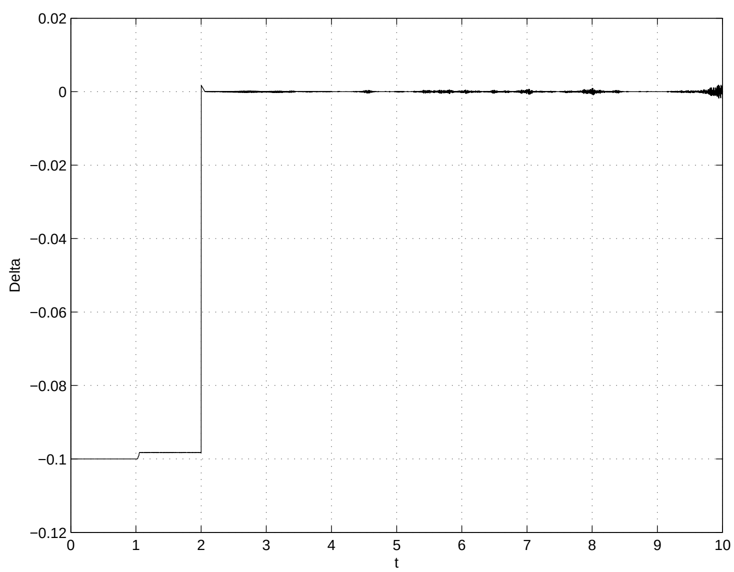

Figure 7: Delta.

Figure 8: Network example.

As a part of the controller located at the node I, we need an observer which can estimate the delays and parameters.

Using time scale at the first node, from (67 - 70) we obtain that the measured variables y_1 and y_2 satisfy the equations:

$$\begin{aligned} \dot{y}_1(t) &= A_{11}y_1(t) + A_{12}y_2(t - h_1 - h'_3) + B_1u_1(t - 2h_1) \\ \dot{y}_2(t) &= A_{21}y_2(t) + A_{22}y_1(t - h_2 - h_3) + B_2u_2(t - 2h_2). \end{aligned} \tag{71}$$

Let us consider a particular example when both subsystems II and III are described by the first order equations with one unknown parameter $a = A_{11} = A_{21}$ and all unknown transmission delays are equal $h_1 = h_2 = h_3 = h'_3 = h/2$. Also, for simplicity, we'll assume that the other parameters are known $a_{12} = 1, a_{22} = -1$, and $B_{ij} = 0$, (or $u_i(t) = 0$). The more general case can be dealt with, in a similar manner.

$$\begin{aligned} \dot{y}_1(t) &= ay_1(t) + y_2(t - h) \\ \dot{y}_2(t) &= ay_2(t) - y_1(t - h). \end{aligned} \tag{72}$$

The right hand side of this system can be written in the integral form

$$\dot{y}(t) = \int_0^h d\mu(\xi)y(t - \xi),$$

where $y = [y_1, y_2]^T$, and μ is a corresponding *unknown* matrix measure

$$d\mu(\xi) = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \delta(\xi)d\xi + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \delta(\xi - h)d\xi.$$

Therefore,

$$d\hat{\mu}(t,\xi) = \begin{bmatrix} \hat{a}(t) & 0 \\ 0 & \hat{a}(t) \end{bmatrix} \delta(\xi) d\xi + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \delta(\xi - \hat{h}(t)) d\xi,$$

or, equivalently, the corresponding switching functions are

$$\begin{aligned} \Delta_1 &= \dot{y}_1(t) - \hat{a}(t) y_1(t) - y_2(t - \hat{h}(t)) \\ \Delta_2 &= \dot{y}_2(t) - \hat{a}(t) y_2(t) + y_1(t - \hat{h}(t)). \end{aligned} \tag{73}$$

In our simulation experiments we used the values $a = -1, h = 1$, and zero initial estimates $\hat{a}(0) = 0$, $\hat{h}(0) = 0$. The plots of the switching functions Δ_1, and Δ_2 entering the sliding mode are shown in Fig. 9, Fig. 10, and corresponding estimates $\hat{a}$ and $\hat{h}$ in Fig. 11, Fig. 12.

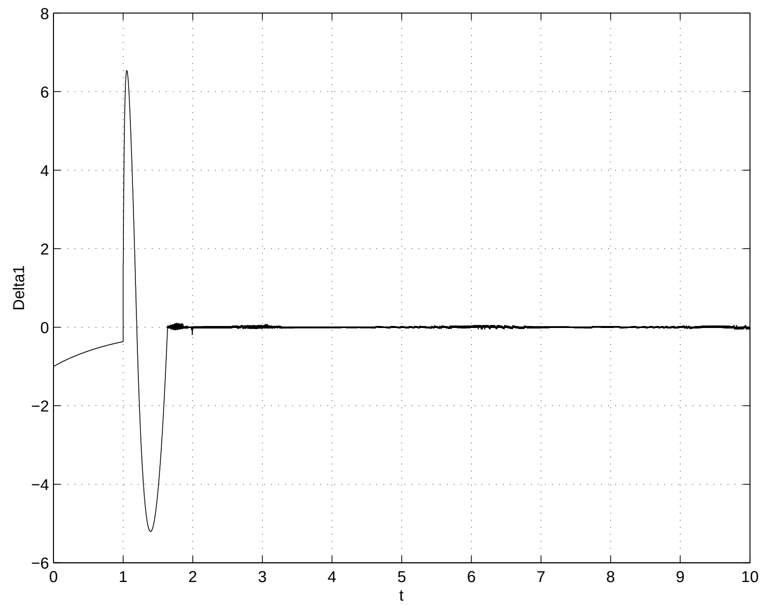

Figure 9: The switching function Δ_1 .

9 Conclusion

We presented variable structure identification algorithms which allow simultaneous delay and parameter estimation. These algorithms are based on the use of system model as sliding surfaces, where instead of the unknown parameters, or delays we substitute their corresponding estimates. Then the estimation algorithm (equation for the estimates) is designed in such a way which guarantees

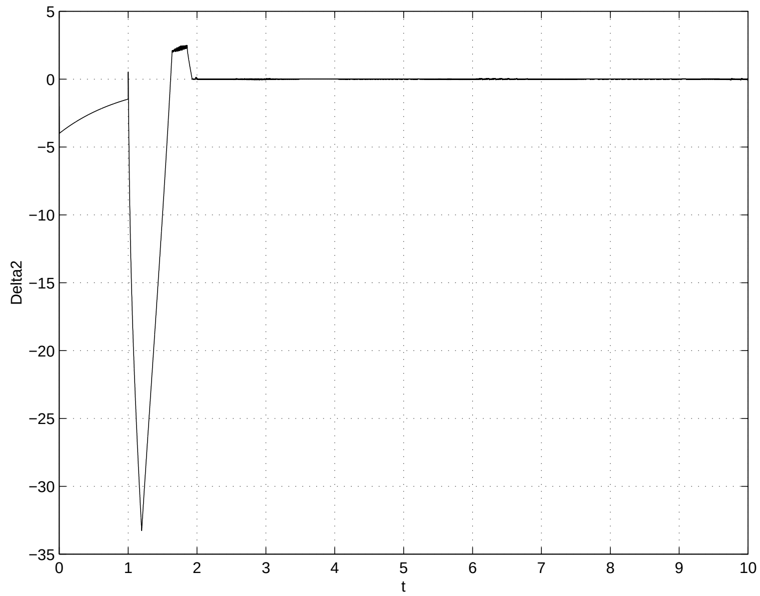

Figure 10: The switching function Δ_2 .

convergence to these sliding surfaces. So, unlike in traditional approaches, where one designs sliding surfaces for the given system, we used the system model as the "sliding surface", and designed "the system" (algorithm) where this sliding mode occurs. The idea of such algorithms can be applied not only to linear time invariant systems, but more generally to systems of almost any form. In particular, it can be extended to identification of nonlinear functional equations, and distributed systems. It is very important to note, that the time convergence of such algorithms is quite small (theoretically, arbitrary small) compared to other methods, where the natural limitations for the convergence speed exist (see, for example [24]).

References

[1] L. Belkoura. Identifiability of systems described by convolution equations. *Automatica*, 41:505–512, 2005.

[2] F. Blanchini and E.P. Ryan. A Razumikhin-type lemma for functional differential equations with application to adaptive control. *Automatica*, 35(5):809–818, 1999.

[3] H.H. Choi and M.J. Chung. Observer-based H_∞ controller design for state delayed linear systems. *Automatica*, 32(7):1073–1075, 1996.

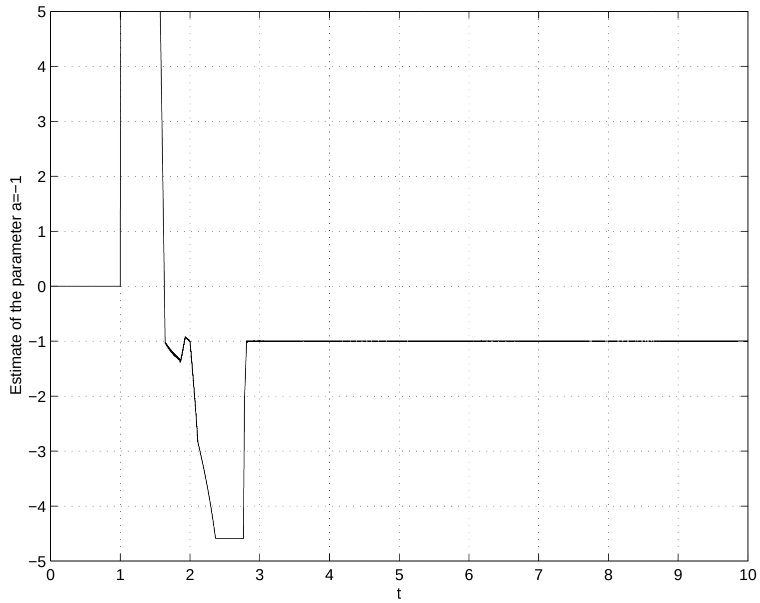

Figure 11: The parameter estimate $\hat{a}$.

[4] F.H. Clarke. Optimization and nonsmooth analysis. *Canadian Mathematical Society Series of Monographs and Advanced Texts, Editor Wiley-Interscince*, 1983.

[5] M. Darouach. Linear functional observers for systems with delays in state variables. *IEEE Trans. Aut. Control*, 46(3):491–496, March 2001.

[6] M. Darouach, P. Pierrot, and E. Richard. Design of reduced-order observers without internal delays. *IEEE Trans. Aut. Control*, 44(9):1711–1713, Sept. 1999.

[7] S. Diop, I. Kolmanovsky, P. Moraal, and M. vanNieuwstadt. Preserving stability/performance when facing an unknown time delay. *Control Eng. Practice*, 9:1319–1325, Dec. 2001.

[8] S. V. Drakunov. Adaptive quasioptimal filters/observers with discontinuous parameters. *Automation and Remote Control*, 44-9:1167–1175, 1984.

[9] S. V. Drakunov. Sliding-mode observers based on equivalent control method. *Proceedings of the 31st IEEE Conference on Decision and Control (CDC), Tucson, Arizona*, pages 2368–2370, 1992.

[10] F.W. Fairmar and A. Kumar. Delay-less observers for systems with delay. *IEEE Trans. Aut. Control*, 31(3):258–259, March 1986.

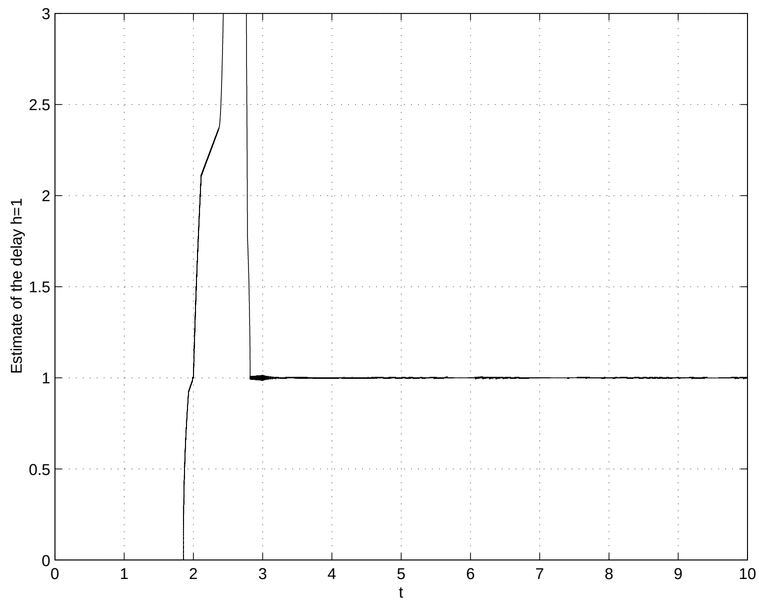

Figure 12: The delay estimate $\hat{h}$.

[11] A. F. Filippov. Differential equations with discontinuous right hand sides. *Dordrecht: Kluwer Academic Publishers*, 1988.

[12] S.G Foda and M.S. Mahmoud. Adaptive stabilization of delay differential systems with unknown uncertainty bounds. *Int. J. Control*, 71(2):259–275, 1998.

[13] A. Germani, C. Manes, and P. Pepe. A state observer for nonlinear delay systems. In 37^{th} *IEEE CDC98 (Conf. on Dec. and Control)*, pages 355–360, Tampa, FL, Dec. 1998.

[14] J.K. Hale and S.M. Verduyn-Lunel. *Introduction to Functional Differential Equations*, volume 99 of *Applied Math. Sciences*. Springer, NY, 1993.

[15] Y.P. Huang and K. Zhou. Robust stability of uncertain time delay systems. *IEEE Trans. Aut. Control*, 45(11):2169–2173, Nov. 2000.

[16] V.B. Kolmanovskii and A. Myshkis. *Introduction to the theory and applications of functional differential equations*. Kluwer Acad., Dordrecht, 1999.

[17] V.B. Kolmanovskii and V.R. Nosov. *Stability of functional differential equations*. Academic Press, London, 1986.

[18] Belkoura L. and Orlov Y. Identifiability analysis of linear delay-differential systems. *IMA J. of Mathematical Control and Information*, 19:73–81, 2002.

[19] J. Leyva-Ramos and A.E. Pearson. An asymptotic modal observer for linear autonomous time lag systems. *IEEE Trans. Aut. Control*, 40:1291–1294, July 1995.

[20] S.M.V. Lunel. Parameter identifiability of differential delay equations. *Int J. of Adapt. Control Signal Process.*, 15:655–678, 2001.

[21] S. Majhi and Atherton D.P. A novel identifcation method for time delay processes. In *ECC99 (*5^{th} *European Control Conf.)*, Karslruhe, Germany, 1999.

[22] A.S. Morse. Ring models for delay differential systems. *Automatica*, 12:529–531, 1976.

[23] S.I. Niculescu. *Delay Effects on Stability*, volume 269 of *LNCIS*. Springer, 2001.

[24] Y. Orlov, L. Belkoura, J.P. Richard, and M. Dambrine. Adaptive identification of linear time-delay systems. *Int. J. Robust and Nonlinear Control*, 13(9), July 2003, to appear. Special Issue on TDS, E. Fridman and U. Shaked.

[25] D.H. Pourboghrat F., Chyung. Parameter identification of linear delay systems. *Int. J. Contro*, 49:595–627, 1989.

[26] DeCarlo R., S. Zak, and S. Drakunov. Variable structure and sliding mode control. *chapter in The Control Handbook a Volume in the Electrical Engineering Handbook Series, CRC Press, Inc.*, 1996.

[27] J.P. Richard. Time delay systems: An overview of some recent advances and open problems. *Automatica*, 39(9):1667–1694, 2003.

[28] Drakunov S. and V. Utkin. Sliding mode control in dynamic systems. *International Journal of Control*, 55-4:1029–1037, 1992.

[29] M. Yamamoto S. Nakagiri. Unique identification of coefficient matrices, time delay and initial functions of functionnal differential equations. *Journal of Mathematical Systems, Estimation and Control*, 5, No 3:323–344, 1995.

[30] O. Sename. New trends in design of observers for time-delay systems. *Kybernetica*, 37(4):427–458, 2001.

[31] K.K. Tan, Q.K. Wang, and T.H. Lee. Finite spectrum assignment control of unstable time delay processes with relay tuning. *Ind. Eng. Chem. Res.*, 37(4):1351–1357, 1998.

[32] J. Tuch, A. Feuer, and Z.J. Palmor. Time delay estimation in continuous linear time-invariant systems. *IEEE Trans. Aut. Control*, 39:823–827, 1994.

[33] V. Utkin. Sliding modes in control optimization. *Springer-Verlag, Berlin*, 1992.

[34] S.M. Verduyn-Lunel. Identification problems in functional differential equations. In 36^{th} *IEEE CDC97 (Conf. on Dec. and Control)*, pages 4409–4413, San Diego, CA, Dec. 1997.

[35] E.I. Verriest. Robust stability and adaptive control of time-varying neutral systems. In 38^{th} *IEEE CDC99 (Conf. on Dec. and Control)*, pages 4690–4695, Phoenix, AZ, Dec. 1999.

[36] Zampieri S. Vettori P. Controllability of systems described by convolutionnal or delay-differential equations. *SIAM Journal of Control and Optimization*, 39, No 3:728–756, 2000.

[37] Z. Wang, B. Huang, and H. Unbehausen. Robust H_∞ observer design of linear state delayed systems with parametric uncertainty: The discrete-time case. *Automatica*, 35(6):1161–1167, 1999.

[38] J.P. Richard Y. Orlov, L. Belkoura and M. Dambrine. On-line parameter identification of linear time-delay. *Systemsin Proc. IEEE CDC 02, Las Vegas*, 2002.

[39] Y. Yamamoto. Reachability of a class of infinite-dimensional linear systems: An external approach with application to general neutral systems. *SIAM J. of Control and Optimization*, 27:217–234, 1989.

[40] Y.X. Yao, Y.M. Zhang, and R. Kovacevic. Functional observer and state feedback for input time-delay systems. *Int. J. Control*, 66(4):603–617, 1997.