Automatica 95 (2018) 455–462

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Useful redundancy in parameter and time delay estimation for continuous-time models ☆

Huong Ha [a,*], James S. Welsh [a], Mazen Alamir [b]

[a] School of Electrical Engineering and Computing, The University of Newcastle, Australia
[b] Gipsa-lab, Control Systems Department. University of Grenoble, France

ARTICLE INFO

Article history:
Received 11 June 2016
Received in revised form 4 March 2018
Accepted 31 May 2018
Available online 21 June 2018

Keywords:
Continuous time identification
Time delay
Useful redundancy
Global optimization

ABSTRACT

In this paper we propose an algorithm to estimate the parameters, including time delay, of continuous time systems based on instrumental variable identification methods. To overcome the multiple local minima of the cost function associated with the estimation of a time delay system, we utilize the useful redundancy technique. Specifically, the cost function is filtered through a set of low-pass filters to improve convexity with the useful redundancy technique exploited to achieve convergence to the global minimum of the optimization problem. Numerical examples are presented to demonstrate the effectiveness of the proposed algorithm.

1. Introduction

The goal of system identification is to estimate the parameters of a model in order to analyse, simulate and/or control a system. Recently, continuous-time (CT) identification methods have received attention due to advantages such as providing insights to the physical system and being independent of the sampling time (Garnier, 2015; Garnier & Young, 2013; Unbehauen & Rao, 1990; Young, 1981). For example, with irregular sample time in CT identification, the system is time-invariant. In reality, irregular sampling occurs in many cases, e.g. when the sampling is event-triggered, when the measurement is manual or in the case of missing data (Åström & Bernhardsson, 2003).

In some CT identification problems, one needs to estimate the system parameters and any unknown time delay. Many practical examples exist including, chemical processes, economic systems and biological systems, that possess time delays. It is important to estimate the delay accurately since a poor estimate can result in unsatisfactory model order selection and inaccurate estimates of the system parameters. There are many approaches to estimate a system time delay (Björklund, 2003). A simple approach is to consider the impulse response data, e.g. estimate the time delay by finding where the impulse response becomes nonzero (Carlemalm, Halvarsson, Wigren, & Wahlberg, 1999) or by noting the delay where the correlation between input and output is maximum (Carlemalm et al., 1999; Carter, 1987). Another approach is to model the delay by a rational polynomial transfer function using a Padé or similar approximation and then estimate the time delay as part of the system parameters (Agarwal & Canudas, 1987; Ahmed, Huang, & Shah, 2006; Gawthrop & Nihtil, 1985). In Baysse, Carrillo, and Habbadi (2011, 2012) and Yang, Iemura, Kanae, and Wada (2007), the time delay and system parameters of a Multiple Input Single Output CT system are estimated in a separable way using an iterative global nonlinear least-squares or instrumental variable method.

Recently, a method of estimating the parameters and time delay of CT systems has been suggested in Chen, Garnier, and Gilson (2015) which is based on a gradient technique. The parameters and the time delay are estimated separately, i.e. when one is estimated, the other is fixed which is then repeated in an iterative manner. In this approach, the Simplified Refined Instrumental Variable (SRIVC) method is used to estimate the parameters whilst the time delay is estimated using the Gauss–Newton method. In addition, due to the effects of multiple minima in the cost function to be minimized for the time delay (Björklund, Nihtilä, & Söderström, 1991; Kaminskas, 1979; Pupeikis, 1989), a low-pass filter is employed to increase convexity. As shown in Björklund et al. (1991), Eckhard, Bazanella, Rojas, and Hjalmarsson (2017) and Ferretti, Maffezzoni, and Scattolini (1996), a suitable low-pass filtering operation on the estimation data can help to extend the global convergence region of the cost function, hence improve the accuracy of the time delay estimate.

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Er-Wei Bai under the direction of Editor Torsten Söderström.

* Corresponding author.
E-mail addresses: huong.ha@uon.edu.au (H. Ha), james.welsh@newcastle.edu.au (J.S. Welsh), mazen.alamir@grenoble-inp.fr (M. Alamir).

https://doi.org/10.1016/j.automatica.2018.06.023

In this paper, we adopt the idea of using a low-pass filter to increase convexity, where instead of using only one filter, we suggest to use multiple low-pass filters and incorporate the useful redundancy technique (Alamir, 2008; Alamir, Welsh, & Goodwin, 2009). Useful redundancy is a technique to avoid local minima when solving a nonlinear inverse problem. The concept is to generate a family of cost functions that have different local minima, but share the same global minimum with the original cost function of the optimization problem. Whenever the algorithm is trapped in a minimum, the solver path is switched to another solver path in a way such that the minimum found using the new solver path corresponds to a decrease in the original cost function. This allows the algorithm to cross local minima and converge to the global minimum, hence improving the accuracy of the estimated parameters. In this paper, multiple cost functions are generated by filtering the original time delay cost function through a number of low-pass filters with different cut-off frequencies that span the system bandwidth. Note that we also published another algorithm to estimate time delay and system parameters that utilizes the multiple filtering idea (Ha & Welsh, 2016). However, the proposed algorithm in Ha and Welsh (2016) requires some priori knowledge of the initial time delay, which we overcome in this paper, by implementing the useful redundancy technique.

The paper is organized as follows. Section 2 describes the model setting and formulates the problem. Section 3 introduces the useful redundancy technique and Section 4 describes the proposed method and provides analysis on its effectiveness. Section 5 presents numerical results for both regular and irregular sampling schemes. Finally, the conclusion will be drawn in Section 6.

2. Model setting and problem formulation

Consider a continuous-time linear, time invariant, single input single output system,

$$x(t) = G_0(p)u(t-\tau_0) = \frac{B(p)}{A(p)}u(t-\tau_0), \qquad y(t) = x(t) + e(t). \tag{1}$$

with

$$B(p) = b_0p^m + b_1p^{m-1} + \cdots + b_m,$$
$$A(p) = p^n + a_1p^{n-1} + \cdots + a_n, \quad n \geq m,$$

where τ_0 is the time delay, $u(t)$, $x(t)$ are the input and deterministic output of the system respectively and p is the differential operator, i.e. $p^{(i)}x(t) = d^ix(t)/dt^i$. In addition, we make the following assumptions:

Assumption 2.1. Polynomials $B(p)$ and $A(p)$ are coprime.

Assumption 2.2. The system is asymptotically stable.

Assumption 2.3. The high frequency gain of $G_0(p)$ is 0, i.e. $G_0(p)$ is strictly proper.

Assumption 2.4. $e(t)$ is a white random process uncorrelated with $u(t)$ having intensity λ.

Furthermore, we consider the sampling time of the input, $u(t_k)$, and output, $y(t_k)$, data as either regular or irregular. The time-varying sampling interval is denoted as,

$$h_k = t_{k+1} - t_k, \qquad k = 1, 2, \ldots, N-1, \tag{2}$$

where N is the length of the data. The objective of a CT system identification problem is to estimate the time delay, τ_0, and the parameters $a_1, a_2, \ldots, a_n, b_0, b_1, \ldots, b_m$ of the CT model in (1), using the measured input $u(t_k)$ and output data $y(t_k)$, $k = 1, \ldots, N$.

There are many algorithms suggested to estimate CT system parameters, with one of the most common being the SRIVC method (Garnier & Wang, 2008; Garnier & Young, 2013; Young, 1981). Based on SRIVC, a technique (Chen et al., 2015) was developed to estimate both the time delay and parameters for the model (1). This technique considers the problem as a separable nonlinear least squares problem, i.e. the SRIVC algorithm is utilized to estimate the system parameters and the Gauss–Newton method is used to estimate the time delay. In this problem, the cost function for the time delay estimation has multiple minima (Chen et al., 2015), hence a low-pass filter is utilized to extend the global convergence region (Chen et al., 2015; Eckhard et al., 2017; Ferretti et al., 1996). A description of the SRIVC-based time delay estimation with filtering can be found in Chen et al. (2015). However, the problem of local minima still exists even with the use of a filter.

In this paper, we propose a new approach to avoid the local minima in the time delay estimation problem. From this, a new algorithm to estimate CT system parameters and time delay with a high accuracy is obtained.

3. The useful redundancy method

In this section we describe the useful redundancy technique (Alamir, 2008; Alamir et al., 2009) that we utilize to develop our proposed algorithm. The technique was originally proposed to avoid local minima when solving a non-linear inverse problem. We define the useful redundancy technique by quoting the definition directly from Alamir et al. (2009).

Definition 1. Consider an optimization problem,

$$\min_{\rho} J_0(\rho).$$

Then it is called *M-safely redundant* if and only if the following conditions hold:

(1) There exists a finite M cost functions J_i sharing the same global minimum $\rho^* \in \mathbb{R}^n$.
(2) There exists a solver (or an iterative scheme) ℓ and a finite number of iterations $r^* \in \mathbb{N}$ such that for some $\gamma \in [0, 1)$ and all $\rho \in \mathbb{R}$ the following inequality holds,

$$\Delta_N^{\gamma}(\rho) = \min_{i\in\{0,\ldots,M\}} [J_0(\ell^{(r^*)}(\rho, J_i)) - \gamma J_0(\rho)] \leq 0 \tag{3}$$

where $\ell^{(r^*)}(\rho, J_i)$ is the candidate solution obtained after r^* iterations of ℓ using the cost function J_i, starting from the initial guess ρ. ■

The solver path (ρ, J_i) is defined as the sequence of iterates $\ell^{(j)}(\rho, J_i)$ for the solver ℓ when the cost function J_i starts from an initial guess ρ. Condition 2 means that for any initial ρ, there always exists a solver path ℓ that corresponds to a decrease in the original cost J_0 after at most r^* iterations. It is proven in Alamir et al. (2009) that if an optimization problem is *M-safely redundant*, then it is possible to define an iterative algorithm such that convergence to the global minimum is guaranteed.

4. The useful redundancy SRIVC time delay estimation

To construct an *M-safely redundant* optimization problem for the estimation of the time delay, we need a cost function J_0 and multiple solver paths that satisfy the conditions in Definition 1. Here, the solver paths are generated by filtering the time delay estimation error using a set of low-pass filters having different cut-off frequencies. The cost function, J_0, is formulated from these filtered cost functions such that there always exists a solver path that corresponds to a decrease in J_0 after a finite number of iterations. Next we analyse the set of filters and cost function, J_0, required to satisfy the conditions in Definition 1.

4.1. Multiple solver paths for useful redundancy

Consider the system described by (1), for any $G(p, \theta)$ and τ, the estimation error $\epsilon(t, \theta, \tau)$ can be computed as, $\epsilon(t, \theta, \tau) = y(t) - G(p, \theta)u(t-\tau)$. The delay can be estimated by minimizing the cost function $J(\theta, \tau)$, where $J(\theta, \tau) = \int_{-\infty}^{\infty} \epsilon(t, \theta, \tau)^2\, \mathrm{d}t$. If $J(\theta, \tau)$ is filtered by the low-pass filter $L(p)$, then an estimate of θ and τ can be computed by minimizing the cost function $\bar{J}(\theta, \tau)$,

$$\bar{J}(\theta, \tau) = \int_{-\infty}^{\infty} \Big\{ L(p)[y(t) - G(p, \theta)u(t-\tau)] \Big\}^2 \,\mathrm{d}t$$

which by Parseval's theorem is equivalent to,

$$\bar{J}(\theta, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big[|G_0(j\omega)e^{-j\omega\tau_0} - G(j\omega, \theta)e^{-j\omega\tau}|^2 \times \Psi_u(\omega) + \Psi_v(\omega) \Big] \big|L(j\omega)\big|^2 \,\mathrm{d}\omega,$$

with $\Psi_u(\omega)$ the spectral density of $u(t)$ and the spectral density $\Psi_v(\omega) = \lambda$ follows from Assumption 2.4.

4.1.1. Known system parameters

When $G_0(p)$ is known and the input signal is white noise, i.e. $\Psi_u(\omega) = 1$; and $\delta\tau = \tau - \tau_0$, the cost function $\bar{J}(\theta, \tau)$ can be rewritten as,

$$\bar{J}(\delta\tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \Big[(1 - \cos(\omega\delta\tau)) \big|G_0(j\omega)\big|^2 + \frac{\lambda}{2} \Big] \big|L(j\omega)\big|^2 \,\mathrm{d}\omega. \quad (4)$$

We next provide necessary theoretical results and suggest a choice of filter set that can generate the multiple solver paths to be used in the useful redundancy framework.

Theorem 4.1. *Consider the system $G_0(p)$ as described in (1). When $\lambda = 0$, $\forall \delta\tau \in \mathbb{R}$, $\forall$ low-pass filters $L(p) \neq 0$, such that $\bar{J}(\delta\tau) \geq \bar{J}(0)$, the equality $\bar{J}(\delta\tau) = \bar{J}(0)$ occurs if and only if $\delta\tau = 0$.*

Proof. The proof is provided in Appendix A.1. ■

Theorem 4.2. *There exists a filter set $L_k(p)$, $k = \overline{1, n_f}$ chosen such that, $\forall\ \beta \geq \frac{5/4 - 1/(5\pi^2)}{3/4 - 1/(3\pi^2)}$, where $L_k(p)G_0(p)$ are ideal low-pass filters chosen with linearly spaced periods, $T_{c,k}$, such that the cut-off frequencies, $\omega_{c,k}$, span from $1/\beta$ to 1 of the system bandwidth, bw, i.e.,*

$$T_{c,k} = \beta T_{bw} - (k-1)\frac{(\beta - 1)T_{bw}}{n_f - 1}, \quad k = \overline{1, n_f}, \quad (5)$$

where $T_{bw} = 2\pi/bw$ and $\omega_{c,k} = 2\pi/T_{c,k}$. Then,

$$\forall \delta\tau_0 \neq 0, \exists L_q(p) : |\xi(L_q(p), \delta\tau_0)| < |\delta\tau_0|, \quad (6)$$

where $\xi(L_q(p), \delta\tau_0)$ is the corresponding minimum found using the filtered time delay cost function generated by $L_q(p)$ with the initial delay $\delta\tau_0$.

Proof. The proof is provided in Appendix A.2. ■

From Theorem 4.2, we see that with the choice of,

$$\beta \geq \frac{5/4 - 1/(5\pi^2)}{3/4 - 1/(3\pi^2)}, \quad (7)$$

there always exists an n_f, that can be computed by (A.8) in Appendix A.2, that ensures the filter set defined in Theorem 4.2 satisfies the necessary part of Definition 1, Condition 2. Note that (A.8) provides a loose lower bound. It is possible to have a smaller value of n_f and still obtain a filter set that satisfies the necessary condition.

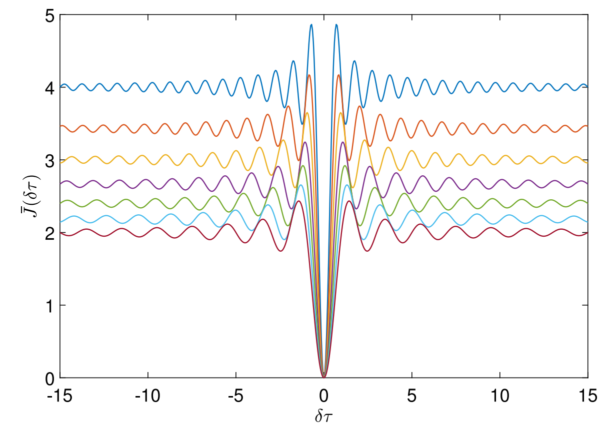


Fig. 1. Plots of $\bar{J}(\delta\tau)$ set of filters with linearly spaced periods.

An empirical method can also be utilized to determine the number of filters. For example, to check if a set of n_f filters with the ratio β satisfies the necessary condition, we can compute the n_f cost functions that follow from (A.2) in Appendix A.2 and check the minima locations using (A.4) to see that for any initial delay, τ_0, there always exists a cost function where the corresponding minima τ_i is found closer to the global minimum with respect to τ_0. Fig. 1 shows a set of six cost functions with the ratio $\beta = 2$ for a system bandwidth of 2π rad/s, i.e. the cut-off frequencies span from 1/2 to 1 (Hz). From Fig. 1, we can see that for any initial value of the time delay, there exists a cost function where a minima is found closer to the global minimum with respect to the initial delay. This can be confirmed by computing the minima of all the cost functions using (A.4) and observing the path to the global minimum.

4.1.2. Unknown system parameters

Considering the case where $G_0(p)$ possesses no resonant peaks, then it does not matter if $G_0(p)$ is known or unknown, as any ideal low pass filter $L_k(p)$ with bandwidth smaller than the bandwidth of $G_0(p)$ will allow $L_k(p)G_0(p)$ to approximate the desired low pass filter behaviour. If $G_0(p)$ has resonant peaks and the bandwidth of $L_k(p)$ is chosen significantly smaller than that of $G_0(p)$ then $L_k(p)G_0(p)$ will approximate the desired low pass filter behaviour.

4.2. Cost function of the time delay redundancy algorithm

Consider we have a value for the delay, τ, then estimating the rational component of (1), we have the following relationship in the Laplace domain,

$$\hat{G}_r(s) = G_0(s)e^{-(\tau_0 - \tau)s}, \quad (8)$$

where $\hat{G}_r(s)$ is a rational estimate only for the system. It is obvious that when $\tau \neq \tau_0$, $\hat{G}_r(s) \neq G_0(s)$. Hence there always exists model error in a rational estimate of the system if $\tau \neq \tau_0$.

We quantify the model error by using a model reduction technique based on an identification method (Gu, 2011) to find a rational transfer function $\hat{G}_r(s)$. The idea is to generate a noise-free data set from the system $G_0(s)e^{-(\tau_0-\tau)s}$, then use the SRIVC method with a reliable initialization to obtain a rational system $\hat{G}_r(s)$ that has the same model order as the true system and satisfies (8). For

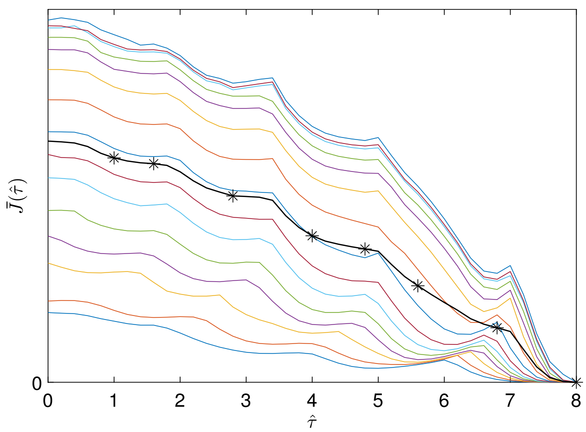


Fig. 2. Plot of $\bar{J}(\tau)$ for filters set $L_k(p)$ defined as in Theorem 4.2.

example, Fig. 2 shows the graphs of 15 filtered time delay cost functions after including the model error for the system,

$$G(s) = \frac{(-6400s + 1600)e^{-8s}}{s^4 + 5s^3 + 408s^2 + 416s + 1600}.$$

We now define the cost function J_0 as,

$$J_0(\tau) = \frac{1}{n_f} \sum_{k=1}^{n_f} \bar{\bar{J}}_k(\tau),\ \bar{\bar{J}}_k(\tau) = 100 \frac{\|y(t) - \hat{\hat{x}}_k(t)\|_2}{\|y(t) - \mathbb{E}\{y(t)\}\|_2} \tag{9}$$

where $\hat{\hat{x}}_k(t)$ is the estimated system output when the filter $L_k(s)$ is used, $\bar{\bar{J}}_k(\tau)$ is the normalized version of the filtered cost function $\bar{J}_k(\tau)$, and $\bar{J}_k(\tau)$ is the cost function obtained when filtering $J(\tau)$ with the filter $L_k(s)$. Here, the averaging is utilized to obtain a smooth decreasing cost function J_0.

It can be seen in Fig. 2 there exists a minimum path through the filtered cost function that corresponds to a decrease in J_0. Therefore, the *M-safely redundant* optimization problem for the time delay estimation is defined as,

$$\min_{\tau} J_0(\tau),$$

with $J_0(\tau)$ defined as in (9). The filter set consists of n_f Butterworth filters with cut-off frequencies, chosen such that the corresponding periods are linearly spaced, and span from $1/\beta$ to 1 of the system bandwidth. In the case the system parameters are known, n_f can be chosen using the empirical methodology discussed in Section 4.1.1 and $\beta \geq \frac{5/4 - 1/(5\pi^2)}{3/4 - 1/(3\pi^2)}$. In a general case where the system parameters are unknown, our suggestion is to choose $\beta = 15$ and n_f to be two or three times the values obtained using the empirical method described in Section 4.1.1.

4.3. Algorithm

The proposed algorithm utilizing useful redundancy with multiple filters and SRIVC (Garnier & Wang, 2008; Garnier & Young, 2013) is described in Algorithm I.

Remark 4.1. Note that Algorithm I is slightly different to the useful redundancy technique described in Alamir et al. (2009). In Alamir et al. (2009), at the ith iteration, the algorithm searches along the kth solver path to find a minimum to reduce the cost function J_0. If it cannot find one, it switches to the $(k+1)$th solver path. In this paper, at the ith iteration, the algorithm searches all solver paths to find all the minima. It then chooses the minima that gives the smallest value of the cost function J_0.

ALGORITHM I

Step 1. Initialization

(1) Select a set of low-pass filters as suggested in Section 4.1.

(2) Set the boundaries τ_{min}, τ_{max}, $\Delta\tau_{min}$, $\Delta\tau_{max}$[a], and the SVF[b] cut-off frequency ω_c^{SVF}.

(3) Choose an initial value $\hat{\tau}_0$ of the time delay.

Step 2. Iterative estimation

for i=1:converge

(1) *for k=1:n_f*

  (a) Choose the low-pass filter $L_k(p)$.

  (b) Set the initial delay $\hat{\tau}_{i,k}^0$ as $\hat{\tau}_{i-1}$

  (c) Use the SRIVC algorithm and Gauss-Newton method as described in (Chen et al. 2015) (without stage 2 of Step 2) to estimate the time delay $\hat{\tau}_{i,k}$[c].

  (d) Compute $J_0(\hat{\tau}_{i,k})$ (follows (9)).

  end

(2) Choose $\hat{\tau}_i = \text{argmin}_{\hat{\tau}_{i,k}} J_0(\hat{\tau}_{i,k})$.

(3) If $|(\hat{\tau}_i - \hat{\tau}_{i-1})/\hat{\tau}_i| \geq \epsilon$, go to Step 1, else break. [d]

end

Step 3. Refine parameter estimation

Repeat Step 2 with only one filter, $L(p) = 1$.

[a] Note that $\Delta\tau_{max}$ is used to constrain the increment $\Delta\hat{\tau}$ in the time delay estimation, i.e. when $\Delta\hat{\tau}^j > \Delta\tau_{max}$, set $\Delta\hat{\tau}^j = \Delta\tau_{max}$. This ensures the Wolfe conditions are satisfied at each iteration in the Gauss-Newton technique (Wolfe, 1969).

[b] SVF is called the State Variable Filter, which is computed as $F(p) = 1/(p + \omega_c^{SVF})^n$. It is a common method in CT system identification to find the initial derivatives of the input and output signals (Garnier & Wang, 2008).

[c] Note that here, we constraint the increment $\Delta\hat{\tau}$ by $\Delta\tau_{max}$

[d] ϵ is a small value selected to obtain the desired accuracy.

5. Numerical examples

In this section, we compare the performance of our proposed method to the one-time filtering method in Chen et al. (2015) and the GSEPNIV method in Yang et al. (2007) using a system based on the Rao–Garnier continuous time benchmark (Garnier & Wang, 2008),

$$G(s) = \frac{(-6400s + 1600)e^{-8s}}{s^4 + 5s^3 + 408s^2 + 416s + 1600}.$$

The experiment is conducted using both regular and irregular sampling schemes,

- For the regular sampling case, the input excitation signal is a PRBS of maximal-length with the number of stages in the shift register being 10, the sample time is 10 ms, and the number of samples, N = 8000.
- For the irregular sampling, the input is also a PRBS of maximal-length with the number of stages in the shift register being 10 and the clock period is 0.5 s. The input and output data are sampled at an irregular time instant t_k, where the sampling interval h_k is uniformly distributed as, $h_k \sim U[0.01, 0.05]$ (s). The number of samples, $N = 4500$.

The additive output disturbance in both schemes is Gaussian distributed white noise with zero mean designed to give a SNR of 5 dB and 15 dB.

Table 1
Global convergence of the time delay estimate as a percentage of 100 Monte Carlo simulations.

Sampling scheme		One-time filtering				GSEPNIV				Proposed method			
	Initial delay	0 s	3 s	7 s	9 s	0 s	3 s	7 s	9 s	0 s	3 s	7 s	9 s
Regular sampled data	SNR = 5 dB	1%	22%	76%	68%	18%	26%	28%	6%	99%	99%	99%	99%
	SNR = 15 dB	0%	6%	70%	69%	21%	24%	31%	12%	100%	100%	100%	100%
Irregular sampled data	SNR = 5 dB	8%	23%	66%	48%	N/A	N/A	N/A	N/A	100%	100%	100%	100%
	SNR = 15 dB	9%	19%	55%	50%	N/A	N/A	N/A	N/A	100%	100%	100%	100%

Table 2
Global convergence of the time delay estimate as a percentage of 100 Monte Carlo simulations with a randomly selected initial delay.

Sampling scheme	SNR	One-time filtering	GSEPNIV	Proposed method
Regular sampled data	5 dB	52%	12%	99%
	15 dB	46%	29%	100%
Irregular sampled data	5 dB	48%	N/A	100%
	15 dB	21%	N/A	100%

Table 3
Parameter estimates from regularly sampled data of the three algorithms (SNR=15 dB) with a random selected initial delay. The statistics are computed only when the time delay estimate converges to the global minimum. Note that as shown in Table 2, the one-time filtering and the GSEPNIV methods converge less than 46% to the global minimum.

Estimated parameters	One-time filtering	GSEPNIV	Proposed method
$\hat{b}_0$	-6405.851 ± 58.828	-6076.464 ± 1811.165	-6405.562 ± 57.132
$\hat{b}_1$	1599.580 ± 44.533	2090.660 ± 302.586	1599.351 ± 44.486
$\hat{a}_1$	4.982 ± 0.112	3.457 ± 0.926	4.990 ± 0.104
$\hat{a}_2$	408.225 ± 3.040	387.634 ± 102.936	408.120 ± 2.839
$\hat{a}_3$	416.204 ± 4.065	416.206 ± 97.081	416.282 ± 3.942
$\hat{a}_4$	1601.066 ± 12.913	1523.544 ± 413.095	1600.836 ± 12.841

Table 4
Computation time of the three algorithms (in seconds) when the initial time delay being 0 and the sampling scheme being regular. Note that as shown in Table 1, the one-time filtering and the GSEPNIV methods converge less than 21% to the global minimum.

One-time filtering	GSEPNIV	Proposed method
3.7	2.1	32.4

The cut-off frequency ω_c^{SVF} of the State Variable Filter used in the SRIVC algorithm is chosen as 25 rad/s, which is approximately the system bandwidth (Garnier & Wang, 2008). The initial values for the time delay, $\hat{\tau}^0$, are selected as 0, 3, 7, 9 as well as a random value from the uniform distribution $U[0, 9](s)$. The delay boundaries τ_{min}, τ_{max} are set to 0 and 10 (s) respectively. The maximum step $\Delta\tau_{max}$ is 1. For the one-time filtering method suggested in Chen et al. (2015), the cut off frequency of the filter is chosen as 1/10 of the system bandwidth (Chen et al., 2015). For the proposed method, the cut off frequencies of the filters, chosen based on linearly spaced periods, span from 1/15 to 1 of the system bandwidth and the number of filters used is 15. The order of the Butterworth filters is set to 10. The maximum number of iterations for the Gauss–Newton method (Step 2 in Algorithm I) is 20 and the threshold ϵ for convergence is 10^{-3}.

For the GSEPNIV method (Yang et al., 2007), the degree of smoothing, $\beta = 1000$, the prefilter passband parameter, $\alpha = 0.5$, and the maximum number of iterations is 100. The GSEPNIV method is initialized with the GSEPNLS method described in the same paper (Yang et al., 2007).

For comparison, 100 different data sets are generated for each noise level of each sampling scheme. We consider the estimated time delay to be the global minimum when the relative error, $\epsilon_r = |\hat{\tau} - \tau_0|/\tau_0$, is less than 0.01, where $\hat{\tau}$ and τ_0 are the estimated delay and the true system delay respectively.

A graph showing the trajectory of a time delay estimate is plotted in Fig. 3 for a single data set. We denote '*' by the local minima obtained after each round of searching using n_f cost functions.

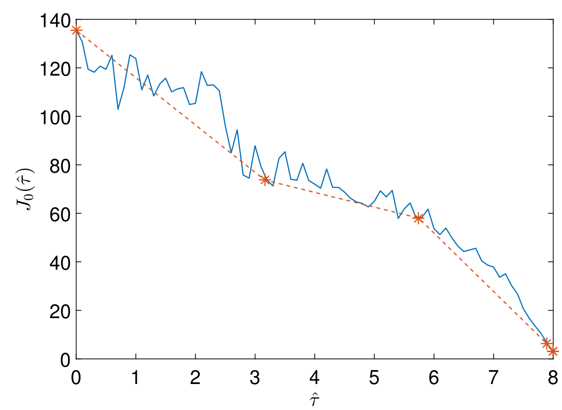


Fig. 3. The original cost function and time delay estimate trajectory for a single data set.

Note that, if only one filter were to be used, there is a high chance of convergence to a local minima. However, by using the useful redundancy technique, the proposed algorithm can avoid the local minima and converge to the global minimum.

The numerical results of the experiments are provided in Tables 1–3. We can see that for both sampling schemes, the proposed method based on the useful redundancy technique performs much better as compared to the other two methods. For all the initial values of delay used in this experiment, the one-time filtering method (Chen et al., 2015) never convergences 100% to the global minimum. However the proposed method utilizing useful redundancy still achieves a very high global convergence percentage, i.e. mostly 100% with any initial delay for both SNRs and sampling schemes. When the delay is estimated correctly, the accuracy of the estimated system parameters are also high as shown in Table 3.

In addition, the computation time of the three algorithms is shown in Table 4. This is the average computation time to estimate the time delay and parameters of the Rao–Garnier system using regular sampled data with initial time delay being 0. The proposed algorithm has a higher computation time as compared to the other two methods, this is the trade-off to achieve a better accuracy of the delay estimates. Note again that the proposed algorithm is the only algorithm that had nearly 100% global convergence, whereas the other algorithms at best achieved less than 21% convergence.

6. Conclusion

The paper presents a new algorithm to estimate the parameters and time delay of a continuous-time system from regularly and irregularly sampled data. The idea is based on Instrumental Variable methods and employing the useful redundancy technique to enhance the global convergence by generating multiple cost functions based on filtering the data several times. The paper also develops some theoretical results related to the minima locations of the filtered delay cost function and the choice of filters to ensure the algorithm can converge to the global minimum for the case when the system parameters are known. Numerical results show a significant improvement in the global convergence rate of the time delay estimation as compared to existing methods irrespective of the SNR.

Appendix A

A.1. Proof of Theorem 4.1

Proof. From (4), when $\lambda = 0, \bar{J}(\delta\tau)$ becomes,

$$\bar{J}(\delta\tau) = \frac{1}{\pi}\int_{-\infty}^{\infty}(1-\cos(\omega\delta\tau))\left|G_0(j\omega)L(j\omega)\right|^2 \mathrm{d}\omega. \tag{A.1}$$

$\forall\delta\tau \in \mathbb{R},\ \omega \geq 0$, we have,

$$1-\cos(\omega\delta\tau) \geq 0,\ \ \forall\delta\tau,$$
$$\left|G_0(j\omega)L(j\omega)\right|^2 \geq 0,\ \ \forall\omega.$$

Therefore, $\bar{J}(\delta\tau) \geq 0,\ \forall\delta\tau \in \mathbb{R},\ \omega \geq 0$. Note that $\bar{J}(0) = 0$, hence we have $\bar{J}(\delta\tau) \geq \bar{J}(0),\ \forall\delta\tau \in \mathbb{R},\ \omega \geq 0$. The equality occurs when $L(j\omega) = 0,\ \forall\omega$ or $G_0(j\omega) = 0,\ \forall\omega$ or $\cos(\omega\delta\tau) = 1,\ \forall\omega$. As the system $G_0(p)$ and the filter $L(p) \neq 0$ (condition of Theorem 4.1), the equality only occurs when $\cos(\omega\delta\tau) = 1,\ \forall\omega$, hence $\delta\tau = 0$. ■

A.2. Proof of Theorem 4.2

Proof. If $L(s)$ is selected such that $L(s)G_0(s)$ is an ideal low-pass filter with cut-off frequency ω_c, then,

$$\begin{aligned}\bar{J}(\delta\tau) &= \frac{1}{2\pi}\int_{-\omega_c}^{\omega_c}|1-e^{-j\omega\delta\tau}|^2\mathrm{d}\omega + C\\ &= \frac{2\omega_c}{\pi} - \frac{2\sin(\omega_c\delta\tau)}{\pi\delta\tau} + C,\end{aligned} \tag{A.2}$$

where $C = \frac{1}{2\pi}\int_{-\infty}^{\infty}\lambda\left|L(j\omega)\right|^2\mathrm{d}\omega$.

The minima and maxima of $\bar{J}(\delta\tau)$ occur at the roots of,

$$\frac{\mathrm{d}\bar{J}(\delta\tau)}{\mathrm{d}\delta\tau} = -\frac{2}{\pi}\frac{\omega_c\cos(\omega_c\delta\tau)\delta\tau - \sin(\omega_c\delta\tau)}{\delta\tau^2}. \tag{A.3}$$

From (A.3), we can see that the extrema of $\bar{J}(\delta\tau)$ are also the extrema of the function $sinc(\omega_c\delta\tau)$. And for the function $sinc(\omega_c\delta\tau)$, the locations of the ith ($i \geq 1$) positive extremum $\tilde{\delta\tau}$ can be approximated by,

$$\tilde{\delta\tau}_i \simeq (i+\frac{1}{2})\frac{\pi}{\omega_c} - \frac{1}{(i+\frac{1}{2})\pi\omega_c} \simeq \frac{2i+1}{4}T_c - \frac{T_c}{(2i+1)\pi^2}. \tag{A.4}$$

where T_c is the corresponding period of the cut-off frequency ω_c. The minima occurs when i is odd and the maxima occurs when i is even. Therefore, for $\bar{J}(\delta\tau)$, the positive extrema can also be approximated using (A.4).

Now we prove the theorem by using (A.4). Since the filtered time delay cost function $\bar{J}(\delta\tau)$ is an even function, we only need to prove the theorem for $\delta\tau_0 > 0$.

Denote $\tilde{\delta\tau}_{i,k}^{min}$ as the ith positive minimum of the filtered time delay cost function $\bar{J}(\delta\tau)$ generated by the filter $L_k(p)$; $\tilde{\delta\tau}_{i,k}^{max}$ as the ith positive maximum of the filtered time delay cost function $\bar{J}(\delta\tau)$ generated by the filter $L_k(p)$. Using (A.4),

$$\begin{aligned}\tilde{\delta\tau}_{i,k+1}^{min} &= (\frac{4i+1}{4} - \frac{1}{(4i+1)\pi^2})\Big(\beta - k\frac{\beta-1}{n_f-1}\Big)T_{bw}\\ \tilde{\delta\tau}_{i,k+1}^{max} &= (\frac{4i-1}{4} - \frac{1}{(4i-1)\pi^2})\Big(\beta - k\frac{\beta-1}{n_f-1}\Big)T_{bw}.\end{aligned} \tag{A.5}$$

From (A.5), for any filter $L_k(p)$, we have the following,

$$\tilde{\delta\tau}_{i,k}^{max} < \tilde{\delta\tau}_{i,k}^{min} < \tilde{\delta\tau}_{i+1,k}^{max},\ \ \forall i \in \mathbb{Z}^+. \tag{A.6}$$

Consider now the two following cases:

Case 1: $\delta\tau_0 \leq \tilde{\delta\tau}_{1,1}^{max}$, then from (A.6), we have, $\xi(L_1(p), \delta\tau_0) = 0$, which is smaller than $\delta\tau_0$ as $\delta\tau_0 \neq 0$. Therefore, $\xi(L_1(p), \delta\tau_0) < \delta\tau_0$. ■

Case 2: $\delta\tau_0 > \tilde{\delta\tau}_{1,1}^{max}$. Denote i_0 such that,

$$\tilde{\delta\tau}_{i_0-1,1}^{max} < \delta\tau_0 \leq \tilde{\delta\tau}_{i_0,1}^{max},\quad i_0 \in \mathbb{N}, i_0 \geq 2. \tag{A.7}$$

Now consider a filter set $L_k(p)$, $k = \overline{1, n_f}$ defined as in Theorem 4.2, with,

$$n_f \geq (1/M + \beta - 2)/(1/M - 1),\quad n_f \in \mathbb{N}, \tag{A.8}$$

where,

$$M = \frac{4i_0 - 3 - 4/((4i_0-3)\pi^2)}{4i_0 - 1 - 4/((4i_0-1)\pi^2)}. \tag{A.9}$$

Next, we will show,

$$\tilde{\delta\tau}_{i+1,k+1}^{max} \geq \tilde{\delta\tau}_{i,k}^{min},\quad 1 \leq i \leq i_0 - 1,\ 1 \leq k \leq n_f - 1,$$

and,

$$\tilde{\delta\tau}_{i,1}^{max} \geq \tilde{\delta\tau}_{i,n_f}^{min},\ i = \overline{1, i_0}.$$

Then, we will show this filter satisfies the condition in (6). First, from (A.8), we have,

$$M \leq 1 - \frac{(\beta-1)/(n_f-1)}{\beta - (n_f-2)(\beta-1)/(n_f-1)}. \tag{A.10}$$

Note that, as $\beta > 1$, then $\forall k = \overline{1, n_f - 1}$, we have,

$$1 - \frac{(\beta-1)/(n_f-1)}{\beta - (k-1)(\beta-1)/(n_f-1)} \geq M. \tag{A.11}$$

Hence, combining (A.9) and (A.11),

$$\begin{aligned}&\frac{\beta - k(\beta-1)/(n_f-1)}{\beta - (k-1)(\beta-1)/(n_f-1)}\\ &\quad\geq \frac{4i_0 - 3 - 4/((4i_0-3)\pi^2)}{4i_0 - 1 - 4/((4i_0-1)\pi^2)},\end{aligned} \tag{A.12}$$

where $k = \overline{1, n_f - 1}$.

Now, consider the function, $f(x) = \frac{4x+1-4/((4x+1)\pi^2)}{4x+3-4/((4x+3)\pi^2)}$. It is easy to show that $f(x)$ is an increasing function for $x \geq 1$, so we have,

$f(i_0-1) \geq f(i),\ i=\overline{1, i_0-1}$, or,

$$\frac{4i_0-3-4/((4i_0-3)\pi^2)}{4i_0-1-4/((4i_0-1)\pi^2)} \geq \frac{4i+1-4/((4i+1)\pi^2)}{4i+3-4/((4i+3)\pi^2)}, \quad i=\overline{1, i_0-1}. \tag{A.13}$$

From (A.12) and (A.13),

$$\Big(\frac{4i+3}{4}-\frac{1}{(4i+3)\pi^2}\Big)\Big(\beta-k\frac{\beta-1}{n_f-1}\Big) \geq \Big(\frac{4i+1}{4}-\frac{1}{(4i+1)\pi^2}\Big)\Big(\beta-(k-1)\frac{\beta-1}{n_f-1}\Big), \tag{A.14}$$

for $i=\overline{1, i_0-1}$, $k=\overline{1, n_f-1}$.

From (A.14) and (A.5), we conclude,

$$\tilde{\delta\tau}^{max}_{i+1,k+1} \geq \tilde{\delta\tau}^{min}_{i,k}, \quad i=\overline{1, i_0-1},\ k=\overline{1, n_f-1}. \tag{A.15}$$

Similar to the previous analysis, it is easy to show that,

$$\tilde{\delta\tau}^{max}_{i,1} \geq \tilde{\delta\tau}^{min}_{i,n_f}, \quad i=\overline{1, i_0}. \tag{A.16}$$

Lastly, we show that for any filter set $L_k(p)(k=\overline{1, n_f})$ defined as in Theorem 4.2 with $n_f \geq (1/M+\beta-2)/(1/M-1)$, $n_f \in \mathbb{N}$,

$\forall \delta\tau_0 \neq 0, \exists L_q(p) : \xi(L_q(p), \delta\tau_0) < \delta\tau_0$.

Case 2-A:

$$\tilde{\delta\tau}^{min}_{i_0-1,1} < \delta\tau_0 \leq \tilde{\delta\tau}^{max}_{i_0,1} \tag{A.17}$$

From (A.6) and (A.17), it is obvious that $\xi(L_1(p), \delta\tau_0) = \tilde{\delta\tau}^{min}_{i_0-1,1}$, and $\tilde{\delta\tau}^{min}_{i_0-1,1} < \delta\tau_0$, hence $\xi(L_1(p), \delta\tau_0) < \delta\tau_0$. ■

Case 2-B:

$$\tilde{\delta\tau}^{max}_{i_0-1,1} < \delta\tau_0 \leq \tilde{\delta\tau}^{min}_{i_0-1,1}. \tag{A.18}$$

From (A.5) we see that, $\forall i \in \mathbb{Z}^+$,

$$\tilde{\delta\tau}^{min}_{i,n_f} < \tilde{\delta\tau}^{min}_{i,n_f-1} < \cdots < \tilde{\delta\tau}^{min}_{i,2} < \tilde{\delta\tau}^{min}_{i,1}, \tag{A.19}$$

hence,

$$\tilde{\delta\tau}^{min}_{i_0-1,n_f} < \tilde{\delta\tau}^{min}_{i_0-1,n_f-1} < \cdots < \tilde{\delta\tau}^{min}_{i_0-1,1}. \tag{A.20}$$

Combining (A.16) and (A.18), we have,

$$\tilde{\delta\tau}^{min}_{i_0-1,n_f} < \delta\tau_0 < \tilde{\delta\tau}^{min}_{i_0-1,1}. \tag{A.21}$$

Considering (A.20) and (A.21), there always exists a value $q \geq 2$ such that $\tilde{\delta\tau}^{min}_{i_0-1,q} < \delta\tau_0 < \tilde{\delta\tau}^{min}_{i_0-1,q-1}$. Following from (A.15), we have $\tilde{\delta\tau}^{min}_{i_0-1,q-1} \leq \tilde{\delta\tau}^{max}_{i_0,q}$. Therefore,

$$\tilde{\delta\tau}^{min}_{i_0-1,q} < \delta\tau_0 < \tilde{\delta\tau}^{max}_{i_0,q}. \tag{A.22}$$

Combining (A.6) and (A.22),

$$\xi(L_q(p), \delta\tau_0) = \tilde{\delta\tau}^{min}_{i_0-1,q}, \tag{A.23}$$

therefore, $\xi(L_q(p), \delta\tau_0) < \delta\tau_0$. ■

References

Agarwal, M., & Canudas, C. (1987). On-line estimation of time-delay and continuous-time process parameters. *Internationl Journal of Control*, *46*(1), 295–311.

Ahmed, S., Huang, B., & Shah, S. L. (2006). Parameter and delay estimation of continuous-time models using a linear filter. *Journal of Process Control*, *16*(4), 323–331.

Alamir, M. (2008). On useful redundancy in dynamic inverse problems related optimization. In *Proceedings of the 17th IFAC world congress, Vol. 41* (2) (pp. 976–981).

Alamir, M., Welsh, J. S., & Goodwin, G. C. (2009). Redundancy versus multiple starting points in nonlinear system related inverse problems. *Automatica*, *45*(4), 1052–1057.

Åström, K. J., & Bernhardsson, B. (2003). Systems with lebesgue sampling. In A. Rantzer, & C. I. Byrnes (Eds.), *Directions in mathematical systems theory and optimization* (pp. 1–13).

Baysse, A., Carrillo, F.J., & Habbadi, A. (2011). Time domain identification of continuous-time systems with time delay using output error method from sampled data. In *Proceeding of the 18th IFAC world congress, Vol. 44* (1) (pp. 9041–9046).

Baysse, A., Carrillo, F.J., & Habbadi, A. (2012). Least squares and output error identification algorithms for continuous time systems with unknown time delay operating in open or closed loop. In *Proceeding of the 16th IFAC symposium on system identification, Vol. 45* (16) (pp. 149–154).

Björklund, S. (2003). *A survey and comparison of time delay estimation methods in linear systems* (Ph.D. thesis), Lund Institute of Technology.

Björklund, M., Nihtilä, M., & Söderström, T. (1991). An algorithm for identification of linear systems with varying time-delay. In *Proceedings of the 9th IFAC symposium on identification and system parameter estimation* (pp. 1254–1259).

Carlemalm, C., Halvarsson, S., Wigren, T., & Wahlberg, B. (1999). Algorithms for time delay estimation using a low complexity exhaustive search. *IEEE Transactions on Automatic Control*, *44*(5), 1031–1037.

Carter, G. C. (1987). Coherence and time delay estimation. *Proceedings of the IEEE*, *75*(2), 236–255.

Chen, F., Garnier, H., & Gilson, M. (2015). Robust identification of continuous-time models with arbitrary time-delay from irregularly sampled data. *Journal of Process Control*, *25*, 19–27.

Eckhard, D., Bazanella, A. S., Rojas, C. R., & Hjalmarsson, H. (2017). Cost function shaping of the output error criterion. *Automatica*, *76*, 53–60.

Ferretti, G., Maffezzoni, C., & Scattolini, R. (1996). On the identifiability of the time delay with least-squares methods. *Automatica*, *32*(3), 449–453.

Garnier, H. (2015). Direct continuous-time approaches to system identification. Overview and benefits for practical applications. *European Journal of Control*, *24*, 50–62.

Garnier, H., & Wang, L. (Eds.). (2008). *Identification of continuous-time models from sampled Data*. London: Springer-Verlag.

Garnier, H., & Young, P. C. (2013). The advantages of directly identifying continuous-time transfer function models in practical applications. *International Journal of Control*, *87*(7), 1319–1338.

Gawthrop, P. J., & Nihtil, M. T. (1985). Identification of time delays using a polynomial identification method. *Systems & Control Letters*, *5*(4), 267–271.

Gu, C. (2011). *Model order reduction of nonlinear dynamical systems* (Ph.D. thesis), Berkeley: University of California.

Ha, H., & Welsh, J. S. (2016). Parameter and delay estimation of continuous-time models utilizing multiple filtering. In *2016 IEEE 55th conference on decision and control* (pp. 1205–1210).

Kaminskas, V. (1979). Parameter estimation in systems with time delay and closed loop systems. In *Proceedings of the 5th IFAC symposium on identification and system parameter estimation* (pp. 669–677).

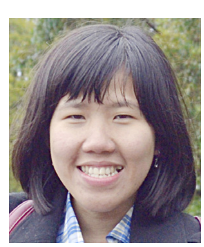
Pupeikis, R. (1989). Recursive estimation of the parameters of linear systems with time delay. In *Proceedings of the 7th IFAC symposium on identification and system parameter estimation* (pp. 787–792).

Unbehauen, H., & Rao, G. P. (1990). Continuous-time approaches to system identification-A survey. *Automatica*, *26*(1), 23–35.

Wolfe, P. (1969). Convergence conditions for ascent methods. *SIAM Review*, *11*(2), 226–235.

Yang, Z., Iemura, H., Kanae, S., & Wada, K. (2007). Identification of continuous-time systems with multiple unknown time delays by global nonlinear least-squares and instrumental variable methods. *Automatica*, *43*(7), 1257–1264.

Young, P. C. (1981). Parameter estimation for continuous-time models - A survey. *Automatica*, *17*(1), 23–39.



Huong Ha was born in 1988 in Vietnam. She received her B.S. degree in Electrical Engineering (cum claude) from Portland State University, USA, in 2011 under Intel Vietnam Study Abroad scholarship. From 2011 to 2013 she was a quality and reliability engineer at Intel Products Vietnam before moving to Australia to work on her Ph.D. degree. After getting her Ph.D. degree in Electrical Engineering from The University of Newcastle, Australia in 2017, Huong works at Freelancer International Pty Limited, Sydney, Australia as a data scientist. Huong's research interests include system identification, machine learning and software analytics.

James S. Welsh was born in Maitland, Australia in 1965. He received a B.E. degree (Hons. I) in electrical engineering in 1997 and his Ph.D. in 2004, both from the University of Newcastle, Australia. His Ph.D. thesis analysed ill-conditioning problems arising in system identification. He spent more than 15 years working in industry before taking up his current position with the University of Newcastle, where, since 2004 he has been employed in the School of Electrical Engineering and Computing. His research interests include system identification, auto-tuning and process control. Most recently he has been actively involved in research projects in the areas of biomedical science, transformer monitoring and medical engineering. He is a member of the Priority Research Centre for Complex Dynamic Systems and Control at the University of Newcastle.

Mazen Alamir graduated in Mechanics (Grenoble, 1990) and Aeronautics (Toulouse, 1992). He received his Ph.D. in Nonlinear Model Predictive Control in 1995. Since 1996, he has been since a CNRS research associate in the Control Systems Department of Gipsa-lab, Grenoble. His main research topics are model predictive control, receding horizon observers, nonlinear hybrid systems, predictive maintenance via Machine Learning, optimal cancer treatment as well as industrial applications. He is member of the IFAC technical committee on Nonlinear Systems and served as head of the ''Nonlinear Systems and Complexity'' research group of the Control Systems Department of Gipsa-lab, Grenoble. Home page: www.mazenalamir.fr