Robust identification of continuous-time models with arbitrary time-delay from irregularly sampled data

Fengwei Chen [a,b], Hugues Garnier [a,b,*], Marion Gilson [a,b]

[a] *Université de Lorraine, Centre de Recherche en Automatique de Nancy (CRAN), UMR 7039, 2 rue Jean Lamour, F-54519 Vandœuvre-les-Nancy, France*
[b] *CNRS, CRAN, UMR 7039, France*

ARTICLE INFO

ABSTRACT

This paper presents a new approach to identify continuous-time systems with arbitrary time-delay from irregularly sampled input–output data. It is based on the separable nonlinear least-squares method which combines in a bootstrap manner the iterative optimal instrumental variable method for transfer function model estimation with an adaptive gradient-based technique that searches for the optimal time-delay. Since the objective function may have several local minima with respect to the unknown parameters (especially the time-delay), the initialization requires special attention. Here, a low-pass filtering strategy is used to widen the convergence region around the global minimum. Simulation results are included to show the performance of the proposed method.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Although dynamic systems in the physical world are naturally described in the continuous-time (CT) domain, most system identification schemes have been based on discrete-time (DT) models [1,2]. The last decade has however witnessed a resurgence of interest in direct CT modeling from sampled data. As discussed in [3,4], direct identification of CT models presents several advantages over traditional DT model identification. Amongst these, direct CT methods can handle in a quite straightforward way irregularly sampled data. The problem of system identification from irregularly sampled data is of importance as this case happens when the data are event-triggered sampled (see *e.g.* [5,3]) and occurs in several applications. The underlying DT model would become time-varying in the case of irregular sampling and cannot be easily handled by traditional DT methods. By contrast, this difficulty can be avoided by using the direct CT identification methods because the differential-equation model is valid whatever the time instants are

considered and, in particular, it does not assume regularly sampled data, as required in the case of the standard difference-equation model.

There are various time-domain approaches available for direct CT model identification. Amongst these, instrumental variable (IV) methods have attracted in the last decade a lot of attention (see *e.g.* [6,7], and the prior references therein). One of the most reliable IV-based estimators (see *e.g.* [4]) is known as the simplified refined instrumental variable method for continuous-time systems (SRIVC), which delivers optimal estimates when the output measurement noise is white. It was first developed by Young and Jakeman [8] and was later extended and evaluated in [9,10]. Since then, a number of papers have exploited this iterative optimal IV concept in various ways (see [11–17]). When the time-delay is *a priori* known, the SRIVC method has been shown to perform well in presence of irregularly sampled data [18].

Many practical systems, such as thermal and chemical processes, biological systems, have inherent time-delay. There has been a continuing interest in methods of identifying CT systems with time-delay. However, there is still no clear agreement on which method is best. A very nice survey of various methods is presented in [19]. A linear filter based method was recently introduced in [20], where the time-delay along with the transfer function model parameters are estimated simultaneously through a simple linear regression. In [21], iterative global non-linear least-squares and instrumental variable methods were proposed, in which the

plant parameters and time-delay are estimated in a separable way. In [22], a combined continuous wavelet transform (CWT) and cross correlation method is used to estimate the time-delay for MIMO dynamical systems, the unbiased time-delay was found out by calculating and handling the cross correlation between the CWT coefficients of the input and output signal. There are also many other methods dedicated to consider step response data, see *e.g.* [23–26].

This paper presents a new method to identify CT delay systems when the measured data are irregularly sampled. The proposed method is based on the separable nonlinear least-squares method and estimates in a combined way the plant parameters by the reliable SRIVC method and the time-delay by an adaptive gradient search. To robustify the time-delay search against its initialization, the ideas presented in [27] are adopted. A low-pass CT filtering technique is introduced to widen the convergence region for parameter estimation. Digital implementation issues of the CT filtering operations require special attention in this irregularly sampled data situation and appropriate methods are suggested. For a properly chosen filter, the proposed method converges to the global minimum for a larger range of initial values. Compared with a fully gradient-based method, the SRIVC-based method reduces the computational load, since the dimension of the unknown parameters estimated by the gradient search is minimized.

This paper is organized as follows. The identification problem is stated in Section 2. Then, independent estimation of the plant parameters and the time-delay is detailed in Sections 3 and 4, respectively. Subsequently, the bootstrap method for time-delay system identification is described in Section 5. After that, the low-pass filtering technique to enlarge the global convergence region is introduced in Section 6. Simulation examples are presented in Section 7 to illustrate the effectiveness of the proposed method. Finally, Section 8 presents the main conclusions.

2. Problem statement

Consider the following CT output error (COE)[1] system

$$\begin{cases} x(t) = G(p, \boldsymbol{\theta}^o)u(t - \tau^o) = \dfrac{B(p, \boldsymbol{\theta}^o)}{A(p, \boldsymbol{\theta}^o)}u(t - \tau^o) \\ y(t_k) = x(t_k) + e(t_k) \end{cases} \tag{1}$$

where $u(t)$, $x(t)$ and τ^o are the excitation signal, the noise-free response and the pure time-delay, respectively. p denotes the differentiation operator, *i.e.* $px = (dx/dt)$. $B(p, \boldsymbol{\theta}^o)$ and $A(p, \boldsymbol{\theta}^o)$ are polynomials assumed to be coprime which have the following structure

$$B(p, \boldsymbol{\theta}^o) = b_0^o p^m + b_1^o p^{m-1} + \cdots + b_m^o$$
$$A(p, \boldsymbol{\theta}^o) = a_0^o p^n + a_1^o p^{n-1} + \cdots + 1, \quad (n \geq m)$$

The true parameters are stacked in a column vector $\boldsymbol{\rho}^o$

$$\boldsymbol{\rho}^o = [\boldsymbol{\theta}^o; \ \tau^o] = [a_0^o \ \ldots \ a_{n-1}^o \ b_0^o \ \ldots \ b_m^o \ \tau^o]^T \in \mathbb{R}^{n+m+2} \tag{2}$$

$u(t)$ and $y(t)$ are assumed to be sampled at time-instants $\{t_k\}_{k=1}^N$, and this gives rise to $u(t_k)$ and $y(t_k)$. Since the noise-free signal is inevitably corrupted by a noise in practical situations, it is assumed that $x(t_k)$ is contaminated by a DT white noise $e(t_k)$, and this results in the measurement $y(t_k)$. The sampled signals $u(t_k)$ and $y(t_k)$ are

---

[1] Since the measurement noise is assumed to be a DT white noise, this model could be called as hybrid CT OE model. However, for the reason of simplicity, we use COE as an acronym.

assumed to be recorded at irregular time-instants, the time-varying sampling interval is denoted as

$$h_k = t_{k+1} - t_k, \quad k = 1, 2, \cdots, N - 1 \tag{3}$$

Thus, the identification problem can be stated as: assuming that the degrees n and m are known, the identification objective is to estimate the unknown plant parameters and time-delay from the irregularly sampled data $Z^N = \{u(t_k), y(t_k)\}_{k=1}^N$.

3. Estimation of the plant parameters when the time-delay is known

In this section, estimation of the plant parameters with the known time-delay is detailed. When the time-delay is assumed to be known, the estimation problem is simplified as a traditional CT model estimation problem. Amongst the available methods, the SRIVC algorithm has proven to be one of the most efficient algorithms in practical applications (see, *e.g.*, some recent examples in [12,3]) and is used in the proposed approach.

3.1. The SRIVC method

Prefiltering is an implicit or explicit component in CT transfer function model estimation. In the case of the SRIVC algorithm, a COE model (which takes the same form as system (1), but parameterized by $\boldsymbol{\theta}$ and τ) to be estimated can be expressed in the *pseudo-linear regression* manner. Let $F(p)$ be a stable filter taking the following form

$$F(p, \boldsymbol{\theta}) = \frac{1}{A(p, \boldsymbol{\theta})} \tag{4}$$

Then, the COE model can be equivalently reformulated in the following regression form (it is assumed here that $\tau = \tau^o$)

$$y_F(t_k, \boldsymbol{\theta}) = \boldsymbol{\phi}_F^T(t_k, \boldsymbol{\theta}, \tau^o)\boldsymbol{\theta} + \varepsilon(t_k, \boldsymbol{\theta}, \tau^o) \tag{5}$$

with

$$\boldsymbol{\phi}_F^T(t_k, \boldsymbol{\theta}, \tau^o) = [-y_F^{(n)}(t_k, \boldsymbol{\theta}) \ \cdots \ -y_F^{(1)}(t_k, \boldsymbol{\theta})$$
$$u_F^{(m)}(t_k - \tau^o, \boldsymbol{\theta}) \ \cdots \ u_F(t_k - \tau^o, \boldsymbol{\theta})] \tag{6}$$

where $\varepsilon(t_k, \boldsymbol{\theta}, \tau^o)$ is the equation error, $[\cdot]_F$ denotes the filtering operation, *i.e.* $[\cdot]_F = F(p, \boldsymbol{\theta})[\cdot]$. In the SRIVC algorithm, the IV vector $\boldsymbol{\psi}_F^T(t_k, \boldsymbol{\theta}, \tau^o)$ is chosen in the following way

$$\boldsymbol{\psi}_F^T(t_k, \boldsymbol{\theta}, \tau^o) = [-x_F^{(n)}(t_k, \boldsymbol{\theta}, \tau^o) \ \cdots \ -x_F^{(1)}(t_k, \boldsymbol{\theta}, \tau^o)$$
$$u_F^{(m)}(t_k - \tau^o, \boldsymbol{\theta}) \ \cdots \ u_F(t_k - \tau^o, \boldsymbol{\theta})] \tag{7}$$

where $x(t_k, \boldsymbol{\theta}, \tau^o)$ is the noise-free response of the following auxiliary model

$$x(t_k, \boldsymbol{\theta}, \tau^o) = \frac{B(p, \boldsymbol{\theta})}{A(p, \boldsymbol{\theta})}u(t_k - \tau^o) \tag{8}$$

From N measured input–output data, the IV estimator is given as

$$\hat{\boldsymbol{\theta}} = [\boldsymbol{\Psi}_F(\boldsymbol{\theta}, \tau^o)\boldsymbol{\Phi}_F^T(\boldsymbol{\theta}, \tau^o)]^{-1}\boldsymbol{\Psi}_F(\boldsymbol{\theta}, \tau^o)\mathbf{y}_F(\boldsymbol{\theta}) \tag{9}$$

with

$$\mathbf{y}_F(\boldsymbol{\theta}) = [y_F(t_s, \boldsymbol{\theta}) \ \cdots \ y_F(t_N, \boldsymbol{\theta})]^T$$
$$\boldsymbol{\Phi}_F(\boldsymbol{\theta}, \tau^o) = [\boldsymbol{\phi}_F(t_s, \boldsymbol{\theta}, \tau^o) \ \cdots \ \boldsymbol{\phi}_F(t_N, \boldsymbol{\theta}, \tau^o)]$$
$$\boldsymbol{\Psi}_F(\boldsymbol{\theta}, \tau^o) = [\boldsymbol{\psi}_F(t_s, \boldsymbol{\theta}, \tau^o) \ \cdots \ \boldsymbol{\psi}_F(t_N, \boldsymbol{\theta}, \tau^o)]$$

where $t_s \geq \tau^o$ and s is a positive integer.

Eq. (9) contains unknown parameters on the right-hand side, therefore the SRIVC algorithm employs an iterative way to find out the optimal estimate, where at each iteration, the auxiliary model
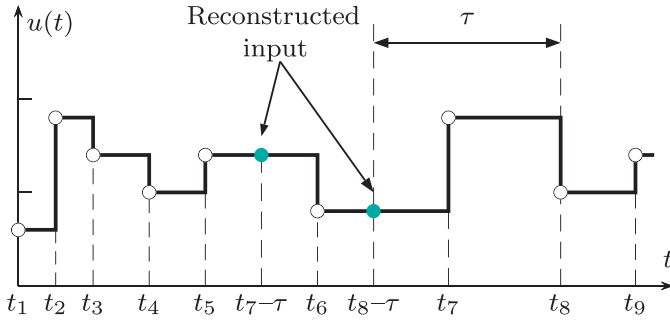
Fig. 1. Reconstruction of the signal at unobserved time-instants using ZOH assumption.

is used to generate the instrument and the prefilter based on the parameters obtained at the previous iteration, see, *e.g.* [9,10], for more details.

3.2. Implementation issues of the CT filtering operation in the presence of arbitrary time-delay

It is worth noticing that the computation of the SRIVC parameter estimates requires the value of prefiltered signals at the irregular time-instants t_k in both regression and instrument vectors, expressed under their developed forms in (6) and (7). The digital implementation issues of the CT filtering operations are well-known in CT model identification but they should be treated in an appropriate way here. Indeed, some components of the both vectors include an arbitrary time-delay and the latter requires special attention. Consider $u_F(t_k - \tau, \boldsymbol{\theta})$ for example

$$u_F(t_k - \tau, \boldsymbol{\theta}) = \frac{1}{A(p, \boldsymbol{\theta})} u(t_k - \tau) \tag{10}$$

The difficulty of computing $u_F(t_k - \tau, \boldsymbol{\theta})$ from irregularly sampled data is that $u(t)$ is unobserved at time-instant $t_k - \tau$. However, if the inter-sample behavior of $u(t_k)$ is *a priori* known, $u(t_k - \tau)$ can be exactly reconstructed from its neighboring data. This idea can be illustrated more clearly in Fig. 1 where the zero-order-hold (ZOH) assumption is used. Of course, other higher order assumptions, *e.g.* first-order-hold (FOH), could also be used.

The digital simulation of $u_F(t_k - \tau)$ is carried out in state-space form. Let the equivalent CT state-space model of (10) denoted as

$$\begin{cases} \dot{\boldsymbol{z}}(t - \tau) = \boldsymbol{F}_c \boldsymbol{z}(t - \tau) + \boldsymbol{G}_c u(t - \tau) \\ u_F(t - \tau) = \boldsymbol{H}_c \boldsymbol{z}(t - \tau) \end{cases} \tag{11}$$

where $\dot{\boldsymbol{z}}(t)$ denotes the time-derivative of $\boldsymbol{z}(t)$. Eq. (11) has the following sampled data model

$$\begin{cases} \boldsymbol{z}(t_{k+1} - \tau) = \boldsymbol{F}(h_k)\boldsymbol{z}(t_k - \tau) + \boldsymbol{M}(t_k) \\ u_F(t_k - \tau) = \boldsymbol{H}_c \boldsymbol{z}(t_k - \tau) \end{cases} \tag{12}$$

where

$$\boldsymbol{F}(h_k) = e^{\boldsymbol{F}_c h_k}, \quad \boldsymbol{M}(t_k) = \int_{t_k - \tau}^{t_{k+1} - \tau} e^{\boldsymbol{F}_c(t_{k+1} - \tau - t)} \boldsymbol{G}_c u(t) dt$$

where $h_k = t_{k+1} - t_k$. The computation of the integral $\boldsymbol{M}(t_k)$ needs more attention, consider the case that $u(t)$ is ZOH, $u(t)$ may change at some time-instants during the interval $[t_k - \tau, t_{k+1} - \tau]$, for example, $u(t)$ changes at time t_6 during $[t_7 - \tau, t_8 - \tau]$ in Fig. 1. Let $t_{k,1}, t_{k,n_k}$, where $n_k \geq 2$ is a positive integer, denote

$t_k - \tau$, $t_{k+1} - \tau$, respectively, and $\{t_{k,i} | i = 2, \cdots, n_k - 1\}$ the transition time-instants of $u(t)$, $\boldsymbol{M}(t_k)$ can be expanded as

$$\begin{aligned} \boldsymbol{M}(t_k) &= \sum_{i=1}^{n_k-1} \int_{t_{k,i}}^{t_{k,i+1}} e^{\boldsymbol{F}_c(t_{k,n_k} - t)} dt \quad \boldsymbol{G}_c u(t_{k,i}) \\ &= \sum_{i=1}^{n_k-1} e^{\boldsymbol{F}_c(t_{k,n_k} - t_{k,i+1})} \int_0^{h_{k,i}} e^{\boldsymbol{F}_c t} dt \quad \boldsymbol{G}_c u(t_{k,i}) \\ &= \sum_{i=1}^{n_k-1} \prod_{j=i+1}^{n_k-1} \boldsymbol{F}(h_{k,j})\boldsymbol{G}(h_{k,i})u(t_{k,i}) \end{aligned} \tag{13}$$

where $h_{k,i} = t_{k,i+1} - t_{k,i}$ and

$$\boldsymbol{G}(h_{k,i}) = \int_0^{h_{k,i}} e^{\boldsymbol{F}_c t} dt \quad \boldsymbol{G}_c \tag{14}$$

The exponentials involved in $\boldsymbol{F}(h_{k,i})$ and $\boldsymbol{G}(h_{k,i})$ could be computed by some standard methods, such as the expm routine in Matlab. However, the computational load in using expm will be very heavy when $h_{k,i}$ is time-varying, therefore, the following approach is derived to provides an fast approximation of $\boldsymbol{F}(h_{k,i})$ and $\boldsymbol{G}(h_{k,i})$

1. Make the following decomposition

   $$h_{k,i} = m_{k,i}\Delta + \delta_{k,i} < (m_{k,i} + 1)\Delta$$

   where Δ is the constant sampling period, $m_{k,i}$ is a positive integer, $\delta_{k,i} \geq 0$.
2. Compute $\boldsymbol{F}(\delta_{k,i})$, $\boldsymbol{G}(\delta_{k,i})$ by a fast approximated approach. Here we adopt the method presented in [28], and use the 4th-order Runge-Kutta (RK4) method to approximate the solutions.
3. Compute $\boldsymbol{F}(h_{k,i})$, $\boldsymbol{G}(h_{k,i})$ using the following relationships

   $$\begin{cases} \boldsymbol{F}(h_{k,i}) = \boldsymbol{F}(m_{k,i}\Delta)\boldsymbol{F}(\delta_{k,i}) \\ \boldsymbol{G}(h_{k,i}) = \boldsymbol{F}(m_{k,i}\Delta)\boldsymbol{G}(\delta_{k,i}) + \boldsymbol{G}(m_{k,i}\Delta) \end{cases}$$

Remark 1. Note that $\{\boldsymbol{F}(j\Delta), \boldsymbol{G}(j\Delta) | j = 1, 2, \cdots\}$ involved in Step 3 can be precomputed and stored in a numerical array. The main computation load is then reduced to the approximation of $\boldsymbol{F}(\delta_{k,i})$, $\boldsymbol{G}(\delta_{k,i})$ using the RK4 method. The bias order of approximated $\boldsymbol{F}(\delta_{k,i})$, $\boldsymbol{G}(\delta_{k,i})$ satisfies that $\boldsymbol{O}(\delta_{k,i}^5) < \boldsymbol{O}(\Delta^5)$, which depends on the chosen Δ.

Similar expressions can be obtained when the input of the filter is linear between the time-instants, as it is the case when the input is driven by first-order hold (FOH). The latter assumption is used to compute the filtered version of the time-derivatives of $y(t_k)$.

4. Estimation of the time-delay when the plant is known

In this section, the plant-independent estimation of the time-delay is investigated, therefore it is assumed that $\boldsymbol{\theta} = \boldsymbol{\theta}^o$. τ is assumed to be a 'pure delay' which can be expressed as an explicit parameter in the CT model. The prediction error as a function of τ is given as

$$\boldsymbol{\epsilon}(\boldsymbol{\theta}^o, \tau) = \boldsymbol{y} - G\left(p, \boldsymbol{\theta}^o\right) \boldsymbol{u}(\tau) \tag{15}$$

with

$$\boldsymbol{u}(\tau) = [u(t_s - \tau) \quad \cdots \quad u(t_N - \tau)]^T$$

$$\boldsymbol{\epsilon}(\boldsymbol{\theta}^o, \tau) = [\epsilon(t_s, \boldsymbol{\theta}^o, \tau) \quad \cdots \quad \epsilon(t_N, \boldsymbol{\theta}^o, \tau)]^T$$

where $t_s \geq \tau$. The estimate $\hat{\tau}$ is obtained by minimizing the following cost function

$$\hat{\tau} = \arg\min_{\tau} J_N(\boldsymbol{\theta}^o, \tau) \tag{16}$$

$$J_N(\boldsymbol{\theta}^o, \tau) = \frac{1}{2(N-s+1)}\boldsymbol{\epsilon}^T(\boldsymbol{\theta}^o, \tau)\boldsymbol{\epsilon}(\boldsymbol{\theta}^o, \tau) \tag{17}$$

The Gauss-Newton method can be used to solve (16), thus $\hat{\tau}$ can be found by the following iterative search

$$\hat{\tau}^{j+1} = \hat{\tau}^j - \mu^j[\nabla^2 J_N(\boldsymbol{\theta}^o, \hat{\tau}^j)]^{-1}\nabla J_N(\boldsymbol{\theta}^o, \hat{\tau}^j) \tag{18}$$

where μ is the step length, $\nabla J_N(\boldsymbol{\theta}^o, \hat{\tau}^j)$ and $\nabla^2 J_N(\boldsymbol{\theta}^o, \hat{\tau}^j)$ denote the gradient and the approximated Hessian matrix, respectively

$$\nabla J_N(\boldsymbol{\theta}^o, \tau) = \boldsymbol{\epsilon}_\tau^T \boldsymbol{\epsilon}(\boldsymbol{\theta}^o, \tau)$$

$$\nabla^2 J_N(\boldsymbol{\theta}^o, \tau) = \boldsymbol{\epsilon}_\tau^T \boldsymbol{\epsilon}_\tau$$

$$\boldsymbol{\epsilon}_\tau = \frac{\partial \boldsymbol{\epsilon}(\boldsymbol{\theta}^o, \tau)}{\partial \tau}$$

5. Separable method for continuous-time system identification with time-delay

Estimation of $\boldsymbol{\theta}$ and τ has been considered independently in the previous sections. The two methods are now combined together in a bootstrap manner to estimate simultaneously the time-delay and transfer function parameters.

5.1. Algorithm

The prediction error as a function of $\boldsymbol{\theta}$ and τ can be given as

$$\boldsymbol{\epsilon}(\boldsymbol{\rho}) = \boldsymbol{y} - G(p, \boldsymbol{\theta})\boldsymbol{u}(\tau)$$

Let us consider the following separable scheme for estimations of $\boldsymbol{\theta}$ and τ

Proposition 2 (**Separable Method**). *Let the estimate $\hat{\boldsymbol{\theta}}$ be a function of τ*

$$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\tau) = [\boldsymbol{\Psi}_F(\boldsymbol{\rho})\boldsymbol{\Phi}_F^T(\boldsymbol{\rho})]^{-1}\boldsymbol{\Psi}_F(\boldsymbol{\rho})\boldsymbol{y}_F(\boldsymbol{\theta}) \tag{19}$$

Then $\hat{\tau}$ can be obtained as

$$\arg\min_{\boldsymbol{\rho}} J_N(\boldsymbol{\rho}) = \arg\min_{\tau} \tilde{J}_N(\tau) \tag{20}$$

$$\hat{\tau} = \arg\min_{\tau} \tilde{J}_N(\tau) \tag{21}$$

where

$$\tilde{J}_N(\tau) = J_N(\boldsymbol{\rho})\big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

$$J_N(\boldsymbol{\rho}) = \frac{1}{2(N-s+1)}\boldsymbol{\epsilon}^T(\boldsymbol{\rho})\boldsymbol{\epsilon}(\boldsymbol{\rho})$$

where s is chosen that $t_s \geq t_0$. The iterative estimation of $\hat{\boldsymbol{\theta}}$ and $\hat{\tau}$ are given as

$$\hat{\tau}^{j+1} = \hat{\tau}^j - \mu^j[\nabla^2 \tilde{J}_N(\hat{\tau}^j)]^{-1}\nabla \tilde{J}_N(\hat{\tau}^j) \tag{22}$$

$$\hat{\boldsymbol{\theta}}^{j+1} = [\boldsymbol{\Psi}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})\boldsymbol{\Phi}_F^T(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})]^{-1}\boldsymbol{\Psi}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})\boldsymbol{y}_F(\hat{\boldsymbol{\theta}}^j) \tag{23}$$

where

$$\nabla \tilde{J}_N(\hat{\tau}^j) = \boldsymbol{\epsilon}_\tau^T \boldsymbol{\epsilon}(\hat{\boldsymbol{\rho}}^j) \tag{24}$$

$$\nabla^2 \tilde{J}_N(\hat{\tau}) \approx \boldsymbol{\epsilon}_\tau^T \boldsymbol{\epsilon}_\tau - \boldsymbol{\epsilon}_\tau^T \boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\boldsymbol{\epsilon}_{\boldsymbol{\theta}}^T \boldsymbol{\epsilon}_{\boldsymbol{\theta}})^{-1}\boldsymbol{\epsilon}_{\boldsymbol{\theta}}^T \boldsymbol{\epsilon}_\tau \tag{25}$$

$$\boldsymbol{\epsilon}_\tau = \frac{\partial \boldsymbol{\epsilon}(\boldsymbol{\rho})}{\partial \tau}\bigg|_{\boldsymbol{\rho}=\hat{\boldsymbol{\rho}}^j} = pG(p, \hat{\boldsymbol{\theta}}^j)\boldsymbol{u}(\hat{\tau}^j) \tag{26}$$

$$\boldsymbol{\epsilon}_{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{\epsilon}(\boldsymbol{\rho})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\rho}=\hat{\boldsymbol{\rho}}^j} = -\boldsymbol{\Psi}_F^T(\hat{\boldsymbol{\rho}}^j) \tag{27}$$

Proof. See [29] and the references therein for the details of the separable non-linear least squares method. □

The proposed algorithm (TDSRIVC) for CT model identification with time-delay can be summarized as

Algorithm 3 (*The basic TDSRIVC algorithm*).

1. Initialization: set the tolerances $\Delta\tau_{min}$, $\Delta\tilde{J}_N(\tau)_{min}$ along with the boundaries τ_{min}, τ_{max}. Based on the chosen initial value $\hat{\tau}^0$ and the initial filter $F(p)$, use the SRIVC method to compute $\hat{\boldsymbol{\theta}}^0$

$$F(p) = \frac{1}{(p/\omega_F + 1)^n} \tag{28}$$

where ω_F is a user provided value, see [9].
2. Iteration:
for $j=0$:convergence
a. Compute the increment $\Delta\hat{\tau}^j$
$$\Delta\hat{\tau}^j = -[\nabla^2 \tilde{J}_N(\hat{\tau}^j)]^{-1}\nabla\tilde{J}_N(\hat{\tau}^j)$$
b. Perform the following
i. Compute $\hat{\tau}^{j+1} = \hat{\tau}^j + \Delta\hat{\tau}^j$.
If $\hat{\tau}^{j+1} \notin [\tau_{min} \ \tau_{max}]$, let $\Delta\hat{\tau}^j = \Delta\hat{\tau}^j/2$ and repeat this step.
If $|\Delta\hat{\tau}^j| < \Delta\tau_{min}$, break.
ii. Estimate $\boldsymbol{\theta}$ using the SRIVC estimator
Update auxiliary model and prefilter
$$G(p, \hat{\boldsymbol{\theta}}^j) = B(p, \hat{\boldsymbol{\theta}}^j)/A(p, \hat{\boldsymbol{\theta}}^j)$$

$$F(p, \hat{\boldsymbol{\theta}}^j) = 1/A(p, \hat{\boldsymbol{\theta}}^j)$$
Compute the filtered signals in
$$\boldsymbol{\Psi}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1}), \quad \boldsymbol{\Phi}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1}), \quad \boldsymbol{y}_F(\hat{\boldsymbol{\theta}}^j)$$
Estimate $\boldsymbol{\theta}$
$$\hat{\boldsymbol{\theta}}^{j+1} = [\boldsymbol{\Psi}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})\boldsymbol{\Phi}_F^T(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})]^{-1}\boldsymbol{\Psi}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})\boldsymbol{y}_F(\hat{\boldsymbol{\theta}}^j)$$
If $\tilde{J}_N(\hat{\tau}^{j+1}) \geq \tilde{J}_N(\hat{\tau}^j)$, let $\Delta\hat{\tau}^j = \Delta\hat{\tau}^j/2$ and go to (i).
c. Check the stopping condition. Go to step (a) if the following condition is satisfied, else break.
$$\tilde{J}_N(\hat{\tau}^j) - \tilde{J}_N(\hat{\tau}^{j+1}) \geq \Delta\tilde{J}_N(\tau)_{min}$$
end

The proposed bootstrap algorithm provides very good estimates when the initial time-delay value is closed to the true one. However, the gradient-based method might be trapped in a local minimum since the cost function $J(\boldsymbol{\theta}, \tau)$ is a nonlinear function of the time-delay τ. To circumvent this problem, a low-pass filtering strategy is introduced in the next section to improve the convergence performance.

6. Expansion of the convergence region by low-pass filtering

6.1. The case without low-pass filtering

Consider the objective function with respect to the CT error

$$J(\boldsymbol{\theta}, \tau) = \frac{1}{2}\int_{-\infty}^{\infty} \epsilon^2(t, \boldsymbol{\theta}, \tau)dt \tag{29}$$

From Parseval's theorem, $\bar{J}(\boldsymbol{\theta}^o, \tau)$ has the following frequency-domain interpretation

$$J(\boldsymbol{\theta}, \tau) = \frac{1}{4\pi}\int_{-\infty}^{+\infty} \left[\left|G(i\omega, \boldsymbol{\theta}^o)e^{-i\omega\tau^o} - G(i\omega, \boldsymbol{\theta})e^{-i\omega\tau}\right|^2 \Phi_u(\omega) \right.$$
$$\left. + \Phi_e(\omega) \right] d\omega \tag{30}$$

where $\Phi_u(\omega)$ and $\Phi_e(\omega)$ are the power spectral density (PSD) functions of $u(t)$ and $e(t)$, respectively. According to [27], $G(i\omega, \theta^o)$ can be formulated in terms of $G(i\omega, \theta)$ and the model error $M(\omega)e^{i\varphi(\omega)}$

$$G(i\omega, \theta^o) = G(i\omega, \theta)[1 + M(\omega)e^{i\varphi(\omega)}] \tag{31}$$

Let the distance between τ and τ^o by $\delta\tau$

$$\delta\tau = \tau - \tau^o \tag{32}$$

substituting (31) and (32) into $J(\theta, \tau)$ and removing the constant term yields the following equation

$$V(\theta, \delta\tau) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left|G(i\omega, \theta)\right|^2 \left|1 - e^{i\omega\delta\tau}[1 + M(\omega)e^{i\varphi(\omega)}]\right|^2 \Phi_u(\omega)d\omega \tag{33}$$

Since the same results can be drawn from both cost functions $J(\theta, \tau)$ and $V(\theta, \delta\tau)$, $V(\theta, \delta\tau)$ will be used in the following sections for further analysis.

To quantify the width of the convergence region, the concept of *admissible region* is given below

Definition 4 (*Admissible Region*). If $\delta\tau_{max}$ is the maximum value which guarantees that $V(\theta, \delta\tau)$ is unimodal when $|\delta\tau| \leq \delta\tau_{max}$, then $\delta\tau_{max}$ is termed as the admissible region. When $\tau \in \left[\tau^o - \delta\tau_{max}, \tau^o + \delta\tau_{max}\right]$, $J(\theta, \tau)$ is unimodal with respect to τ.

However, when $\theta \neq \theta^o$, the admissible region for $V(\theta, \delta\tau)$ is quite hard to determine because $M(\omega) \neq 0$ (see [27] for more discussions). In this paper, we only analyze the simple case when $\theta = \theta^o$. The following theorem is used to compute $V(\theta^o, \delta\tau)$ and the corresponding $\delta\tau_{max}$ based on the output signal.

Theorem 5. *Assuming that the PSD function of $u(t)$ is denoted by $\Phi_u(\omega)$. Let $x(t)$ denote the response of the true system*

$$x(t) = G(p, \theta^o)u(t - \tau^o)$$

Then $V(\theta^o, \delta\tau)$ can be computed by

$$V(\theta^o, \delta\tau) = R_x(0) - R_x(\delta\tau) \tag{34}$$

where R_x is the autocorrelation function of $x(t)$

$$R_x(\delta\tau) = \mathbb{E}\left\{x(t)x(t - \delta\tau)\right\}$$

Proof. This theorem is quite similar to Theorem 1 in [27], the difference is that we do not apply a low-pass filter here. By setting $M(\omega)$ to zero, $V(\theta^o, \delta\tau)$ can be given as

$$
\begin{aligned}
V(\theta^o, \delta\tau) &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left|G(i\omega, \theta^o)\right|^2 \left|1 - e^{i\omega\delta\tau}\right|^2 \Phi_u(\omega)d\omega \\
&= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \Phi_x(\omega)(2 - e^{-i\omega\delta\tau} - e^{i\omega\delta\tau})d\omega
\end{aligned} \tag{35}
$$

where $\Phi_x(\omega) = \left|G(i\omega, \theta^o)\right|^2 \Phi_u(\omega)$ is the PSD function of $x(t)$. By using the well-known *Wiener-Khinchin* theorem, (35) can be reformulated as
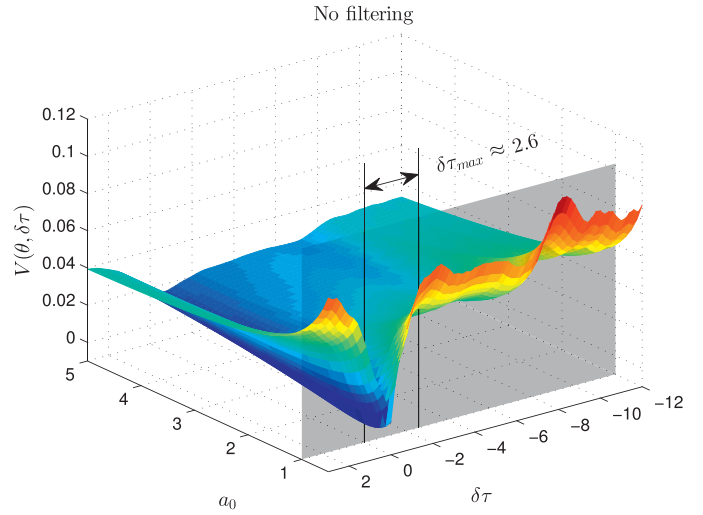
$$V(\theta^o, \delta\tau) = R_x(0) - \frac{1}{2}R_x(-\delta\tau) - \frac{1}{2}R_x(\delta\tau) = R_x(0) - R_x(\delta\tau) \tag{36}$$

The second equation in (36) uses the fact that $R_x(\delta\tau)$ is an even function. □

The following example gives a graphical illustration of the relationship between $V(\theta, \delta\tau)$ and the unknown parameters.

Example 6. Consider the following noise-free system

$$x(t_k) = \frac{1}{a_0^o p + 1}u(t_k - \tau^o) \tag{37}$$

Fig. 2. The surface plot of $V(\theta, \delta\tau)$ as a function of a_0 and $\delta\tau$. $\delta\tau_{max}$ is computed when $a_0 = 1$. The intersection line of the transparent plane and the surface plot is $V(\theta^o, \delta\tau)$.

where $a_0^o = 1$, $\tau^o = 12\,s$. The system is excited by $u(t)$ with $\Phi_u(\omega) = 1$. The number of observations is $N = 1000$, the sampling interval h_k is uniformly distributed in the following interval

$$h_k \sim U[0.01, 0.09]\,s$$

The surface plot of $V(\theta, \delta\tau)$ as a function of a_0 and $\delta\tau$ is shown in the Fig. 2. When $a_0 = a_0^o = 1$, the admissible region is $\delta\tau_{max} \approx 2.6$, it indicates that the time-delay search will converge to the true value when it is initialized within the region $[9.4, 14.6]\,s$. Note that the intersection line $V(\theta^o, \delta\tau)$ of the surface plot and the transparent plane can be obtained by Theorem 5, thus $\delta\tau_{max} \approx 2.6$ can be obtained without any knowledge of the true parameters.

6.2. Widening of the convergence region by low-pass filtering

To increase the chance of converging to the global minimum, the use of low-pass filter was initially suggested in [27]. It was shown that the use of filtered signals in estimation can widen the global convergence region in searching the time-delay.

Let the filtered error and cost function be denoted by

$$\overline{\epsilon}(t, \theta, \tau) = L(p)\epsilon(t, \theta, \tau) \tag{38}$$

$$\overline{J}(\theta, \tau) = \frac{1}{2} \int_{-\infty}^{\infty} \overline{\epsilon}^2(t, \theta, \tau)dt \tag{39}$$

where $L(p)$ is a CT low-pass filter. Even if the theory is attractive, from a practical side, there is an important issue to be discussed which is the choice of $L(i\omega)$. Amongst the possible options, an ideal low-pass filter is easier to specify and will be used in the proposed approach. In the situation of ideal low-pass filtering, Theorem 5 should be modified as follows

Proposition 7. *Assume $L(i\omega)$ is an ideal low-pass filter with the cutoff frequency of ω_L*

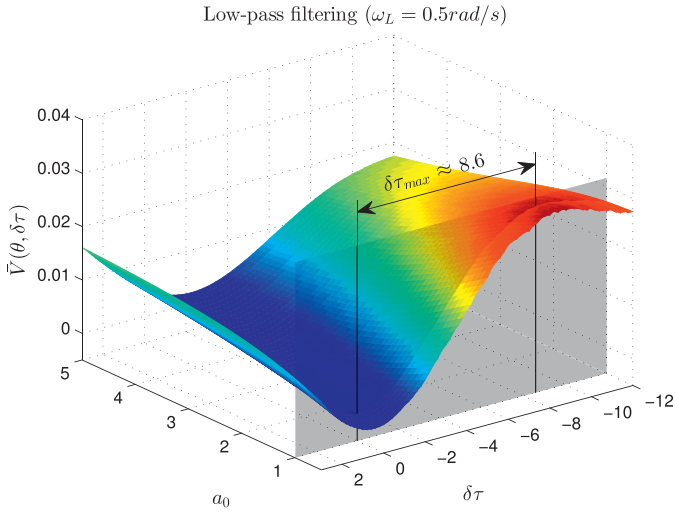$$L(i\omega) = \begin{cases} 1 & |\omega| \leq \omega_L \\ 0 & \text{otherwise} \end{cases} \tag{40}$$

Let

$$\overline{x}(t) = L(p)x(t) \tag{41}$$

where $x(t)$ is defined in Theorem 5. $[\overline{\cdot}]$ denotes the filtering operation. Based on this filtered signal, Eq. (34) in Theorem 5 is modified as

$$\overline{V}(\theta^o, \delta\tau) = R_{\overline{x}}(0) - R_{\overline{x}}(\delta\tau) \tag{42}$$

Low-pass filtering ($\omega_L = 0.5\,rad/s$)

Fig. 3. The filtered index term $\overline{V}(\boldsymbol{\theta}, \delta\tau)$ as a function of a_0 and $\delta\tau$ ($\omega_L = 0.5\,rad/s$). $\delta\tau_{max}$ is computed when $a_0 = 1$. The intersection line of the transparent plane and the surface plot is $V(\boldsymbol{\theta}^o, \delta\tau)$

The following example is used to illustrate the impact of the low-pass filter.

Example 8. Consider the same system as Example 6. An ideal low-pass filter $L(i\omega)$ having the cutoff frequency of $\omega_L = 0.5\,rad/s$ is applied to expand the convergence region. The filtered index term $\overline{V}(\boldsymbol{\theta}, \tau)$ as a function of τ and a_0 is plotted in Fig. 3. When $a_0 = a_0^o = 1$, the admissible region is widened to $\delta\tau_{max} \approx 8.6$, the corresponding convergence region is [3.4, 20.6] s. This figure also illustrates that the sharp peaks of the cost function are 'spread out' by the low-pass filter, which makes the global minimum easier to reach.

6.3. The robust TDSRIVC algorithm

Since the low-pass filter can widen the convergence region, the TDSRIVC algorithm is then modified to take this advantage. We suggest to run the TDSRIVC algorithm in the first m iterations by minimizing the filtered cost function $\bar{J}(\boldsymbol{\theta}, \tau)$, which could yield a relatively accurate estimate from a coarse initialization, then we move to minimize the unfiltered cost function $J(\boldsymbol{\theta}, \tau)$ until the convergence occurs. The modified algorithm is summarized below

Algorithm 9 (The robust TDSRIVC algorithm).

1. Initialization.
2. Iteration:
- The first stage:
  for j=0:m
  Perform the estimation by minimizing the filtered cost function

$$\hat{\boldsymbol{\rho}} = \arg\min_{\boldsymbol{\rho}} \bar{J}_N(\boldsymbol{\rho}) \tag{43}$$

where [$\bar{}$] denotes the additional filtering operation by $L(p)$. Eqs. (22) and (23) are modified as

$$\hat{\tau}^{j+1} = \hat{\tau}^j - \mu^j [\nabla^2 \tilde{\bar{J}}_N(\hat{\tau}^j)]^{-1} \nabla \tilde{\bar{J}}_N(\hat{\tau}^j) \tag{44}$$

$$\hat{\boldsymbol{\theta}}^{j+1} = [\overline{\boldsymbol{\Psi}}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1}) \overline{\boldsymbol{\Phi}}_F^T(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1})]^{-1} \overline{\boldsymbol{\Psi}}_F(\hat{\boldsymbol{\theta}}^j, \hat{\tau}^{j+1}) \overline{\boldsymbol{y}}_F(\hat{\boldsymbol{\theta}}^j) \tag{45}$$

end
- The second stage:
  for j=m:convergence

Perform the estimation by minimizing the unfiltered cost function

$$\hat{\boldsymbol{\rho}} = \arg\min_{\boldsymbol{\rho}} J_N(\boldsymbol{\rho}) \tag{46}$$

Eqs. (22) and (23) are used to estimate the delay and the plant parameters.

end

Remark 10. The question remains here to know how ω_L should be chosen. The previous statements conclude that the global minimum is easier to access for smaller values of ω_L. Then, one may ask, can ω_L be arbitrarily small? Our answer is no, the reason can be given as: (1) The NUFFT is used to implement the ideal low-pass filtering, which has a minimum resolution. So it does not make sense when ω_L is lower than this resolution. (2) Smaller values for ω_L mean that more information of the sampled data is removed, which can deteriorate the plant parameter estimation in Algorithm 9. A rule of thumb can be to choose the cut-off frequency of the ideal low-pass filter as (1/10) to (1/2) of the system bandwidth. Another problem is the choice of m, a rule of thumb is to chose a value between 5 and 10, the convergence region can always be extended.

6.4. Implementation of the ideal low-pass filtering from irregularly sampled data

Once the cut-off frequency of the ideal filter is chosen, there is a last issue to be investigated: how to implement the ideal low-pass filtering operations in the present irregularly sampled data situation? One efficient way is to carry out the latter by using the non-uniform fast Fourier transform (NUFFT) and its inverse, by adding a rectangular window to the frequency data. Different approximated methods for fast Fourier transform (FFT) from irregularly sampled data were reviewed in [30]. The basic idea is to reconstruct a sequence of *regularly sampled data*, then the traditional fast Fourier transform can be applied. Here we use the so called *polynomial interpolation through neighboring grid points* suggested in [30] where the nearest 2 points are used to construct the regularly resampled data.

Assuming that the sampled data is $\{y(t_k)\}_{k=1}^N$, where t_k is irregularly spaced. Let $\{y'_\ell\}_{\ell=1}^{N'}$ denote the regularly resampled data, with constant sampling period Δ

$$N' = 2^M \geq N > 2^{M-1}, \qquad M = 1, 2, 3, \cdots \tag{47}$$

$$y'_\ell = \frac{y(t_{k+1}) - y(t_k)}{t_{k+1} - t_k}(\ell\Delta - t_k) + y(t_k), \qquad t_k \leq \ell\Delta < t_{k+1} \tag{48}$$

Based on y'_ℓ, apply fast Fourier transform

$$F_q = \sum_{\ell=1}^{N'} y'_\ell e^{-i2\pi \frac{(\ell-1)(q-1)}{N'}}, \qquad q = 1, \cdots, N' \tag{49}$$

Assuming that the ideal low-pass filter $L(i\omega)$ has the cutoff frequency of ω_L, N_c is an integer ($1 \leq N_c < N'/2$)
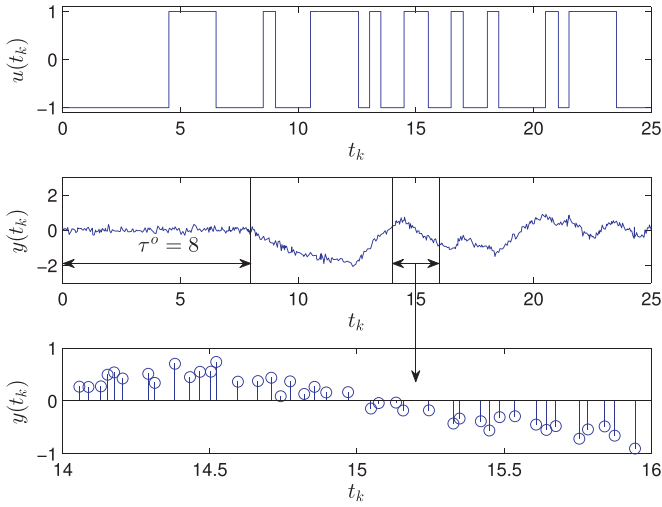
$$\frac{N_c}{N'} \cdot \frac{2\pi}{\Delta} \leq \omega_L < \frac{N_c + 1}{N'} \cdot \frac{2\pi}{\Delta} \tag{50}$$

We define the following signal

$$G_q = \begin{cases} 0, & N_c + 1 < q \leq N' - N_c \\ 1, & \text{otherwise} \end{cases}$$

Then the filtered version of y'_ℓ by $L(i\omega)$ can be obtained by the inverse Fourier transform

$$\overline{y}'_\ell = \frac{1}{N'} \sum_{q=1}^{N'} F_q G_q e^{i2\pi \frac{(\ell-1)(q-1)}{N'}}, \qquad \ell = 1, \cdots, N' \tag{51}$$

Fig. 4. Sampled input–output data for System 1. Top – The input signal. Middle – The output signal where the presence of the delayed response can be observed and bottom – Zoom part of the sampled output where the irregular sampling can be observed.

where $[\bar{\cdot}]$ denotes the filtering operation by $L(i\omega)$. Finally $\bar{y}(t_k)$ can be obtained by the following interpolation

$$\bar{y}(t_k) = \frac{\bar{y}'_{\ell+1} - \bar{y}'_\ell}{\Delta}(t_k - \ell\Delta) + \bar{y}'_\ell, \quad \ell\Delta \le t_k < (\ell+1)\Delta \tag{52}$$

7. Numerical examples

7.1. Identification from irregularly sampled data

In this section, the identification problem for systems of different orders is considered. In the simulation, the following conditions are assumed

1. The systems to be identified take the general form of (1), with true parameters given in Table 1. The input signal is chosen to be a pseudo random binary sequence (PRBS) switching between ± 1, which is generated from a 9-stage shift register with the clock period of 0.5 s. The input–output signals are observed at $N = 1500$ time-instants, the sampling interval h_k is uniformly distributed in the following interval

$$h_k \sim U[0.01, 0.09]\, s \tag{53}$$

The variance of $e(t_k)$ is adjusted to obtain a signal-to-noise ratio (SNR) of 15 dB. The SNR is defined below

$$\text{SNR} = 10\log\frac{P_x}{P_e}$$

where P_x and P_e represent the average power of the noise-free signal $x(t_k)$ and the disturbance $e(t_k)$, respectively. A portion of the irregularly sampled data for System 1 is plotted in Fig. 4.

2. Monte-Carlo (MC) simulation with 100 runs is performed to obtain the statistical estimation results. Each run has random realizations of h_k and $e(t_k)$. To initiate the TDSRIVC algorithm, one needs to provide ω_F and $\hat{\tau}^0$ (see Algorithm 3). It has been declared in [10] (p250) that the choice of ω_F is not critical for the SRIVC algorithm since convergence occurs over a fairly wide range of values. In this section, ω_F, ω_L and $\hat{\tau}^0$ are chosen as

$$\omega_F = 1\, rad/s, \omega_L = 0.1\, rad/s, \quad \hat{\tau}^0 = \{1, 3, 5, 7, 9\}\, s \tag{54}$$

3. The normalized root-mean-square error (NRMSE) is used as the benchmark for performance evaluation, it is computed by the following formula

$$\text{NRMSE} = 100\left(1 - \frac{\left\|y(t) - \hat{x}(t)\right\|_2}{\left\|y(t) - \mathbb{E}\{y(t)\}\right\|_2}\right) \tag{55}$$

where $\hat{x}(t)$ is the noise-free response of the estimated model. When SNR = 15 dB, NRMSE satisfies the following inequality

$$\text{NRMSE} \le 100\left(1 - \frac{\sqrt{P_e}}{\sqrt{P_x + P_e}}\right) = 100\left[1 - (1 + 10^{\text{SNR}/10})^{-1/2}\right]$$

$$\approx 82.5$$

In this paper, the criterion for global convergence is chosen as NRMSE > 81.

4. The following estimation strategies are considered:
   A. Identification without low-pass filtering (Algorithm 3). τ_{min} and τ_{max} are set to 0 and 10, respectively. The stopping rule algorithm is that the relative change in $J_N(\hat{\rho})$ is lower than 0.01%. The basic TDSRIVC algorithm is allowed to have 50 iterations at most.
   B. Identification with low-pass filtering (Algorithm 9). τ_{min}, τ_{min} keep the same. The maximum number of iterations of the first estimation stage (With low-pass filtering) is set to 10. During these 10 iterations, the algorithm will be terminated if the relative change of $\bar{J}_N(\rho)$ is smaller than 0.01%. Subsequently, the algorithm proceeds to the second estimation stage (No filtering), the maximum number of iterations in this stage is set to 50.

MC simulations (each has 100 realizations) are performed by using different initializations, more specifically, ω_F and ω_L are fixed and $\hat{\tau}^0$ picks up a value from $\{1, 3, 5, 7, 9\}$ sequentially. The global convergence ratios of the MC simulations are gathered in Table 1. It can be observed that Strategy A (No filtering) is more probable to converge when good initial guesses (Around $\tau^o = 8\, s$) of the time-delay are provided, while Strategy B (Low-pass filtering) can converge from a coarse initialization.

7.2. Identification from uniformly sampled data: a comparison with PROCEST

The proposed method can also handle uniformly sampled data by assuming that the sampling interval h_k is constant. In this case,

Table 1
Monte-Carlo simulation results for the proposed method with and without low-pass filtering from irregularly sampled data. P_{glob}–the ratio of global convergence.

No.	True system	P_{glob} without filtering (%)					P_{glob} with filtering (%)				
	$\hat{\tau}^0$	1 s	3 s	5 s	7 s	9 s	1 s	3 s	5 s	7 s	9 s
1.	$\frac{2e^{-8s}}{2s+1}$	5	3	9	49	97	100	99	99	100	98
2.	$\frac{3e^{-8s}}{0.25s^2+s+1}$	0	1	15	100	100	100	100	100	99	93
3.	$\frac{2e^{-8s}}{0.25s^2+0.7s+1}$	12	10	13	98	61	98	100	99	98	62
4.	$\frac{(-4s+1)e^{-8s}}{9s^2+2.4s+1}$	0	8	49	97	20	98	98	99	90	76

Table 2

Parameter estimates from regularly sampled data with low-pass filtering (Strategy B). $\overline{N}_{iter}$-The averaged number of iterations for global convergence, P_{glob}-the ratio of global convergence. The mean value, the standard deviation are computed from the estimated models that converge to the global minimum.

No.	True system	Method	$\hat{a}_0$	$\hat{a}_1$	$\hat{b}_0$	$\hat{b}_1$	$\hat{\tau}$	$\overline{N}_{iter}$	P_{glob} (%)
1.	$\frac{2e^{-8s}}{2s+1}$	TDSRIVC	1.9999 ± 0.0207	–	2.0005 ± 0.0129	–	7.9996 ± 0.0038	11.7	100
		PROCEST	1.9996 ± 0.0205	–	2.0003 ± 0.0127	–	7.9996 ± 0.0038	15.2	100
2.	$\frac{3e^{-8s}}{0.25s^2+s+1}$	TDSRIVC	0.2488 ± 0.0077	0.9984 ± 0.0086	3.0007 ± 0.0180	–	8.0011 ± 0.0068	11.3	100
		PROCEST	0.2489 ± 0.0042	0.9984 ± 0.0080	3.0007 ± 0.0168	–	8.0010 ± 0.0049	21.2	96
3.	$\frac{2e^{-8s}}{0.25s^2+0.7s+1}$	TDSRIVC	0.2492 ± 0.0049	0.6992 ± 0.0058	2.0011 ± 0.0116	–	8.0009 ± 0.0057	10.7	100
		PROCEST	0.2489 ± 0.0044	0.6987 ± 0.0057	2.0005 ± 0.0113	–	8.0013 ± 0.0053	21.5	73
4.	$\frac{(-4s+1)e^{-8s}}{9s^2+2.4s+1}$	TDSRIVC	8.9988 ± 0.0563	2.3975 ± 0.0293	-3.9947 ± 0.0351	0.9985 ± 0.0158	7.9994 ± 0.0062	12.0	93
		PROCEST	9.0009 ± 0.0548	2.3964 ± 0.0290	-3.9938 ± 0.0344	0.9986 ± 0.0155	7.9993 ± 0.0060	15.7	82

the PROCEST routine available in the Matlab System Identification (SID) Toolbox (Version 2014a) is also used to give a comparison. PROCEST is a variation of the prediction error minimization (PEM) method, which can estimate up to a 3rd order simple process model plus delay from regularly sampled data only (see [31]).

Compared with the previous section, the simulation conditions keep the same except the following two

- The constant sampling interval is chosen as $h_k = 0.05\,s$.
- The initial value of the time-delay is generated from a uniform distribution $\hat{\tau}^0 \sim U[0, 10]\,s$.

The settings for the PROCEST routine are given here: the minimum and maximum values of the time-delay τ are chosen as 0 and 10, respectively. 'SearchMethod'='gna', 'Tolerance'=0.01%. When Strategy B is used in estimation, a focus filter (low-pass filter) is performed in the first stage by setting 'Focus'=[0, ω_L], the maximum number of iterations is set to 'MaxIter'=10. Then in the second stage, 'Focus', 'MaxIter' are changed to 'Prediction' and 50, respectively. The same filter $F(p)$ defined in (28) and time-delay $\hat{\tau}^0$ are used to initialize the PROCEST routine in identifying System $1 \sim 3$. In identifying System 4, the initial model is $(-3p+1)F(p)$. The number of iterations is the total iterations in the both estimation stages.

The results of strategy B are presented in Table 2. For the first order system, TDSRIVC and PROCEST routines provide identical and very good global convergence ratios. When the system has more parameters to estimate (System 3 and 4), the global convergence ratio slightly drops for the PROCEST routine while the proposed TDSRIVC method is still able to provide fairly good results. The proposed TDSRIVC seems however to converge more quickly. The robustness of TDSRIVC comes from the fact that it has only one dimension (the time-delay) to be estimated by numerical search, the other parameters are estimated by the robust SRIVC method. By contrast, all the parameters are estimated by numerical search in PROCEST.

8. Conclusion

In this paper, an approach to continuous-time system identification with arbitrary time-delay from irregularly sampled input–output data has been presented. The proposed approach combines the SRIVC method in a bootstrap manner with an adaptive gradient-based search algorithm. The implementation issues of CT filtering from irregularly sampled data have been addressed. In order to improve the convergence property, a low-pass filter strategy has been introduced to widen the convergence region. Simulation results have shown that the proposed method is of low complexity and converges to the true parameters from a relatively wide range of time-delay initial values. Numerical examples have shown that the proposed method has similar or superior convergence property to the explicit method available in the System Identification toolbox for simple process plus delay estimation

which can handle regularly sampled data only. This new algorithm provides therefore more choices for CT delayed system identification and constitutes a very useful addition to the CONTSID[2] toolbox.

References

[1] T. Söderström, P. Stoica, System Identification, Series in Systems and Control Engineering, Prentice Hall, Englewood Cliffs, 1989.
[2] L. Ljung, System Identification – Theory for the User, 2nd edn., Prentice-Hall, Upper Saddle River, 1999.
[3] H. Garnier, P.C. Young, The advantages of directly identifying continuous-time transfer function models in practical applications, Int. J. Control 87 (7) (2014) 1319–1338.
[4] G.P. Rao, H. Unbehauen, Identification of continuous-time systems, IET Control Theory Appl. 153 (2) (2006) 185–220.
[5] K.J. Åström, B. Bernhardsson, Systems with Lebesgue sampling, in: A. Rantzer, C.I. Byrnes (Eds.), Directions in Mathematical Systems Theory and Optimization, Springer-Verlag, Berlin Heidelberg, 2003, pp. 1–13.
[6] P.C. Young, Some observations on instrumental variable methods of time-series analysis, Int. J. Control 23 (5) (1976) 593–612.
[7] T. Söderström, How accurate can instrumental variable models become? in: L. Wang, H. Garnier (Eds.), System Identification, Environmental Modelling, and Control System Design, Springer-Verlag, London, 2012, pp. 3–26.
[8] P.C. Young, A. Jakeman, Refined instrumental variable methods of recursive time-series analysis Part I. Single input, single output systems Int. J. Control 29 (1) (1979) 1–30.
[9] P.C Young, H. Garnier, M. Gilson, Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models, in: H. Garnier, L. Wang (Eds.), Identification of Continuous-Time Models from Sampled Data, Springer-Verlag, London, 2008, pp. 91–132.
[10] P.C. Young, Recursive Estimation and Time-series Analysis: An Introduction for the Student and Practitioner, Springer-Verlag, Berlin, 2011.
[11] P.C. Young, H. Garnier, Identification and estimation of continuous-time data-based mechanistic (DBM) models for environmental systems, Environ. Model. Softw. 21 (8) (2006) 1055–1072.
[12] H. Garnier, M. Gilson, P.C. Young, E. Huselstein, An optimal IV technique for identifying continuous-time transfer function model of multiple input systems, Control Eng. Pract. 15 (4) (2007) 471–486.
[13] M. Gilson, H. Garnier, P.C. Young, P. Van den Hof, Instrumental variable methods for closed-loop continuous-time model identification, in: H. Garnier, L. Wang (Eds.), Identification of Continuous-Time Models from Sampled Data, Springer, London, 2008, pp. 133–160.
[14] V. Laurain, M. Gilson, R. Toth, H. Garnier, Refined instrumental variable methods for identification of LPV Box-Jenkins models, Automatica 46 (6) (2010) 959–967.
[15] B. Ni, M. Gilson, H. Garnier, A refined instrumental variable method for Hammerstein-Wiener continuous-time model identification, IET Control Theory Appl. 7 (9) (2013) 1276–1286.
[16] J. Schorsch, H. Garnier, M. Gilson, P.C. Young, Instrumental variable methods for identifying partial differential equation models, Int. J. Control 86 (12) (2013) 2325–2335.
[17] J. Wang, W. Zheng, T. Chen, Identification of linear dynamic systems operating in a networked environment, Automatica 45 (12) (2009) 2763–2772.
[18] F. Chen, H. Garnier, M. Gilson, Refined instrumental variable identification of continuous-time OE and BJ models from irregularly sampled data, in: 11th IFAC International Workshop on Adaptation and Learning in Control and Signal Processing (ALCOSP'2013), Caen, France, 2013.
[19] S. Björklund, A Survey and Comparison of Time-Delay Estimation Methods in Linear Systems, Linköping University, Sweden, 2003 (Ph.D. thesis).
[20] S. Ahmed, B. Huang, S.L. Shah, Parameter and delay estimation of continuous-time models using a linear filter, J. Process Control 16 (4) (2006) 323–331.
[21] Z. Yang, H. Iemura, S. Kanae, K. Wada, Identification of continuous-time systems with multiple unknown time delays by global nonlinear least-squares and instrumental variable methods, Automatica 43 (7) (2007) 1257–1264.

[22] B. Ni, D. Xiao, S.L. Shah, Time delay estimation for MIMO dynamical systems – With time – frequency domain analysis, J. Process Control 20 (1) (2010) 83–94.
[23] Q. Wang, Y. Zhang, Robust identification of continuous systems with dead-time from step responses, Automatica 37 (3) (2001) 377–390.
[24] T. Liu, F. Gao, Robust step-like identification of low order process model under nonzero initial conditions and disturbance, IEEE Trans. Autom. Control 53 (11) (2008) 2690–2695.
[25] T. Liu, F. Gao, A frequency domain step response identification method for continuous-time processes with time delay, J. Process Control 20 (7) (2010) 800–809.
[26] S. Ahmed, B. Huang, S.L. Shah, Novel identification method from step response, Control Eng. Pract. 15 (5) (2006) 545–556.
[27] G. Ferretti, C. Maffezzoni, R. Scattolini, On the identifiability of the time delay with least-squares methods, Automatica 32 (3) (1996) 449–453.
[28] C.F. Van Loan, Computing integrals involving the matrix exponentials, IEEE Trans. Autom. Control 23 (3) (1978) 395–404.
[29] L.S.H. Ngia, Separable nonlinear least-squares methods for efficient off-line and on-line modeling of systems using Kautz and Laguerre filters, IEEE Trans. Circ. Syst. II: Analog Digit. Signal Process. 48 (6) (2001) 562–579.
[30] A.F. Ware, Fast approximate Fourier transforms for irregularly spaced data, SIAM Rev. 40 (4) (1998) 838–856.
[31] L. Ljung, Version 8 of the system identification toolbox, in: 16th IFAC Symposium on System Identification, Brussels, Belgium, 2012.