



HAL
open science

Parameter and differentiation order estimation in fractional models

Stephane Victor, Rachid Malti, Hugues Garnier, Alain Oustaloup

► **To cite this version:**

Stephane Victor, Rachid Malti, Hugues Garnier, Alain Oustaloup. Parameter and differentiation order estimation in fractional models. *Automatica*, 2013, 49 (4), pp.926-935. 10.1016/j.automatica.2013.01.026 . hal-00804784

HAL Id: hal-00804784

<https://hal.science/hal-00804784v1>

Submitted on 16 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Parameter and differentiation order estimation in fractional models

Stéphane Victor ^a Rachid Malti ^a Hugues Garnier ^b and Alain Oustaloup ^a

^a *Université de Bordeaux, IMS UMR 5218 CNRS, 351 cours de la Libération, F-33400 Talence, France*
{stephane.victor, rachid.malti, alain.oustaloup}@ims-bordeaux.fr

^b *Université de Lorraine, Centre de Recherche en Automatique de Nancy, CNRS, 2 rue Jean Lamour, 54519*
Vandoeuvre-lès-Nancy Cedex, France – hugues.garnier@univ-lorraine.fr

Abstract

This paper deals with continuous-time system identification using fractional differentiation models. An adapted version of the simplified refined instrumental variable method is first proposed to estimate the parameters of the fractional model when all the differentiation orders are assumed known. Then, an optimization approach based on the use of the developed instrumental variable estimator is presented. Two variants of the algorithm are proposed. Either, all differentiation orders are set as integral multiples of a commensurate order which is estimated, or all differentiation orders are estimated. The former variant allows to reduce the number of parameters and can be used as a good initial hit for the latter variant. The performances of the proposed approaches are evaluated by Monte Carlo simulation analysis. Finally, the proposed identification algorithms are used to identify thermal diffusion in an experimental set-up.

Key words: continuous-time models; fractional calculus; instrumental variable; system identification; model order optimization.

1 Introduction

Fractional calculus is a generalization of the traditional calculus and dates back to (Liouville, 1832) and (Riemann, 1876). It remained for a long time an abstract mathematical concept until used by (Oldham and Spanier, 1970; Oldham and Spanier, 1974) for modeling general diffusive phenomena in semi-infinite planar, spherical, and cylindrical media. The authors showed that diffusive systems can be modeled by transfer functions involving differentiation orders multiples of 0.5. Moreover, in electrochemistry for example, it is proven that the diffusion of charges in acid batteries is governed by Randles models (Rodrigues *et al.*, 2000; Sabatier *et al.*, 2006) that involve a half order integrator. In semi-infinite thermal systems, (Battaglia *et al.*, 2001) have shown that the exact solution of heat equation links thermal flux to a half order derivative of surface temperature on which the flux is applied. In rheology, stress in viscoelastic materials is proportional to a non integer derivative of deformation (Heymans and Bauwens, 1994). Foucault currents inside rotor bars in induction machines obey to diffusive phenomena modeled by fractional operators (Benchellal *et al.*, 2004).

Time-domain system identification using fractional models was initiated in the late nineties and the beginning of this century (Mathieu *et al.*, 1995; Cois *et al.*, 2001; Aoun *et al.*, 2007; Malti *et al.*, 2008; Narang *et al.*, 2011; Gabano and Poinot, 2011*a*). (Malti *et al.*, 2009) and (Gabano and Poinot, 2011*b*), have identified thermal systems using fractional models in an output error context. An advantage of continuous-time modeling is in straightforwardly estimating physical parameters even when the physics reveals systems governed by fractional differential equations. In this context and based on a fractional model, (Gabano and Poinot, 2011*a*) have successfully estimated thermal conductivity and diffusivity in a homogeneous medium.

Instrumental variable methods for system identification have been developed for decades (see e.g. (Young, 1981; Young, 2011; Söderström and Stoica, 1983; Söderström, 2012)). **R**efined **I**nstrumental **V**ariables for **C**ontinuous system (*rivc*), and its simplified version (*srivc*) when the additive measurement noise is white, were developed in (Young and Jakeman, 1980; Young, 1981; Young, 2002). When the model lies in the same class as the true system, the *rivc* method and consequently the *srivc* can be interpreted in optimal statistical terms yielding consistent estimates with minimum variance. Both methods

use an iterative-adaptive prefiltering based on a quasi-optimal statistical solution and provide an estimation of the parametric error covariance matrix. Recent developments aimed at extending the *srivc* method to handle multi-input transfer function model identification (Garnier *et al.*, 2007), hybrid Box-Jenkins model identification (Young *et al.*, 2008), closed-loop identification (Gilson *et al.*, 2008), irregularly sampled data (Wang *et al.*, 2009) and linear parameter varying model identification (Laurain *et al.*, 2011).

One very important aspect of fractional differential equation modeling is the determination of the differentiation orders. In system identification with rational models, where only the coefficients are estimated, the model order remains unchanged. Estimating coefficients and differentiation orders in fractional models is not a trivial problem: in an iterative algorithm, it indeed changes the model order at each iteration. In this paper, the *srivc* algorithm is first extended to fractional models as the *srivcf* for linear coefficient estimation, by assuming all differentiation orders known. Section 4 of the paper aims at proposing a gradient-based approach which combines simultaneously the coefficients and the differentiation order estimation. Two variants are proposed. In the first one, all differentiation orders are set as integral multiples of a commensurate order which is estimated. Then, this commensurability constraint is released and all differentiation orders are further adjusted. Finally, these algorithms are applied to the identification of a real thermal process, where the transfer function linking the temperature to the heat density flux is modeled by a fractional model.

The main advantage of fractional models is that they allow modeling infinite dimensional systems with a reduced number of parameters. Since it is always possible to find a high order rational model equivalent to a low dimension fractional model (Oustaloup, 1995), the use of high dimension fractional models is of limited interest. That is why, the proposed algorithms will be tested on low dimension fractional systems. The fractional behavior of a given system can be recognized either from a preliminary experiment on a step response when the convergence time is polynomial or from prior knowledge when theoretical models reveal the presence of fractional derivatives or integrals.

Mathematical background

A fractional mathematical model is based on a differential equation:

$$y(t) + a_1 p^{\alpha_1} y(t) + \dots + a_N p^{\alpha_N} y(t) = b_0 p^{\beta_0} u(t) + b_1 p^{\beta_1} u(t) + \dots + b_M p^{\beta_M} u(t), \quad (1)$$

where $u(t)$ is the system input, $y(t)$ is the noise-free system output, p is the differentiation operator

$p = \frac{d}{dt}$, and so $p^\mu u(t)$ denotes the μ -th time-derivative of the continuous-time signal $u(t)$, $(a_j, b_i) \in \mathbb{R}^2 \forall i \in \{0, 1, \dots, m\}$ and $\forall j \in \{1, \dots, n\}$, and the differentiation orders

$$0 < \alpha_1 < \alpha_2 < \dots < \alpha_N, \quad 0 \leq \beta_0 < \beta_1 < \dots < \beta_M, \quad (2)$$

allowed to be non-integer positive numbers, are ordered for identifiability purposes. The concept of differentiation to an arbitrary order (non-integer) γ , with $\gamma \in \mathbb{R}$ was defined by Grünwald-Letnikov (see e.g. (Podlubny, 1999, chapter 2)), which results from the generalization of integer order derivatives:

$$p^\gamma x(t) = \lim_{h \rightarrow 0} \frac{1}{h^\gamma} \sum_{k=0}^{\lfloor \frac{t}{h} \rfloor} (-1)^k \binom{\gamma}{k} x(t - kh), \quad (3)$$

where $\lfloor \cdot \rfloor$ stands for the floor operator, and $\binom{\gamma}{k}$ stands for the Newton binomial coefficient generalized to real numbers:

$$\binom{\gamma}{k} = \frac{\Gamma(\gamma + 1)}{\Gamma(k + 1)\Gamma(\gamma - k + 1)} = \frac{\gamma(\gamma - 1)\dots(\gamma - k + 1)}{k!}.$$

Since $\binom{\gamma}{k} = 0$ when $(\gamma - k) \in \mathbb{N}$, the definition (3) reduces, for $\gamma = 1$, to the standard derivative definition:

$$p x(t) = \lim_{h \rightarrow 0} \frac{x(t) - x(t - h)}{h}. \quad (4)$$

The non integer derivative of a function has a global characterization, as the whole past is taken into account.

For numerical evaluation of the fractional derivative, the h parameter in (3) is replaced by the sampling period and hence the limit is dropped:

$$p^\gamma x(t) = \frac{1}{h^\gamma} \sum_{k=0}^{\lfloor \frac{t}{h} \rfloor} (-1)^k \binom{\gamma}{k} x(t - kh) + O(h). \quad (5)$$

In doing so, the error terms are proportional to the sampling period (Podlubny, 1999, section 7.4). Consequently, the sampling period should be small enough for the approximation error to be negligible.

The Laplace transform of $p^\gamma x(t)$ when $x(t)$ is relaxed at $t = 0$ ($x(t)$ equals 0 for all $t < 0$), is given by (Podlubny, 1999, section 2.8.4):

$$\mathcal{L}\{p^\gamma x(t)\} = s^\gamma X(s). \quad (6)$$

This result is coherent with the classical case when γ is an integer. Consequently, it is easy to define a symbolic representation of the fractional dynamic system

governed by (1) using the transfer function:

$$G(s) = \frac{B(s)}{A(s)} = \frac{\sum_{i=0}^M b_i s^{\beta_i}}{1 + \sum_{j=1}^N a_j s^{\alpha_j}}. \quad (7)$$

Moreover, if $G(s)$ is commensurate of order γ , then $G(s)$ can be rewritten as:

$$G(s) = \frac{\sum_{i=0}^m \tilde{b}_i s^{i\gamma}}{1 + \sum_{j=1}^n \tilde{a}_j s^{j\gamma}}, \quad (8)$$

where $m = \frac{\beta_M}{\gamma}$ and $n = \frac{\alpha_N}{\gamma}$ are integers and $\forall i' \in \{0, 1, \dots, m\}, \forall j' \in \{1, \dots, n\}$:

$$\begin{cases} \tilde{b}_{i'} = b_i \text{ if } \exists i \in \{0, 1, \dots, M\} \text{ such that } i'\gamma = \beta_i \\ \tilde{b}_{i'} = 0 \text{ otherwise} \\ \tilde{a}_{j'} = a_j \text{ if } \exists j \in \{1, \dots, N\} \text{ such that } j'\gamma = \alpha_j \\ \tilde{a}_{j'} = 0 \text{ otherwise.} \end{cases} \quad (9)$$

In rational transfer functions, γ equals 1 and usually numerator α_N and denominator β_M orders are both fixed, then all coefficients $b_i, i = 0, \dots, M$ and $a_j, j = 1, \dots, N$ are estimated. Generally, no care is taken to check whether any intermediate coefficient, as in (9), equals zero.

(Khemane *et al.*, 2012) have proven that the fractional transfer function representation (7) is unique, provided that all the differentiation orders are ordered as in (2) and provided that all the coefficients $a_1, \dots, a_N, b_0, \dots, b_M$ are non zero. This property is interesting in system identification context as it shows that the global minimum is unique when the fractional model is in the same class as the true system.

Time-domain simulation of fractional systems is an extensively studied topic in the literature (Oustaloup, 1995; Aoun *et al.*, 2004; Krishna, 2011). System identification algorithms proposed in this paper may be used with any time-domain simulation algorithm of fractional systems, at the discretion of the user. Fractional systems should be correctly simulated with negligible simulation errors, in order to be able to consistently estimate the fractional model parameters.

Stability of fractional systems was treated in different contexts (linear, non linear, commensurate, non commensurate, time-variant, time invariant, delayed, non delayed, analytical, numerical) by different authors as presented in the state of the art by (Sabatier *et al.*, 2010). The most well-known stability criterion for commensurate fractional systems is (Matignon, 1998, theorem

1) which allows to check the stability of a commensurate fractional system through the location of its s^γ -poles. Matignon's stability theorem was initially proven for commensurate orders between 0 and 1, extended to commensurate orders between 1 and 2 in (Moze and Sabatier, 2005). A proof of instability when $\gamma \geq 2$ for all s^γ -poles is provided in (Malti *et al.*, 2011b).

Stability theorem ((Matignon, 1998), extended): A commensurate transfer function, with a commensurate order γ as in (8), $G(s) = \frac{T(s^\gamma)}{R(s^\gamma)}$, where T and R are coprime polynomials, is BIBO (bounded input bounded output) stable if and only if :

$$0 < \gamma < 2 \quad (10)$$

and

$$|\arg(s_k)| > \gamma \frac{\pi}{2}, \quad \forall s_k \in \mathbb{C} / R(s_k) = 0. \quad (11)$$

The stability conditions are used later when estimating the commensurate order, which must be in the interval (10) for stable systems. Additionally, analytical formulae have been provided in (Malti *et al.*, 2011a) for computing the \mathcal{H}_2 -norm of fractional commensurate transfer functions.

2 Problem formulation

It is assumed that the input $u(t)$ and the noise-free output $y(t)$ are related by the constant coefficient differential equation (1), which can also be written in the following compact model form:

$$G(p) = \frac{B(p)}{A(p)} = \frac{\sum_{i=0}^M b_i p^{\beta_i}}{1 + \sum_{j=1}^N a_j p^{\alpha_j}}. \quad (12)$$

$B(p)$ and $A(p)$ are assumed to be coprime, and the system asymptotically stable.

The data are collected at regular time-instants from $t = 0$ to $t = T_{final}$ (K samples), with a sampling period, h , small enough so that the approximation errors while computing numerically the fractional derivatives, as in (5), are negligible. Moreover, the number of samples is assumed to be large enough to guarantee convergence of the estimated parameters to the true ones. The quasi-stationary input signal $\{u(t), 0 \leq t \leq T_{final}\}$ applied to the system is persistently exciting, and gives rise to an output signal $\{y(t), 0 \leq t \leq T_{final}\}$. The noise-free output y is supposed to be corrupted by an additive white measurement noise ξ , normally distributed with a zero

mean and σ^2 variance, considered at discrete instants. The complete equation can be written in the form:

$$\begin{cases} y(t) = G(p)u(t) \\ y^*(t_k) = y(t_k) + \xi(t_k), \end{cases} \quad (13)$$

where $y^*(t_k)$ is the sampled value of the unobserved noisy output $y(t)$. Given the discrete-time sampled nature of the data, the usual assumption is that a discrete-time noise $\xi(t_k)$ is associated with the sampled data (see e.g. (Young *et al.*, 2008)). It is easier to consider white noise in the discrete-time case, since the concept of continuous-time white noise is a non-trivial extension of the discrete-time case.

When the model (12) is used, the parameter vector

$$\theta = \begin{bmatrix} \rho \\ \mu \end{bmatrix} \quad (14)$$

is composed of a vector of $N + M + 1$ transfer function coefficients,

$$\rho = [b_0, b_1, \dots, b_M, a_1, \dots, a_N]^T, \quad (15)$$

and additionally a vector of $N + M + 1$ ordered differentiation orders, as in (2),

$$\mu = [\beta_0, \dots, \beta_M, \alpha_1, \dots, \alpha_N]^T. \quad (16)$$

Estimating all differentiation orders necessitates to determine $N + M + 1$ terms along with the $N + M + 1$ transfer function coefficients. Moreover, if N and/or M are high, the number of local minima may increase and the nonlinear optimization algorithms, with high complexity, may fail to converge to the global minimum. To reduce the number of parameters, a commensurate model can be chosen, instead of (12):

$$G(p) = \frac{\sum_{i=0}^M b_i p^{i\gamma}}{1 + \sum_{j=1}^N a_j p^{j\gamma}}, \quad (17)$$

where only one differentiation order and hence $N + M + 2$ parameters are estimated. The parameter vector θ in (14) is then reduced to the $N + M + 1$ transfer function coefficients as in (15) added to a single commensurate differentiation order:

$$\mu = \gamma. \quad (18)$$

In this case, the searching domain of the commensurate order is held for $\gamma \in (0, 2)$ as specified in the extended Matignon's theorem.

3 Simplified refined instrumental variable for continuous-time fractional models

In this section all differentiation orders are assumed to be known and fractional transfer function coefficients are estimated. The iterative *srivc* algorithm, known to deliver optimal estimates when the additive measurement noise is white, is first extended to fractional systems and then validated through a Monte Carlo analysis.

The *srivc* algorithm derives from the *sriv* algorithm for discrete-time transfer function model identification. This was evolved by converting the maximum likelihood (ML) estimation equations to a pseudo-linear form involving optimal prefilters (see e.g. (Young, 2011)). A similar analysis can be utilized in the present situation because the problem is very similar, in both algebraic and statistical terms. For space constraint, a simpler development of the *srivc* algorithm for fractional transfer function model identification is presented. Following the *Prediction Error Method*, a suitable error function ε is given by the *Output Error*:

$$\varepsilon(t, \rho) = y^*(t) - \frac{B(p, \rho)}{A(p, \rho)}u(t), \quad (19)$$

which can be rewritten as:

$$\begin{aligned} \varepsilon(t, \rho) &= A(p, \rho) \left(\frac{1}{A(p, \rho)} y^*(t) \right) - B(p, \rho) \left(\frac{1}{A(p, \rho)} u(t) \right) \\ &= A(p, \rho) y_f^*(t) - B(p, \rho) u_f(t) \end{aligned} \quad (20)$$

$$\text{where } y_f^* = \frac{1}{A(p, \rho)} y^*(t) \text{ and } u_f(t) = \frac{1}{A(p, \rho)} u(t).$$

Hence, as for rational systems, a linear transformation (low-pass filter) is applied to input and output signals instead of a direct differentiation. As shown in (20), the filter required to convert the output error into an equation error, which turns out to lead to optimal Instrumental Variable (IV) estimates, takes the form (see e.g. (Young, 2011)):

$$F^{\text{opt}}(p) = \frac{1}{A(p, \rho)}. \quad (22)$$

In practical cases, $A(p, \rho)$ being unknown, an estimation $\hat{A}(p, \hat{\rho}^{\text{iter}})$ is computed iteratively:

$$F^{\text{iter}}(p) = \frac{1}{\hat{A}(p, \hat{\rho}^{\text{iter}})} = \frac{1}{1 + \sum_{j=1}^N \hat{a}_j p^{\alpha_j}}, \quad (23)$$

where $\text{iter} = 1, 2, \dots$ stands for the iteration number and \hat{a}_j the estimated a_j coefficient at iteration iter .

The noise-free output estimates required to build the instrument are obtained from an auxiliary model, with transfer function coefficients computed iteratively:

$$y_{iv}^{\text{iter}}(t) = \frac{\hat{B}(p, \hat{\rho}^{\text{iter}})}{\hat{A}(p, \hat{\rho}^{\text{iter}})} u(t). \quad (24)$$

The filtered derivatives of the input, the output, and the instruments are computed respectively with:

$$p^{\beta_i} u_f(t) = F_{\beta_i}^{\text{iter}}(p) u(t), \quad (25)$$

$$p^{\alpha_j} y_f^*(t) = F_{\alpha_j}^{\text{iter}}(p) y_f^*(t), \quad (26)$$

$$p^{\alpha_j} y_{iv,f}^{\text{iter}}(t) = F_{\alpha_j}^{\text{iter}}(p) y_{iv}^{\text{iter}}(t), \quad (27)$$

where

$$F_{\sigma}^{\text{iter}}(p) = p^{\sigma} F^{\text{iter}}(p) = \frac{p^{\sigma}}{\hat{A}(p, \hat{\rho}^{\text{iter}})}, \quad (28)$$

and gathered in the regression $\varphi_f(t)$ and the instrumental variable $\varphi_{iv,f}^{\text{iter}}(t)$ vectors :

$$\varphi_f(t) = \begin{bmatrix} p^{\beta_0} u_f(t), \dots, p^{\beta_M} u_f(t), \\ -p^{\alpha_1} y_f^*(t), \dots, -p^{\alpha_N} y_f^*(t) \end{bmatrix}^T, \quad (29)$$

$$\varphi_{iv,f}^{\text{iter}}(t) = \begin{bmatrix} p^{\beta_0} u_f(t), \dots, p^{\beta_M} u_f(t), \\ -p^{\alpha_1} y_{iv,f}^{\text{iter}}(t), \dots, -p^{\alpha_N} y_{iv,f}^{\text{iter}}(t) \end{bmatrix}^T. \quad (30)$$

From different time-instants, the **srivcf** optimization problem can be stated as

$$\hat{\rho}^{\text{iter}+1} = \arg \min_{\rho} \left\| [\Phi_{iv,f}^{\text{iter}} \Phi_f^T] \rho - [\Phi_{iv,f}^{\text{iter}} \mathbf{Y}_f^*] \right\|^2, \quad (31)$$

with

$$\Phi_{iv,f}^{\text{iter}} = [\varphi_{iv,f}^{\text{iter}}(0) \quad \dots \quad \varphi_{iv,f}^{\text{iter}}(T_{final})]^T, \quad (32)$$

$$\Phi_f = [\varphi_f(0) \quad \dots \quad \varphi_f(T_{final})]^T, \quad (33)$$

$$\mathbf{Y}_f^* = [y_f^*(0) \quad \dots \quad y_f^*(T_{final})]^T. \quad (34)$$

The **srivcf**-based solution is hence given at each iteration by:

$$\hat{\rho}^{\text{iter}+1} = (\Phi_{iv,f}^{\text{iter}} \Phi_f^T)^{-1} \Phi_{iv,f}^{\text{iter}} \mathbf{Y}_f^*, \quad (35)$$

and the algorithm is iterated until convergence, when

$$\max_j \left| \frac{\hat{\rho}_j^{\text{iter}+1} - \hat{\rho}_j^{\text{iter}}}{\hat{\rho}_j^{\text{iter}+1}} \right| < \epsilon, \quad (36)$$

where $\hat{\rho}_j^{\text{iter}}$ is the j -th element of the estimated parameter vector at the iteration iter . The **sriv** (Young and Jakeman, 1980; Young, 1981) and **srivc** (Young, 2002; Garnier and Wang, 2008) algorithms have been successfully used for more than three decades in various contexts. Moreover, it has recently been established in the rational continuous-time case in (Liu *et al.*, 2011, theorem 1) that the **srivc** algorithm converges asymptotically in one iteration under some *mild assumptions*¹. In the fractional case, the convergence of the **srivc** algorithm follows exactly the same scheme under the same assumptions.

The iterative **srivcf** algorithm can be initialized using the least squares-based state-variable filter (**lssvf**) or the suboptimal instrumental variable **svf** (**ivsvf**) method for fractional models, as detailed in (Cois *et al.*, 2001; Malti *et al.*, 2008). In both methods a non-optimal filter is used, instead of (28). In the case of the state-variable filter, it takes the form:

$$F_{\sigma}^{\text{svf}}(p) = \frac{p^{\sigma}}{\left(1 + \frac{p}{\omega_c}\right)^{N_c}}, \quad (37)$$

where two user parameters need to be *a priori* chosen: the integer order $N_c > \alpha_N$ usually set as the integer above the system order α_N , i.e. $N_c = \lfloor \alpha_N \rfloor + 1$, and the cut-off frequency ω_c usually chosen to emphasize the frequency band of interest, which is a little bit larger than the frequency bandwidth of the system to be identified (Garnier *et al.*, 2003). The basic idea in using the state-variable filter is to obtain the behavior of a differentiator in low frequencies and to filter noise in high frequencies.

As in the rational case, if the quasi-stationary input signal is persistently exciting and the noise signal is white with zero-mean, the **srivcf** estimator is asymptotically unbiased thanks to the **iv** mechanism. When the system is in the true model class, the **iv** estimate provides a consistent estimate under the following two conditions

$$\begin{cases} \mathbb{E} \left[\varphi_{iv,f}^{\text{iter}}(t_k) \varphi_f^T(t_k) \right] \text{ is non-singular,} \\ \mathbb{E} \left[\varphi_{iv,f}^{\text{iter}}(t_k) \xi(t_k) \right] = 0. \end{cases} \quad (38)$$

Additionally, when the noise signal is normally distributed, an estimation of the asymptotic covariance matrix of the estimation error associated with the estimate of $\hat{\rho}$ in (35) follows straightforwardly from the rational continuous-time case (Young *et al.*, 2008) by replacing integer derivatives by fractional derivatives and yields:

$$\hat{\mathbf{P}}_{\rho} = \hat{\sigma}^2 \left(\Phi_{iv,f} \Phi_{iv,f}^T \right)^{-1}, \quad (39)$$

¹ as qualified by the authors of the paper.

where $\hat{\sigma}^2$ is the empirical estimation of the noise variance, and $\Phi_{iv,f}$ corresponds to $\Phi_{iv,f}^{\text{iter}}$ computed at the last iteration.

In any parameter estimation problem, the estimator performance can be assessed by comparing the estimator covariance matrix to the best possible accuracy given by the Cramér-Rao lower-bound (CRB) (Kay, 1993; Stojica and Moses, 2005; Young *et al.*, 2008). Under a white Gaussian noise and sufficient persistency of the input signal assumptions, the CRB is computable and often tight, meaning that there exists an iv estimator that asymptotically reaches the CRB (Wellstead, 1978). In this regard, a direct extension of the rational continuous-time case (Young *et al.*, 2008)², by replacing integer derivatives by fractional derivatives, enables computing the CRB:

$$\mathbf{P}_\rho^{\text{opt}} = \frac{\sigma^2}{N} \left(\mathbb{E} \left[\varphi_{iv,f}^{\text{opt}}(t_k) \varphi_{iv,f}^{\text{opt}}(t_k)^T \right] \right)^{-1}, \quad (40)$$

obtained with the true design variables and the true noise variance σ^2 . The instrumental vector $\varphi_{iv,f}^{\text{opt}}(t_k)$ is the optimally noise-free pre-filtered iv vector, with $F^{\text{opt}}(p)$ defined in (22) and

$$\varphi_{iv,f}^{\text{opt}}(t_k) = F^{\text{opt}}(p) \begin{bmatrix} p^{\beta_0} u(t_k), \dots, p^{\beta_M} u(t_k), \\ -p^{\alpha_1} y(t_k), \dots, -p^{\alpha_N} y(t_k) \end{bmatrix}^T. \quad (41)$$

It is assumed that the **srivcf** algorithm converges in the sense that $\hat{\rho} \rightarrow \rho$, then the covariance matrix (39) will tend in probability to (40).

Example 1 – **srivcf** algorithm

To illustrate the efficacy of this algorithm, a pseudo random binary sequence (prbs) is applied to the following data-generating system:

$$G_0(s) = \frac{0.5}{0.5s^3 + 1.5s^{1.5} + 1} = \frac{1}{(s^{1.5} + 1)(s^{1.5} + 2)}. \quad (42)$$

The output, $y(t)$, is corrupted at the sampling instants (13), by an additive Gaussian white noise $\xi(t_k)$ with zero mean and a signal to noise ratio (SNR= $10 \log \frac{\|y(t_k)\|^2}{\|\xi(t_k)\|^2}$), with $\|\cdot\|^2$ standing for the squared Euclidean norm) of 5dB. Input and output data, sampled with a rate of $h = 0.02$ s, are plotted in Fig. 1. The simulation algorithm used in this example is based on the numerical discretization of the Grünwald operator (5).

² The rational continuous-time case in (Young *et al.*, 2008) results from an original justification in (Pierce, 1972).

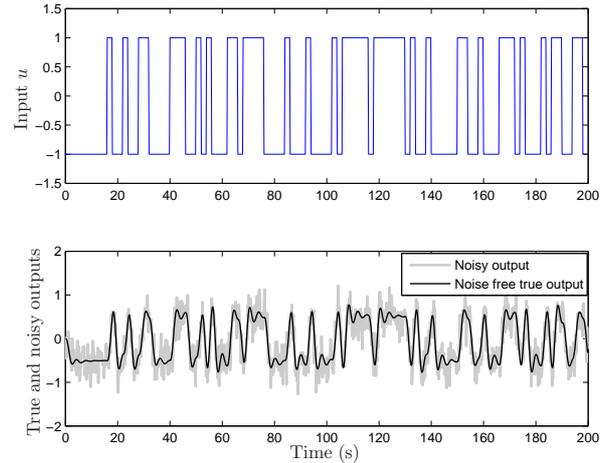


Fig. 1. Input, noise-free and noisy outputs, used as identification data, zoomed between $t = 0$ and 200s, identification data going up to 600s (SNR = 5dB and $h = 0.02$ s)

	true	ivsvf		srivcf	
	ρ	$\bar{\hat{\rho}}$	$\hat{\sigma}_\theta$	$\bar{\hat{\rho}}$	$\hat{\sigma}_\theta$
b₀	0.500	0.501	0.016	0.500	0.007
a₂	0.500	0.487	0.148	0.500	0.010
a₁	1.500	1.500	0.020	1.500	0.012

Table 1

Monte Carlo simulation results for SNR = 5 dB in example 1. $\bar{\hat{\rho}}$ denotes the mean and $\hat{\sigma}_\theta$ the standard deviation of the parameter estimates. For the **ivsvf** method, the filter used is given by (37), ω_c is set to 100 rad/s and N_c to 4.

The model structure is chosen as:

$$G(s) = \frac{b_0}{a_2 s^{2\gamma} + a_1 s^\gamma + 1}, \quad (43)$$

and all parameters are estimated, except the commensurate order, set to $\gamma = 1.5$, in order to be in the same model class as the true system. To study the statistical properties of the estimators, a Monte Carlo (MC) simulation of 500 runs is implemented with different noise realizations. For each MC run, the simulated system is identified with the **ivsvf** method (Cois *et al.*, 2001; Malti *et al.*, 2008), and the **srivcf** method. The results of the MC simulation are presented in Table 1 where the estimated parameter mean and standard deviation are given for each method. In such a high-level noisy context, the **srivcf** and **ivsvf** algorithms are both asymptotically unbiased thanks to the iv mechanism. The **srivcf** parameter estimates have, as expected, a smaller variance than the suboptimal **ivsvf** parameter estimates. This is further confirmed by computing the following Euclidean norms of the covariance matrices, obtained for a single

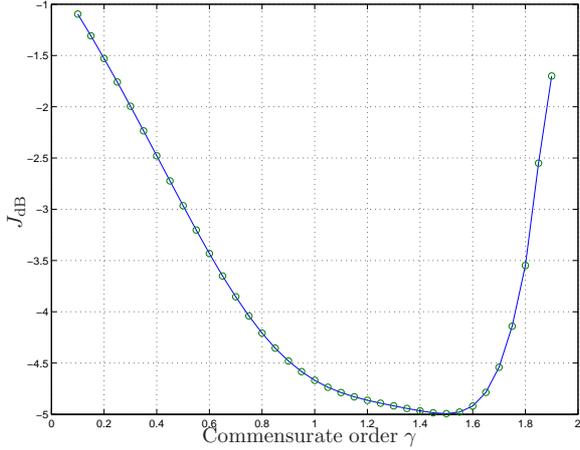


Fig. 2. Output error ℓ_2 -norm versus commensurate differentiation order in example 1.

run in the Monte Carlo simulation:

$$\left\| \hat{\mathbf{P}}_{\rho}^{ivsvf} \right\|_2 = 2.81 \times 10^{-4}, \quad (44)$$

$$\left\| \hat{\mathbf{P}}_{\rho}^{srivcf} \right\|_2 = 1.88 \times 10^{-4}. \quad (45)$$

Since, the true parameters, the noise-free output, and the noise variance are known, the CRB \mathbf{P}_{ρ}^{opt} can be computed:

$$\left\| \mathbf{P}_{\rho}^{opt} \right\|_2 = 1.79 \times 10^{-4}.$$

As a consequence, the covariance matrix $\left\| \hat{\mathbf{P}}_{\rho}^{srivcf} \right\|_2$ is closer to the Cramér-Rao lower bound as compared to $\left\| \hat{\mathbf{P}}_{\rho}^{ivsvf} \right\|_2$:

$$\left\| \hat{\mathbf{P}}_{\rho}^{ivsvf} \right\|_2 > \left\| \hat{\mathbf{P}}_{\rho}^{srivcf} \right\|_2 \geq \left\| \hat{\mathbf{P}}_{\rho}^{opt} \right\|_2. \quad (46)$$

Up to now, all differentiation orders have been as known. However, if differentiation orders are unknown, then a computation of the ℓ_2 -norm (in dB) of the normalized output error,

$$J_{dB} = 10 \log \left(\frac{\|\varepsilon(t, \hat{\theta}_{srivcf})\|^2}{\|y(t)\|^2} \right), \quad (47)$$

where $\varepsilon(t, \hat{\theta}_{srivcf}) = y^*(t) - G(p, \hat{\theta}_{srivcf})u(t)$, versus the commensurate order shows, in Fig. 2, that the criterion depends on the differentiation order and that its optimum is obtained at the true value. This normalized criterion equals ($-\text{SNR} = -5\text{dB}$) at the optimum when there is no modeling error. Otherwise, it allows to quantify the modeling error which equals, in this example, 0.4dB for a rational model with $\gamma = 1$.

4 Differentiation order estimation

When differentiation orders are unknown, as it is often the case in practice, it is helpful to consider order estimation along with transfer function coefficient estimation. An algorithm is proposed to identify the parameters in two stages. This algorithm, named order-optimization-*srivcf* or *oosrivcf* for short, uses the *srivcf* algorithm for coefficient estimation and is combined with a gradient-based algorithm for differentiation order optimization. Such two-stage algorithms have been used in (Spinelli *et al.*, 2006; Bai, 1998; Mukhopahyay *et al.*, 1991; Rao *et al.*, 1984; Rao and Unbehauen, 2006). Although they do not have a proof of convergence, it has been noted in the aforementioned references that they do often converge to a minimum. Two variants of order optimization algorithm are presented. Either all differentiation orders are set as integral multiples of a commensurate order, as in the model (17), and the best commensurate order is computed in the interval (0, 2), or all differentiation orders, as in (12), are estimated. The former variant is used as a good initial hit for the latter variant. The estimation problem is formulated as a minimization problem of the ℓ_2 -norm:

$$J = \frac{1}{2} \|\varepsilon(t, \theta)\|^2, \quad (48)$$

of the output error

$$\varepsilon(t, \theta) = y^*(t) - G(p, \theta)u(t), \quad (49)$$

with respect to the vector μ , as in (16) or (18).

A Gauss-Newton algorithm (see e.g. (Ljung, 1999)) is used for the estimation of the differentiation order(s). This iterative algorithm allows to compute the differentiation order(s) μ^{iter} at iteration iter with respect to $\mu^{\text{iter}-1}$ and a correction factor:

$$\mu^{\text{iter}} = \mu^{\text{iter}-1} - \lambda H^{-1} \frac{\partial J}{\partial \mu}. \quad (50)$$

Here, the gradient $\frac{\partial J}{\partial \mu}$ and the approximated Hessian H , respectively given by

$$\frac{\partial J}{\partial \mu} = \frac{\partial \varepsilon^T}{\partial \mu} \varepsilon \quad \text{and} \quad H = \frac{\partial \varepsilon^T}{\partial \mu} \frac{\partial \varepsilon}{\partial \mu} \quad (51)$$

are evaluated for $\mu = \mu^{\text{iter}-1}$.

In the Gauss-Newton method, the advantage of computing the approximated Hessian is that the second derivative of ε is not required; hence it is less time consuming. Moreover, it allows to obtain always a positive semi-definite matrix which is a necessary convergence condition. If the criterion is convex, the algorithm reaches the

global minimum. Otherwise, only a convergence to a local minimum is guaranteed. The λ parameter, usually set to , may be used to adjust the step in case of slow convergence or oscillations around a minimum.

The error sensitivity function, required for computing the gradient and the Hessian in (51), is given by:

$$\frac{\partial \varepsilon}{\partial \mu} = \left[-\frac{\partial \hat{y}}{\partial \beta_0}, \dots, -\frac{\partial \hat{y}}{\partial \beta_M}, -\frac{\partial \hat{y}}{\partial \alpha_1}, \dots, -\frac{\partial \hat{y}}{\partial \alpha_N} \right]^T, \quad (52)$$

when all differentiation orders are estimated, and by:

$$\frac{\partial \varepsilon}{\partial \mu} = \frac{\partial \varepsilon}{\partial \gamma} = -\frac{\partial \hat{y}}{\partial \gamma}, \quad (53)$$

when the commensurate order is estimated. The output sensitivity functions with respect to all differentiation orders are given by:

$$\frac{\partial \hat{y}}{\partial \beta_i} = \ln(p) \frac{\hat{b}_i p^{\hat{\beta}_i}}{1 + \sum_{j=1}^N \hat{a}_j p^{\hat{\alpha}_j}} u(t), \quad \forall i = 0, \dots, M, \quad (54)$$

$$\frac{\partial \hat{y}}{\partial \alpha_k} = \ln(p) \hat{a}_k p^{\hat{\alpha}_k} \frac{\sum_{i=0}^M \hat{b}_i p^{\hat{\beta}_i}}{\left(1 + \sum_{j=1}^N \hat{a}_j p^{\hat{\alpha}_j}\right)^2} u(t), \quad \forall k = 1, \dots, N. \quad (55)$$

The output sensitivity function with respect to the commensurate differentiation order is given by:

$$\frac{\partial \hat{y}}{\partial \gamma} = \ln(p) \frac{\sum_{i=0}^M i \hat{b}_i p^{i\hat{\gamma}} + \sum_{i=0}^M \sum_{j=1}^N (i-j) \hat{b}_i \hat{a}_j p^{(i+j)\hat{\gamma}}}{\left(1 + \sum_{j=1}^N \hat{a}_j p^{j\hat{\gamma}}\right)^2} u(t). \quad (56)$$

Sensitivity functions $\frac{\partial \hat{y}}{\partial \beta_i}$, $\frac{\partial \hat{y}}{\partial \alpha_k}$, and $\frac{\partial \hat{y}}{\partial \mu}$ depend on $\ln(p)$ (log of the differential operator) which is not trivial to simulate in the time-domain. Thus, an alternative method consists of computing $\frac{\partial \hat{y}}{\partial \mu}$ numerically, using for example the central difference method.

By assuming that the **oosrivcf** algorithm converges in the sense that $\hat{\theta} \rightarrow \theta$, then an estimate of the covariance matrix can be obtained on the basis of the Gauss-Newton algorithm (Ljung, 1999, Theorem 9.1) by:

$$\hat{\mathbf{P}}_{\theta} = \hat{\sigma}^2 \mathcal{H}^{-1}, \quad (57)$$

where $\hat{\sigma}^2$ is, as previously, the empirical estimate of noise variance and \mathcal{H} is the approximate Hessian computed

towards all the estimated parameters:

$$\mathcal{H} = \frac{\partial \varepsilon^T}{\partial \theta} \frac{\partial \varepsilon}{\partial \theta} = \frac{\partial \varepsilon^T}{\partial [\rho^T \mu^T]^T} \frac{\partial \varepsilon}{\partial [\rho^T \mu^T]^T}, \quad (58)$$

with $\frac{\partial \varepsilon}{\partial \mu}$ defined in (52) or (53) depending on whether all differentiation orders are estimated or only the commensurate order and

$$\frac{\partial \varepsilon}{\partial \rho} = -\frac{\partial \hat{y}}{\partial \rho} = -\left[\frac{\partial \hat{y}}{\partial b_0}, \dots, \frac{\partial \hat{y}}{\partial b_M}, \frac{\partial \hat{y}}{\partial a_1}, \dots, \frac{\partial \hat{y}}{\partial a_N} \right]^T, \quad (59)$$

where

$$\frac{\partial \hat{y}}{\partial b_i} = \frac{\hat{p}^{\hat{\beta}_i}}{1 + \sum_{j=1}^N \hat{a}_j p^{\hat{\alpha}_j}} u(t), \quad \forall i = 0, \dots, M, \quad (60)$$

$$\frac{\partial \hat{y}}{\partial a_k} = -\frac{\sum_{i=0}^M \hat{b}_i p^{\hat{\beta}_i + \hat{\alpha}_k}}{\left(1 + \sum_{j=1}^N \hat{a}_j p^{\hat{\alpha}_j}\right)^2} u(t), \quad \forall k = 1, \dots, N. \quad (61)$$

Summary of the oosrivcf algorithm

Both variants of the **oosrivcf** algorithm are summarized here. Either the commensurate order is estimated when μ is defined as in (18) or all orders are estimated when μ is defined as in (16).

Step 1 Initialization

iter = 0

Initialize μ^0 and compute ρ^0 with the **srivcf** method.

From $\theta^0 = [\rho^0, \mu^0]$, compute $J(\theta^0)$.

Step 2 Gauss-Newton optimization

do

Initialize $\lambda = \Lambda$ (usually to 1)

do

(i) Refine the order estimate:

$$\mu^{\text{iter}+1} = \mu^{\text{iter}} - \lambda \left[H^{-1} \frac{\partial J}{\partial \mu} \right] \Big|_{\mu^{\text{iter}}}.$$

(ii) Compute ρ^{iter} using the **srivcf** algorithm.

(iii) Evaluate the error criterion: $J(\theta^{\text{iter}+1})$.

(iv) $\lambda = \lambda/2$

while $J(\theta^{\text{iter}+1}) > J(\theta^{\text{iter}})$

iter = *iter* + 1

while $\max_l \left| \frac{\mu_l^{\text{iter}} - \mu_l^{\text{iter}-1}}{\mu_l^{\text{iter}}} \right| > \epsilon_2$

where μ_l^{iter} corresponds to l -th element of the order vector μ^{iter} at iteration *iter*. l equals one in case of commensurate order estimation.

Step 3 Parametric error estimation

After convergence, the covariance matrix $\hat{\mathbf{P}}_\theta$ associated to the estimate $\hat{\theta}$ can be estimated from (57).

In case all differentiation orders are estimated, it is highly recommended to estimate a commensurate model first, and then to release the commensurability constraint in order to estimate all differentiation orders. The commensurate model constitutes generally a good initial hit when estimating all differentiation orders.

An alternative algorithm to the *oosrivcf* is the Gauss-Newton algorithm (or the Levenberg-Marquardt variant) with a joint estimation of coefficients and differentiation orders (Malti *et al.*, 2008). The main advantage of the *oosrivcf* algorithm is that it simplifies considerably the initialization problem, as it requires only initializing the commensurate order, whereas the Gauss-Newton algorithm requires initializing additionally all the $N + M + 1$ coefficients in (15). Consequently, the *oosrivcf* algorithm is more likely to converge to the global minimum as compared to the Gauss-Newton algorithm where the number of local minima increases with the increase of N and/or M . Another advantage of the *oosrivcf* algorithm is its low computational complexity, as it requires computing only one sensitivity function, in case of commensurate order estimation, whereas the Gauss-Newton algorithm, with a joint estimation of coefficients and differentiation orders, requires computing additionally $N + M + 1$ sensitivity functions with respect to all the coefficients in (15). However, if the Gauss-Newton algorithm is initialized in the convex region containing the global minimum, then it will converge to that global minimum more likely in less iterations, as compared to the *oosrivcf* algorithm, because it uses at each iteration the global covariance matrix which contains the cross-covariance elements between the coefficients and the differentiation order(s).

Example 2 – Commensurate order estimation combined to the *srivcf* algorithm

The proposed method is applied to the system described in example 1. The model structure is set to (43), with an unknown commensurate order γ . For each commensurate order γ between 0.1 and 1.9, it has been previously shown, in Fig. 2 for a single noise realization, that the ℓ_2 -norm of the output error depends on the commensurate order and that its optimum is on the true value. This example intends to analyze the performance of the proposed algorithm when the commensurate order is estimated using Monte Carlo simulations with 500 runs in a noisy context with a SNR of 5dB, as in example 1. Each Monte Carlo run is initialized at a commensurate order $\gamma^0 = 1.8$. The estimation results are presented in Table 2, where the estimated parameter mean and standard deviation are given.

	True	Estimated	
	θ	$\hat{\theta}$	σ_θ
b_0	0.500	0.499	0.008
a_2	0.500	0.495	0.015
a_1	1.500	1.500	0.013
γ	1.500	1.504	0.012

Table 2

Monte Carlo simulation results – *oosrivcf* algorithm with commensurate order estimation in example 2.

The Monte Carlo simulation results reveal an estimation of the commensurate order highly accurate with a mean value close to the real commensurate order and with a low variance. Based on the efficacy of the *srivcf* method, the commensurate order estimate is asymptotically unbiased including for the commensurate differentiation order.

Example 3 – All order estimation combined to the *srivcf* algorithm

Consider now the case of estimating transfer function coefficients and all differentiation orders of the following data-generating system:

$$G_1(s) = \frac{0.5}{0.5s^{2.8} + 1.5s^{1.2} + 1}. \quad (62)$$

The input signal is the PRBS of Fig.1. The output, $y(t)$, is corrupted at the sampling instants (13) by an additive Gaussian white noise $\xi(t_k)$ with zero mean and a SNR of 5dB. As stated previously, it is recommended that a commensurate model, as in (43), be estimated first and used as an initial hit for estimating all orders.

The ℓ_2 -norm of the output error versus the commensurate order shows, in Fig. 3, that the criterion depends on the differentiation order. Even if the model is considered commensurate, in which case the model does not lie in the same class as the true system, the system identification algorithm behaves well and the minimum of the ℓ_2 -norm of the output error is found at a commensurate order of $\gamma = 1.25$. In this case, the estimated model for a given run is:

$$\hat{G}_1(s) = \frac{0.896}{0.695s^{2.5} + 2.15s^{1.25} + 1}. \quad (63)$$

As it can be checked, the transfer function coefficients are not equal to the true ones, which is a normal fact because the model is not in the same class as the true system. The obtained model constitutes however a good initial hit for all order estimation.

Fig. 3 shows further that the ℓ_2 -norm of the modeling error is around 0.6dB at the estimated commensurate order obtained after convergence and that it equals 0.8dB

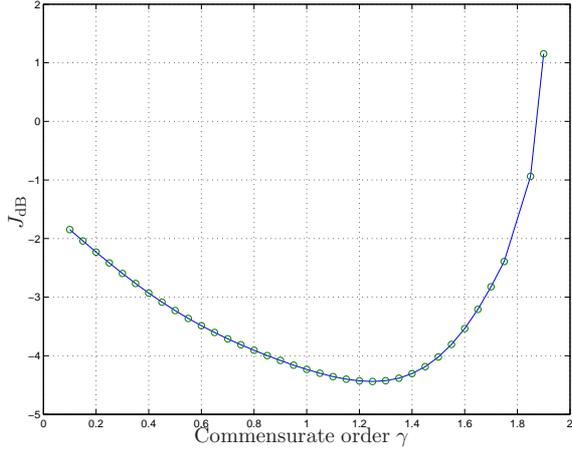


Fig. 3. Output error ℓ_2 -norm versus commensurate differentiation order in example 3.

	True	Estimated	
	θ	$\hat{\theta}$	σ_θ
b_0	0.500	0.501	0.010
a_2	0.500	0.503	0.018
a_1	1.500	1.500	0.032
α_2	2.800	2.794	0.028
α_1	1.200	1.196	0.019

Table 3
Monte Carlo simulation results – *oosrivcf* algorithm with all differentiation order estimation in example 3.

for a rational model with $\gamma = 1$. When there is no modeling error, J_{dB} equals $(-\text{SNR} = -5\text{dB})$ as explained in example 2. To improve the parametric estimation, the commensurability constraint is now released so that all differentiation orders are estimated separately. Hence, the model structure is now chosen as:

$$G(s) = \frac{b_0}{a_2 s^{\alpha_2} + a_1 s^{\alpha_1} + 1}. \quad (64)$$

A Monte Carlo simulation of 500 runs is held for a SNR of 5dB, as in examples 1 and 2. For each Monte Carlo run, the commensurability constraint is released and all differentiation orders are estimated. For each run, the algorithm is initialized with the commensurate model (63). The estimation synthesis is presented in Table 3, where the mean and the standard deviation of the parameter estimates are given. All estimated parameters converge to the true ones.

5 Application to a thermal plant identification

A long aluminium rod (Fig. 4) heated by a resistor at one end is considered in this experiment. The input signal is a

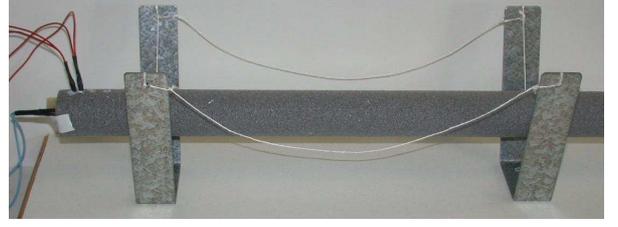


Fig. 4. Insulated long aluminium rod heated by a resistor

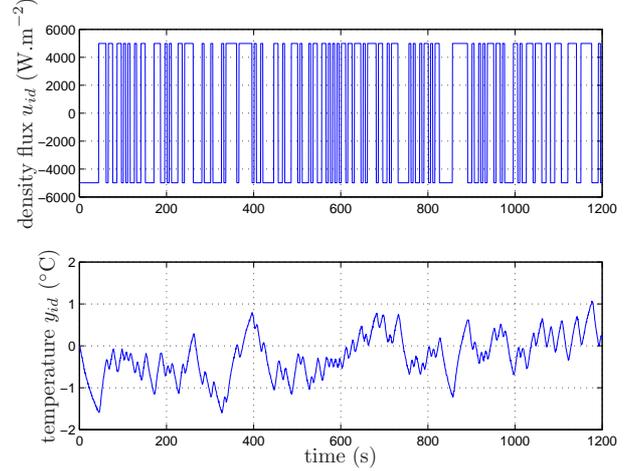


Fig. 5. Pretreated input/output signals of the thermal system with a constant flux of 5kWm^{-2} and a constant temperature of 42.2°C subtracted and a shift of the output signal to eliminate the delay

thermal flux generated by a resistor glued at one end and the output signal is the temperature of the rod measured at a distance $x = 0.5\text{cm}$ from the heated end using a platinum probe and an amplifier with a quantification error of 0.125 degree. To ensure a unidirectional heat transfer, the entire surface of the rod is insulated.

The sampling period is set to 0.5s. The system is driven to a steady-state temperature by injecting a constant flux density of $\phi = 5\text{kWm}^{-2}$ for a sufficiently long period. Then a prbs is applied with a flux variation of $\pm 5\text{kWm}^{-2}$ around the constant flux of 5kWm^{-2} . A delay of 4 samples (2s) is observed between the output and the input. The input density flux and the output temperature are pretreated to eliminate the constant parts and the delay. The pretreated signals are plotted in Fig.5.

In (Malti *et al.*, 2009), a theoretical model of the aluminium rod, obtained from heat equation under some simplifying assumptions, was compared to experimental data. Among the simplifying assumptions, the rod was assumed to be perfectly isolated. The theoretical model was found to be commensurate of order 0.5. It was also shown that fractional models are more compact than rational models: higher order rational models are required to get comparable results to fractional models.

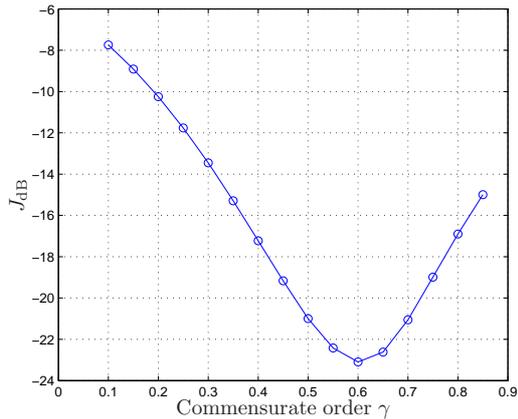


Fig. 6. Output error ℓ_2 -norm versus commensurate differentiation; the gap between 0.90 and 1.30 corresponds to unstable models.

Order estimation

Since the aluminum rod reaches a steady state temperature, it is not perfectly insulated. Consequently, there is no reason to have a commensurate order of 0.5 nor to have an integrator as in the physical model (Malti *et al.*, 2009). Thus, the selected model for the system identification is set to:

$$G_1(s) = \frac{b_0}{a_2 s^{\gamma_2} + a_1 s^{\gamma_1} + 1} \times e^{-2s}. \quad (65)$$

First of all, a commensurate model is considered with $\gamma_2 = 2\gamma_1$ and the *srivcf* algorithm is applied for a commensurate order γ_1 varying from 0.1 to 1.9 with a step of 0.05 yielding transfer function coefficients for each commensurate order. The ℓ_2 -norm (in dB) of the normalized output error,

$$J_{dB} = 10 \log \left(\frac{\|y^*(t) - \hat{y}(t)\|^2}{\|y^*(t)\|^2} \right), \quad (66)$$

is plotted versus the commensurate order in Fig. 6. For commensurate orders greater than 0.85, either unstable or poor quality models are obtained due to important modeling errors. The commensurate order leading to the minimum of the ℓ_2 -norm of the normalized output error is around $\gamma_1 = 0.60$. The *srivcf* method applied to a rational model, with $\gamma = 1$, does not converge to a stable model.

Hence applying the *oosrivcf* algorithm to estimate the commensurate order, initialized according to the prior knowledge to 0.5, leads to the global minimum $\gamma_1 = 0.593$ and to the following commensurate model:

$$G_1(s) = \frac{3.09 \times 10^{-3}}{263.4s^{1.186} + 88.78s^{0.593} + 1} e^{-2s}, \quad (67)$$

	Estimated	
	$\hat{\theta}$	σ_{θ}
b_0	5.89×10^{-3}	0.13×10^{-3}
a_2	954	20
a_1	168	2
γ_2	1.474	0.004
γ_1	0.557	0.002

Table 4
Thermal process parameter estimation with the *oosrivcf* algorithm.

with a corresponding ℓ_2 -norm of the normalized output error:

$$J_{dB} = -23.1\text{dB}. \quad (68)$$

Then, the commensurability constraint is released so that all differentiation orders and transfer function coefficients are estimated with the *oosrivcf* algorithm. Model (67) is used to initialize the *oosrivcf* algorithm for all order estimation and it converges to the parameters of Table 4, where the estimated standard deviations are obtained from the covariance matrix (57). In comparison with (68), the corresponding ℓ_2 -norm of the normalized output error drops to:

$$J_{dB} = -25.6\text{dB}. \quad (69)$$

By releasing the commensurability constraint, the ℓ_2 -norm of the normalized output error decreases by 2.5dB. The time-domain responses of the commensurate (67) and the non commensurate (65) models with the parameters of Table 4 are both compared on a set of validation data in Fig. 7 together with the output error. As expected, the estimated model (65) with the parameters of Table 4 outperforms the commensurate model (67).

This kind of model has been successfully used in (Victor *et al.*, 2009b; Victor *et al.*, 2011) for temperature path generation in open-loop by using flatness principles, and also for temperature path tracking in closed-loop by using a CRONE controller that guarantees a robust path planning in presence of input/output disturbances and in presence of parametric variations.

6 Conclusion and prospects

This paper has presented an optimal instrumental variable method for estimating transfer function coefficients of fractional differential systems when the output is corrupted by an additive white noise. The well-known *srivcf* algorithm has been extended to estimating transfer function coefficients of fractional models when all differentiation orders are set according to prior knowledge. Then, when the prior knowledge is not available, a very important aspect of fractional differential models has been

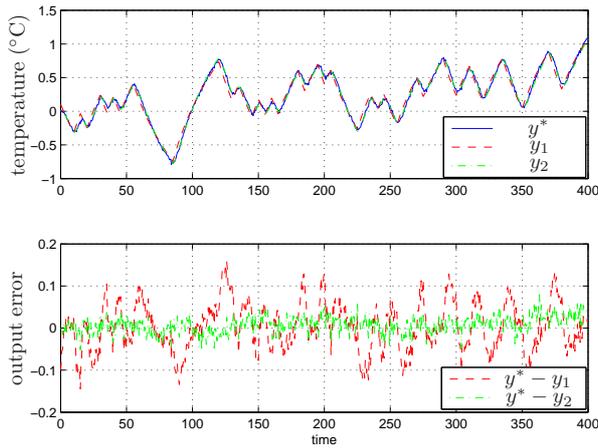


Fig. 7. Time-domain responses plotted on validation data (y_1 stands for the commensurate model output (67) and y_2 for the non commensurate model output of Table 4)

treated: the determination of differentiation orders. For that purpose a gradient-based algorithm, which combines simultaneously the parameter and the differentiation order estimation, has been proposed with two variants: either a commensurate order is estimated, or all differentiation orders are further adjusted. An initialization procedure is proposed consisting of estimating a commensurate model first, and then of using the obtained model as an initial hit for all order estimation. The performance of the proposed algorithms has been evaluated by Monte Carlo simulation analysis. Last but not least, the proposed algorithms have been successfully applied to identify thermal diffusion in an aluminum rod.

In a future work, it would be interesting to develop techniques to help fixing the number of parameters when the model structure is not known. It will also be interesting to extend this study to deal with colored output noise by using hybrid Box-Jenkins models with continuous-time fractional input-output models and discrete-time noise models.

References

Aoun, M., R. Malti, F. Levron and A. Oustaloup (2004). Numerical simulations of fractional systems: an overview of existing methods and improvements. *Nonlinear Dynamics* **38**(1-4), 117–131.

Aoun, M., R. Malti, F. Levron and A. Oustaloup (2007). Synthesis of fractional Laguerre basis for system approximation. *Automatica* **43**, 1640–1648.

Bai, E.W. (1998). An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica* **34**(3), 333–338.

Battaglia, J.-L., O. Cois, L. Puigsegur and A. Oustaloup (2001). Solving an inverse heat conduction problem using a non-integer identified model. *Int. J. of Heat and Mass Transfer* **44**(14), 2671–2680.

Benchellal, A., S. Bachir, T. Poinot and J.-C. Trigeassou (2004). Identification of a non-integer model of induction machines. In: *1st IFAC Workshop on Fractional Differentiation and its Applications (FDA)*. Bordeaux, France.

Cois, O., A. Oustaloup, T. Poinot and J.-L. Battaglia (2001). Fractional state variable filter for system identification by fractional model. In: *6th European Control Conference ECC'01*. Porto, Portugal.

Gabano, J.-D. and T. Poinot (2011a). Estimation of thermal parameters using fractional modelling. *Signal Processing* **91**(4), 938 – 948.

Gabano, J.-D. and T. Poinot (2011b). Fractional modelling and identification of thermal systems. *Signal Processing* **91**(3), 531 – 541. *Advances in Fractional Signals and Systems*.

Garnier, H. and L. Wang (2008). *Identification of continuous-time models from sampled data*. Springer-Verlag.

Garnier, H., M. Gilson, P.C. Young and E. Huselstein (2007). An optimal IV technique for identifying continuous-time transfer function model of multiple input systems. *Control Engineering Practice* **46**(15), 471–486.

Garnier, H., M. Mensler and A. Richard (2003). Continuous-time model identification from sampled data. Implementation issues and performance evaluation. *International Journal of Control* **76**(13), 1337–1357.

Gilson, M., H. Garnier, P.C. Young and P. Van den Hof (2008). Instrumental Variable Methods for Closed-Loop Continuous-time Model Identification. In: *Identification of continuous-time models from sampled data* (H. Garnier and L. Wang, Eds.). Springer-Verlag.

Heymans, N. and J.C. Bauwens (1994). Fractal rheological models and fractional differential equations for viscoelastic behavior. *Rheologica Acta* **33**, 219.

Kay, S.M. (1993). *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall. Upper Saddle River, NJ.

Khemane, F., R. Malti, T. Raïssi and X. Moreau (2012). Robust estimation of fractional models in the frequency domain using set membership methods. *Signal Processing* **92**, 1591–1601.

Krishna, B.T. (2011). Studies on fractional order differentiators and integrators: A survey. *Signal Processing* **91**(3), 386–426.

Laurain, V., R. Toth, M. Gilson and H. Garnier (2011). Direct identification of continuous-time LPV input/output models. *IET Control Theory & Applications* **5**(7), 878–888.

Liouville, J. (1832). Mémoire sur quelques questions de géométrie et de mécanique et sur un nouveau genre de calcul pour résoudre ces équations. *Journal de l'Ecole Polytechnique* **13**, 71–162.

Liu, X., J. Wang and W.X. Zheng (2011). Convergence analysis of refined instrumental variable method for continuous-time system identification. *IET Control Theory & Applications* **5**(7), 868–877.

Ljung, L. (1999). *System identification – Theory for the user*. 2 ed.. Prentice-Hall. Upper Saddle River, N.J., USA.

Malti, R., J. Sabatier and H. Akçay (2009). Thermal modeling and identification of an aluminium rod using fractional calculus. In: *15th IFAC Symposium on System Identification (SYSID'2009)*. St Malo, France.

Malti, R., M. Aoun, F. Levron and A. Oustaloup (2011a). Analytical computation of the \mathcal{H}_2 -norm of fractional commensurate transfer functions. *Automatica*. Accepted.

Malti, R., S. Victor and A. Oustaloup (2008). Advances in system identification using fractional models. In: *Journal*

- of *Computational and Nonlinear Dynamics* (ASME, Ed.). Vol. 3. New York, USA. pp. 021401.1–021401.7.
- Malti, R., X. Moreau, F. Khemane and A. Oustaloup (2011b). Stability and resonance conditions of elementary fractional transfer functions. *Automatica* **47**(11), 2462–2467.
- Mathieu, B., A. Oustaloup and F. Levron (1995). Transfer function parameter estimation by interpolation in the frequency domain. In: *European Control Conference, ECC'95*. Rome, Italie.
- Matignon, D. (1998). Stability properties for generalized fractional differential systems. *ESAIM proceedings - Systèmes Différentiels Fractionnaires - Modèles, Méthodes et Applications*.
- Moze, M. and J. Sabatier (2005). LMI tools for stability analysis of fractional systems. In: *20th ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, IDETC/CIE'05*. Long Beach, CA. pp. 1–9.
- Mukhopahyay, S., A. Patra and G. P. Rao (1991). Irreducible model estimation for MIMO systems. *International Journal of Control* **53**(1), 223–253.
- Narang, A., S.L. Shah and T. Chen (2011). Continuous-time model identification of fractional-order models with time delays. *IET Control Theory & Applications* **5**(7), 900–912.
- Oldham, K.B. and J. Spanier (1970). The replacement of fick's laws by a formulation involving semi-differentiation. *Electroanal. Chem. Interfacial Electrochem* **26**, 331–341.
- Oldham, K.B. and J. Spanier (1974). *The fractionnal calculus - Theory and Applications of Differentiation and Integration to Arbitrary Order*. Academic Press, New-York and London.
- Oustaloup, A. (1995). *La dérivation non-entière*. Hermès. Paris.
- Pierce, D.A. (1972). Least squares estimation in dynamic disturbance time-series models. *Biometrika* **5**, 73–78.
- Podlubny, I. (1999). *Fractional Differential Equations*. Academic Press. San Diego.
- Rao, G.P. and H. Unbehauen (2006). Identification of continuous-time systems. *Electrical Engineering* **153**, 185 – 220.
- Rao, G.P., K. Diekmann and H. Unbehauen (1984). Parameter estimation in large-scale interconnected systems. In: *9th IFAC World Congress*. Budapest, Hungary. pp. 118–122.
- Riemann, B. (1876). Versuch einer allgemeinen Auffassung der Integration und Differentiation. *Gesammelte Mathematische Werke und Wissenschaftlicher* pp. 331–344.
- Rodrigues, S., N. Munichandraiah and A.-K. Shukla (2000). A review of state of charge indication of batteries by means of A.C. impedance measurements. *Journal of Power Sources* **87**, 12–20.
- Sabatier, J., M. Aoun, A. Oustaloup, G. Grégoire, F. Ragot and P. Roy (2006). Fractional system identification for lead acid battery state charge estimation. *Signal processing* **86**(10), 2645–2657.
- Sabatier, J., M. Moze and Ch. Farges (2010). LMI stability conditions for fractional order systems. *Computers & Mathematics with Applications* **59**(5), 1594 – 1609. Fractional Differentiation and Its Applications.
- Söderström, T. (2012). How accurate can instrumental variable models become?. In: *System Identification, Environmental Modelling, and Control System Design* (L. Wang and H. Garnier, Eds.). Springer-Verlag.
- Söderström, T. and P. Stoica (1983). *Instrumental variable methods for system identification*. Springer Verlag. New York.
- Spinelli, W., L. Piroddi and M. Lovera (2006). A two-stage algorithm for structure identification of polynomial NARX models. *American Control Conference Vols 112* **1-12**, 2387–2392.
- Stoica, P. and R. Moses (2005). *Spectral Analysis of Signals*. Pearson Prentice-Hall. Upper Saddle River, NJ.
- Victor, S., P. Melchior, R. Malti and A. Oustaloup (2011). Path tracking with flatness and crone control for fractional systems. In: *The 18th IFAC World Congress (IFAC'11)*. Milan, Italy.
- Victor, S., R. Malti and A. Oustaloup (2009a). Instrumental variable method with optimal fractional differentiation order for continuous-time system identification. In: *15th IFAC Symposium on System Identification (SYSID'2009)*. Saint Malo, France.
- Victor, S., R. Malti, P. Melchior and A. Oustaloup (2009b). From system identification to path planning using fractional approach: a thermal application example. In: *21th ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, IDETC/CIE'09*. San Diego, USA.
- Wang, J., W. X. Zheng and T. Chen (2009). Identification of linear dynamic systems operating in a networked environment. *Automatica* **45**(12), 2763–2772.
- Wellstead, P. E. (1978). An instrumental product moment test for model order estimation. *Automatica* **14**, 89–91.
- Young, P.C. (1981). Parameter estimation for continuous-time models – a survey. *Automatica* **17**(1), 23–29.
- Young, P.C. (2002). Optimal IV identification and estimation of continuous-time TF models. In: *15th World IFAC Congress*. IFAC. Elsevier. Barcelona (Spain).
- Young, P.C. (2011). *Recursive estimation and Time-Series Analysis*. 2nd ed.. Springer-Verlag. Berlin.
- Young, P.C. and A.J. Jakeman (1980). Refined instrumental variable methods of time-series analysis: Part III, extensions. *International Journal of Control* **31**, 741–764.
- Young, P.C., H. Garnier and M. Gilson (2008). Identification of continuous-time models from sampled data. In: *Identification of continuous-time models from sampled data* (H. Garnier and L. Wang, Eds.). Chap. Refined Instrumental Variable Identification of Continuous-time Hybrid Box-Jenkins Models, pp. 91–131. Springer-Verlag.