2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Model order identification for fractional models

Stéphane Victor, Rachid Malti

Université de Bordeaux, IMS, CNRS UMR 5218,
351 cours de la Libération, 33405 Talence Cedex, France
{stephane.victor, rachid.malti}@ims-bordeaux.fr

***Abstract*— This paper deals with continuous-time system identification using fractional differentiation models. So far, no algorithm exists concerning model order identification of fractional models. The "simplified" refined instrumental variable method is proposed to estimate parameters of fractional differential equation models when all the fractional orders are assumed known. Then an optimization approach based on this instrumental variable estimator is presented. This two-stage algorithm, called *oosrivcf* is then used for model order identification by using simultaneously two criteria: the Young Information Criterion (YIC) and the R_T^2 criterion.**

I. Introduction

Fractional calculus dates back to Liouville and Riemann in the ninetieth century and excels in modeling diffusive phenomena. Oldham and Spanier [1] expressed diffusive phenomena in semi-infinite media where diffusive systems are modeled with transfer functions involving differentiation orders multiples of 0.5. In electrochemistry, it is proven that the diffusion of charges in acid batteries is governed by Randles models [2], [3] that involve a half order integrator. In semi-infinite thermal systems, Battaglia et al. [4] 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. Foucault currents inside rotor bars in induction machines obey to diffusive phenomena modeled by fractional models [5].

Instrumental variable methods for system identification have been developed since the seventies. For a detailed description refer to [6], [7], [8] and all references therein. **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 [6], [9]. 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.

In system identification with rational models, where only the coefficients are estimated, the model order remains unchanged, whereas estimating coefficients and differentiation orders in fractional models is not a trivial task given that estimating the differentiation orders in an iterative algorithm changes the model order at each iteration. In this paper, the ***srivc*** algorithm helps for linear coefficient estimation, and a gradient-based approach allows the differentiation order estimation. [10] have exposed this two-stage algorithm with an application to a thermal system. However, to the best of authors' knowledge, no algorithm has ever been proposed in the literature to fix the model structure, and consequently to fix optimally the number of parameters.

Fractional models allow modeling infinite dimensional systems with reduced number of parameters. Since it is always possible to find a high order rational model equivalent to a low dimension fractional model [11], the use of high dimension fractional models is of limited interest.

Mathematical background

Differentiation to an arbitrary order γ, with $\gamma \in \mathbb{R}$ was defined by Riemann and Liouville as being an integer derivative of order $\lfloor\gamma\rfloor+1$ ($\lfloor.\rfloor$ stands for the floor operator) of a non-integer integral $\mathbf{I}$ [12]:

$$p^{\gamma}x(t) = p^{\lfloor\gamma\rfloor+1}\left(\mathbf{I}^{\lfloor\gamma\rfloor+1-\gamma}x(t)\right) \triangleq \left(\frac{d}{dt}\right)^{\lfloor\gamma\rfloor+1}\left(\frac{1}{\Gamma\left(\lfloor\gamma\rfloor+1-\gamma\right)}\int_0^t \frac{x\left(\tau\right)d\tau}{\left(t-\tau\right)^{\gamma-\lfloor\gamma\rfloor}}\right), \quad (1)$$

where $t>0$, $\gamma \in \mathbb{R}_+^*$. The Euler's Γ function is defined in the set of strictly positive real numbers $x \in \mathbb{R}_+^*$ as:

$$\Gamma(x) = \int_0^{\infty} e^{-t}t^{x-1}dt, \quad (2)$$

and extended to the rest of the real numbers excepting negative integers $x \in \mathbb{R}^* \setminus \mathbb{N}^-$ by analytical continuity.

It is assumed that the system input $u(t)$ and the noise-free system output $y(t)$ are related by the following fractional differential equation,

$$y\left(t\right)+a_1p^{\alpha_1}y\left(t\right)+\cdots+a_Np^{\alpha_N}y\left(t\right) = b_0p^{\beta_0}u\left(t\right)+b_1p^{\beta_1}u\left(t\right)+\cdots+b_Mp^{\beta_M}u\left(t\right), \quad (3)$$

where p stands for the differentiation operator $p=\frac{\mathrm{d}}{\mathrm{d}t}$, and so $p^{\mu}u(t)$ denotes the μ-th time-derivative of the continuous-time signal $u(t)$, $(a_j,b_i) \in \mathbb{R}^2$, and the differentiation orders

$$0<\alpha_1<\ldots<\alpha_N, \quad 0\leq\beta_0<\beta_1<\ldots<\beta_M, \quad (4)$$

are ordered for identifiability purposes.

It is easy to define a symbolic representation of the fractional dynamic system governed by (3) using the transfer

function:

$$G(p) = \frac{B(p)}{A(p)} = \frac{\sum_{i=0}^{M} b_i p^{\beta_i}}{1 + \sum_{j=1}^{N} a_j p^{\alpha_j}}. \quad (5)$$

$G(p)$ is said commensurate of order γ, when rewritten as:

$$G(p) = \frac{\sum_{i=0}^{m} \tilde{b_i} p^{i\gamma}}{1 + \sum_{j=1}^{n} \tilde{a_j} p^{j\gamma}}, \quad (6)$$

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

$$\begin{cases} \tilde{b_{i'}} = b_i \text{ if } \exists i \in \{0, 1, \ldots, m\} \text{ such that } i'\gamma = \beta_i \\ \tilde{b_{i'}} = 0 \text{ otherwise} \\ \tilde{a_{j'}} = a_j \text{ if } \exists j \in \{1, \ldots, n\} \text{ such that } j'\gamma = \alpha_j \\ \tilde{a_{j'}} = 0 \text{ otherwise.} \end{cases} \quad (7)$$

In rational transfer functions, γ equals 1 and usually numerator α_N and denominator β_M orders are both fixed, then all coefficients $b_i, i = 0, \ldots, M$ and $a_j, j = 1, \ldots, N$ are estimated. Generally, no care is taken to check whether any intermediate coefficient, as in (7), equals zero.

Time-domain simulation of fractional systems is an extensively studied topic in the literature [11], [13], [14]. System identification algorithms proposed in this paper may be used with any time-domain simulation algorithm of fractional systems. It is however definitely of his responsibility to correctly simulate the fractional system with negligible simulation errors, in order to be able to consistently estimate the fractional model parameters. Otherwise, simulation errors would affect the estimated parameters. The simulation algorithm used in this paper is based on [11] approximation of fractional operator in a desired frequency band by a recursive distribution of zeros and poles.

Stability theorem ([15], extended): *A commensurate transfer function, with a commensurate order γ as in* (6)*, $G(p) = \frac{T(p^\gamma)}{R(p^\gamma)}$, where T and R are coprime polynomials, is* BIBO *stable if and only if :*

$$0 < \gamma < 2 \quad (8)$$

and

$$|\arg(s_k)| > \gamma\frac{\pi}{2}, \quad \forall s_k \in \mathbb{C} \,/\, R(s_k) = 0. \quad (9)$$

The stability conditions are used later when estimating the commensurate order, which must be in the interval (8) for stable systems.

II. Problem formulation

For simplicity, the formulation and solution of the estimation problem is restricted to the case of linear, single-input, single-output system, the analysis being extendable to multivariable fractional systems (see e.g.[16]).

It is assumed that the input $u(t)$ and the noise-free output $y(t)$ are related by the constant coefficient differential equation (3), which can also be written in the compact model form (5), where $B(p)$ 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 are negligible. 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 $\{u(t), 0 \le t \le T_{final}\}$ applied to the system is persistently exciting, and gives rise to an output signal $\{y(t), 0 \le t \le T_{final}\}$. 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:

$$\begin{cases} y(t) = G(p)u(t) \\ y^*(t_k) = y(t_k) + \xi(t_k), \end{cases} \quad (10)$$

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. [17]). 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 (5) is used, the parameter vector

$$\theta = \begin{bmatrix} \rho \\ \mu \end{bmatrix} \quad (11)$$

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

$$\rho = [b_0, b_1, \ldots, \ b_M, a_1, \ldots, a_N]^{\mathrm{T}}, \quad (12)$$

and additionally a vector of $N+M+1$ ordered differentiation orders, as in (4),

$$\mu = [\beta_0, \ldots, \beta_M, \alpha_1, \ldots, \alpha_N]^{\mathrm{T}}. \quad (13)$$

Estimating all differentiation orders requires 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 (6) can be chosen, where only one differentiation order and hence $N + M + 2$ parameters are estimated. The parameter vector θ in (11) is then reduced to the $N+M+1$ transfer function coefficients as in (12) added to a single commensurate differentiation order:

$$\mu = \gamma. \quad (14)$$

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

III. Model Order Identification for Continuous-Time Fractional Models

A. 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 ***srivc*** [9], [18] algorithm derives from the ***sriv*** [19], [6] algorithm for discrete-time transfer function model identification and have been successfully used for more than three decades in various contexts. This was evolved by converting the maximum likelihood (ML) estimation equations to a pseudo-linear form involving optimal prefilters (see e.g. [7]). The iterative ***srivcf*** algorithm (***srivc*** for fractional models), known to deliver optimal estimates when the additive measurement noise is white, is recalled (see [20], [10] for the details). Moreover, it has recently been established in the rational continuous-time case in [21, theorem 1] that the ***srivc*** algorithm converges asymptotically in one iteration under some *mild assumptions*[1]. In the fractional case, the convergence of the ***srivcf*** algorithm follows exactly the same scheme under the same assumptions.

All linear transformations can be interpreted as lowpass filtering. Not only direct differentiation of noised signals is avoided, but also it allows to specify the frequency band into which the model and system adequacy is sought. Moreover the filter frequency response should match its system transfer function. The ***srivcf*** uses a filter updated iteratively which avoids choosing *a priori* its design parameters.

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

$$\begin{cases} \mathrm{E}\left[\varphi_{\mathrm{iv},f}^{\mathrm{iter}}(t_k)\varphi_f^{\mathrm{T}}(t_k)\right] \text{ is non-singular,} \\ \mathrm{E}\left[\varphi_{\mathrm{iv},f}^{\mathrm{iter}}(t_k)\xi(t_k)\right] = 0. \end{cases} \tag{15}$$

For lack of space, only the main steps of the ***srivcf*** algorithm are recalled as follows. The interested reader may refer to [10] for more details.

Step 1 ***Initialization***:
iter=0
Generate an initial[2] estimate of coefficient parameter vector $\hat{\rho}^0$.

Step 2 ***Iterative estimation***:
do

(i) Generate the auxiliary model output

$$y_{\mathrm{iv}}^{\mathrm{iter}}(t) = \frac{\hat{B}\left(p,\hat{\rho}^{\mathrm{iter}}\right)}{\hat{A}\left(p,\hat{\rho}^{\mathrm{iter}}\right)}u(t). \tag{16}$$

(ii) Update the filter

$$F_\sigma^{\mathrm{iter}}(p) = \frac{p^\sigma}{\hat{A}(p,\hat{\rho}^{\mathrm{iter}})}, \tag{17}$$

and compute the filtered derivatives of the input, the output, and the instruments:

$$\begin{aligned} p^{\beta_i}u_f(t) &= F_{\beta_i}^{\mathrm{iter}}(p)u(t), \\ p^{\alpha_j}y_f^*(t) &= F_{\alpha_j}^{\mathrm{iter}}(p)y^*(t), \\ p^{\alpha_j}y_{\mathrm{iv},f}^{\mathrm{iter}}(t) &= F_{\alpha_j}^{\mathrm{iter}}(p)y_{\mathrm{iv}}^{\mathrm{iter}}(t). \end{aligned}$$

(iii) Generate the regression $\varphi_f(t)$ and the instrumental variable $\varphi_{\mathrm{iv},f}^{\mathrm{iter}}(t)$ vectors :

$$\varphi_f(t) = \left[\begin{array}{c} p^{\beta_0}u_f(t),\ldots,p^{\beta_M}u_f(t), \\ -p^{\alpha_1}y_f^*(t),\ldots,-p^{\alpha_N}y_f^*(t) \end{array}\right]^T$$

$$\varphi_{\mathrm{iv},f}^{\mathrm{iter}}(t) = \left[\begin{array}{c} p^{\beta_0}u_f(t),\ldots,p^{\beta_M}u_f(t), \\ -p^{\alpha_1}y_{\mathrm{iv},f}^{\mathrm{iter}}(t),\ldots,-p^{\alpha_N}y_{\mathrm{iv},f}^{\mathrm{iter}}(t) \end{array}\right]^T$$

(iv) Based on these prefiltered data, compute the new estimates $\hat{\theta}_{\mathrm{iter}}^{\boldsymbol{srivcf}}$

$$\hat{\rho}^{\mathrm{iter}+1} = \left(\boldsymbol{\Phi}_{\mathrm{iv},f}^{\mathrm{iter}}\boldsymbol{\Phi}_f^{\mathrm{T}}\right)^{-1}\boldsymbol{\Phi}_{\mathrm{iv},f}^{\mathrm{iter}}\mathbf{Y}_f^*, \tag{18}$$

with

$$\begin{aligned} \boldsymbol{\Phi}_{\mathrm{iv},f}^{\mathrm{iter}} &= \left[\varphi_{\mathrm{iv},f}^{\mathrm{iter}}(0) \quad \ldots\varphi_{\mathrm{iv},f}^{\mathrm{iter}}\left(T_{final}\right)\right]^{\mathrm{T}}, \\ \boldsymbol{\Phi}_f &= \left[\varphi_f(0) \quad \ldots\varphi_f\left(T_{final}\right)\right]^{\mathrm{T}}, \\ \mathbf{Y}_f^* &= \left[y_f^*(0) \quad \ldots y_f^*\left(T_{final}\right)\right]^{\mathrm{T}}. \end{aligned}$$

while $\max_j \left|\frac{\hat{\rho}_j^{\mathrm{iter}+1}-\hat{\rho}_j^{\mathrm{iter}}}{\rho_j^{\mathrm{iter}+1}}\right| > \epsilon_1$,
where $\hat{\rho}_j^{\mathrm{iter}}$ is the j-th element of the estimated parameter vector at the iteration iter.

Step 3 ***Parametric error estimation***
Compute the estimated parametric error covariance matrix

$$\hat{\mathbf{P}}_\rho = \hat{\sigma}^2\left(\boldsymbol{\Phi}_{\mathrm{iv},f}\boldsymbol{\Phi}_{\mathrm{iv},f}{}^{\mathrm{T}}\right)^{-1}, \tag{19}$$

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

B. 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 estimate 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. Two variants of order optimization algorithm are recalled [10]. Either all differentiation orders are set as integral multiples of a commensurate order, as in the model (**??**), and the best commensurate order is computed in the interval (0, 2), or all differentiation orders, as in (5), are estimated. The former variant is used as a good initial hit

[1] as qualified by the authors of the paper.

[2] Least-squares or instrumental variable methods combined with fractional state variable filters can be used (see [20]).

for the latter variant. The estimation problem is formulated as a minimization problem of the ℓ_2-norm:

$$J(\theta) = \frac{1}{2} \left\| \varepsilon \left(t, \theta \right) \right\|^2, \tag{20}$$

of the output error

$$\varepsilon(t, \theta) = y^*(t) - G(p, \theta) u(t), \tag{21}$$

with respect to the vector μ, as in (13) or (14).

A Gauss-Newton algorithm (see e.g. [22]) is used for the estimation of all differentiation orders μ as defined in (13). For lack of space, only the main steps of the ***oosrivcf*** algorithm are recalled as follows:

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) Compute the error sensitivity function:

$$\frac{\partial \varepsilon}{\partial \mu} = \left[-\frac{\partial \hat{y}}{\partial \beta_0}, \ldots, -\frac{\partial \hat{y}}{\partial \beta_M}, -\frac{\partial \hat{y}}{\partial \alpha_1}, \ldots, -\frac{\partial \hat{y}}{\partial \alpha_N} \right]^{\mathrm{T}}, \tag{22}$$

where the output sensitivity functions are computed numerically.

(ii) Compute the gradient

$$\frac{\partial J}{\partial \mu} = \frac{\partial \varepsilon}{\partial \mu}^{\mathrm{T}} \varepsilon, \tag{23}$$

and the approximated Hessian

$$H = \frac{\partial \varepsilon}{\partial \mu}^{\mathrm{T}} \frac{\partial \varepsilon}{\partial \mu}. \tag{24}$$

(iii) Compute the new order estimate:

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

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

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

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

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

iter = iter + 1

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

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 [22, Theorem 9.1] by:

$$\hat{\mathbf{P}}_\theta = \hat{\sigma}^2 \mathcal{H}^{-1}, \tag{25}$$

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:

$$\mathcal{H} = \frac{\partial \varepsilon}{\partial \theta}^{\mathrm{T}} \frac{\partial \varepsilon}{\partial \theta} = \frac{\partial \varepsilon^{\mathrm{T}}}{\partial \left[\rho^{\mathrm{T}} \mu^{\mathrm{T}} \right]^{\mathrm{T}}} \frac{\partial \varepsilon}{\partial \left[\rho^{\mathrm{T}} \mu^{\mathrm{T}} \right]^{\mathrm{T}}}, \tag{26}$$

with $\frac{\partial \varepsilon}{\partial \mu}$ defined in (22) and

$$\frac{\partial \varepsilon}{\partial \rho} = - \left[\frac{\partial \hat{y}}{\partial b_0}, \ldots, \frac{\partial \hat{y}}{\partial b_M}, \frac{\partial \hat{y}}{\partial a_1}, \ldots, \frac{\partial \hat{y}}{\partial a_N} \right]^{\mathrm{T}}, \tag{27}$$

where $\forall r = 0, \ldots, M$, and $\forall k = 1, \ldots, N$

$$\frac{\partial \hat{y}}{\partial b_r} = \frac{\hat{p}^{\hat{\beta}_r}}{1 + \sum_{j=1}^{N} \hat{a}_j p^{\hat{\alpha}_j}} u(t), \tag{28}$$

$$\frac{\partial \hat{y}}{\partial a_k} = - \frac{\sum_{i=0}^{M} \hat{b}_i p^{\hat{\beta}_i + \hat{\alpha}_k}}{\left(1 + \sum_{j=1}^{N} \hat{a}_j p^{\hat{\alpha}_j} \right)^2} u(t). \tag{29}$$

Remarks:

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 the convergence to a local minimum is guaranteed. The λ parameter is used to adjust the step in case of slow convergence or oscillations around a minimum.
2) The output sensitivity functions ($\forall r = 0, \ldots, M$ and $\forall k = 1, \ldots, N$)

$$\frac{\partial \hat{y}}{\partial \beta_r} = \ln(p) \frac{\hat{b}_r p^{\hat{\beta}_r}}{1 + \sum_{j=1}^{N} \hat{a}_j p^{\hat{\alpha}_j}} u(t), \tag{30}$$

$$\frac{\partial \hat{y}}{\partial \alpha_k} = \ln(p) \hat{a}_k p^{\hat{\alpha}_k} \frac{\sum_{i=0}^{M} \hat{b}_i p^{\hat{\beta}_i}}{\left(1 + \sum_{j=1}^{N} \hat{a}_j p^{\hat{\alpha}_j} \right)^2} u(t). \tag{31}$$

depend on $\ln(p)$ (log of the differential operator) which is not trivial to simulate in the time-domain. That is why they are computed numerically, using for example the central difference method.
3) The ***oosrivcf*** algorithm works similarly when considering the commensurate order; μ is then defined as in (14). Instead of using (22), the following error sensitivity function will be used:

$$\frac{\partial \varepsilon}{\partial \mu} = \frac{\partial \varepsilon}{\partial \gamma} = - \frac{\partial \hat{y}}{\partial \gamma}, \tag{32}$$

where

$$\frac{\partial \hat{y}}{\partial \gamma} = \ln(p) \frac{\sum_{i=0}^{M} i\hat{b}_i p^{i\hat{\gamma}} + \sum_{i=0}^{M}\sum_{j=1}^{N} (i-j)\,\hat{b}_i \hat{a}_j p^{(i+j)\hat{\gamma}}}{\left(1 + \sum_{j=1}^{N} \hat{a}_j p^{j\hat{\gamma}}\right)^2} u(t) \tag{33}$$

depends on $\ln(p)$. As previously, it will also be computed numerically.

In case all differentiation orders are estimated, it is highly recommended to estimate a commensurate model, 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. Monte Carlo simulation results have been successfully carried out and presented in [10].

C. Model Order estimation

In the identification procedure, a key point involves the model order structure selection: the number of parameters N and M in the transfer function (5) need to be fixed. This point is named as Model order selection in rational systems. However, in fractional systems, model order α_N in (5) is estimated along with the coefficients. Hence, it is preferable to talk about structure selection. The method developed for rational SISO systems [9] is extended to fractional systems. Two statistical measures are computed and used to choose between a range of orders M and N, which are defined as:

$$R_T^2 = 1 - \frac{\hat{\sigma}_\xi^2}{\hat{\sigma}_y^2}, \tag{34}$$

$$YIC = \log\left\{\frac{\hat{\sigma}_\xi^2}{\hat{\sigma}_y^2}\right\} + \log\left\{\frac{1}{n_\theta}\sum_{i=1}^{n_\theta} \frac{\hat{\sigma}_\xi^2 \hat{p}_{ii}}{\hat{\theta}_i^2}\right\}, \tag{35}$$

where $\hat{\sigma}_\xi^2$ and $\hat{\sigma}_y^2$ denote respectively the variance of the simulation error and the variance of the measured output, $\hat{p}_{ii}$ is the i-th diagonal element of the estimated parametric covariance matrix given in (25) and θ_i^2 is the square of the i-th parameter estimate in the θ vector so that the ratio $\hat{p}_{ii}/\theta_i^2$ is a normalized measure of the uncertainty on the i-th parameter estimate.

R_T^2 is recognized as the coefficient of determination based on the simulation error. It is a normalized measure of how much of the output variance is explained by the deterministic system part of the estimated model and it will be close to 1 in low noise situations. However, this measure is not sufficient to avoid over-parametrization and identify a parsimonious model. The Young's information criterion YIC is more complex and provides a measure of how well parameters are defined statistically: the more negative YIC, the better the definition. However, it may lead to underestimate the orders M and N. Both criteria are inspected to find the orders for which R_T^2 is sufficiently high to indicate a good explanation of the data and the YIC is sufficiently negative to indicate well-defined parameter estimates.

D. Simulation example

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

$$G_1(p) = \frac{0.5}{0.5p^{2.8} + 1.5p^{1.2} + 1}. \tag{36}$$

The input signal is a PRBS (see Fig.1).

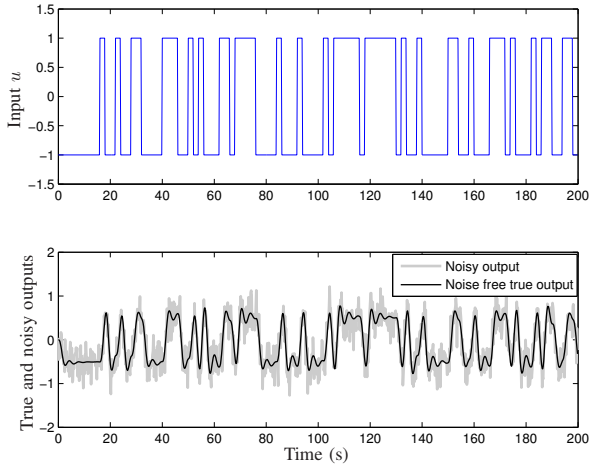

Fig. 1. Input, noise-free and noisy outputs, used as identification data, zoomed between $t = 0$ and 200s, identification data going up to 600s (SNR = 10dB and $h = 0.02$s)

The output $y(t)$ is corrupted at the sampling instants (10) by an additive Gaussian white noise $\xi(t_k)$ with zero mean and a SNR of 10dB. It is highly recommended that a commensurate model, as in (6), be estimated first and used as an initial hit for estimating all orders.

The model structure is chosen as in (5) with $\beta_0 = 0$. As initial guess, the coefficients are fixed to 1, and the differentiation orders are multiple of 1.3.

The estimation synthesis is presented in Table I, where the first two columns define the model structure. Here the right model class is defined for $N = 3$ and $M = 1$ (marked in yellow in Table I). When the model is chosen in the right model class, the ***oosrivcf*** converged to the true parameters (up to to the noise effect):

$$G_1(p) = \frac{0.502}{0.502p^{2.796} + 1.503p^{1.197} + 1}. \tag{37}$$

As a comparison, the AIC criterion:

$$AIC = \log\left\{\left(1 + \frac{2n_\theta}{K}\right)\left\|\varepsilon\left(t, \hat{\theta}\right)\right\|^2\right\}, \tag{38}$$

and the quadratic criteria have been computed and displayed respectively on the last columns of Table I. Both criteria are not adapted for choosing the most suited fractional model class. Indeed, with the AIC criterion, multiple model structure seem to suit (the lowest criterion value are marked in green in Table I), and the quadratic criterion do not give the right model structure (marked in cyan in Table I).

		YIC criterion		AIC criterion	quadratic criterion
N	**M**	J_{YIC} (dB)	R_T^2	J_{AIC} (dB)	J (dB)
2	1	-6.886	0.802	-1.293	-7.031
3	1	-8.020	0.909	-1.632	-10.419
3	2	-4.021	0.909	-1.632	-10.420
4	1	-7.451	0.908	-1.627	-10.368
4	2	-6.174	0.909	-1.632	-10.416
4	3	-4.582	0.909	-1.629	-10.393
5	1	-2.661	0.908	-1.627	-10.369
5	2	-5.031	0.909	-1.629	-10.386
5	3	-4.574	0.909	-1.631	-10.412
5	4	-4.052	0.909	-1.631	-10.408

TABLE I

BEST MODEL STRUCTURES ESTIMATED WITH THE ***oosrivcf*** ALGORITHM ACCORDING TO YIC, AIC AND QUADRATIC CRITERIA.

The YIC and R_T^2 criteria have the advantage to directly give the most suited model structure (marked in red in Table I): the YIC should be the lowest and R_T^2 the highest.

IV. CONCLUSION

The well-known ***srivc*** algorithm extended to fractional models is used to estimate transfer function coefficients and it is combined to a gradient-based algorithm for fractional differentiation order estimation. This algorithm is presented as the ***oosrivcf*** algorithm, which combines simultaneously the transfer function coefficient and the differentiation order estimation. 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. In order to get the best model structure, two criteria are required. The R_T^2 criterion, recognized as the coefficient of determination based on the simulate error is not sufficient to avoid over-parametrization and identify a parsimonious model. However, the Young's information criterion YIC provides a measure of how well parameters are defined statistically: the more negative YIC, the better the definition. It may lead to underestimate model orders. Both criteria are inspected to find the orders for which R_T^2 is sufficiently high to indicate a good explanation of the data and the YIC is sufficiently low to indicate well-defined parameter estimates. The performance of the proposed algorithms has been evaluated on a simulation example and compared to the AIC and quadratic criteria.

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

[1] K.B. Oldham and J. Spanier. *The fractionnal calculus - Theory and Applications of Differentiation and Integration to Arbitrary Order*. Academic Press, New-York and London, 1974.

[2] S. Rodrigues, N. Munichandraiah, and A.-K. Shukla. A review of state of charge indication of batteries by means of A.C. impedance measurements. *Journal of Power Sources*, 87:12–20, 2000.

[3] J. Sabatier, M. Aoun, A. Oustaloup, G. Grégoire, F. Ragot, and P. Roy. Fractional system identification for lead acid battery state charge estimation. *Signal processing*, 86(10):2645–2657, 2006.

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

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

[6] P.C. Young. Parameter estimation for continuous-time models – a survey. *Automatica*, 17(1):23–29, 1981.

[7] P.C. Young. *Recursive estimation and Time-Series Analysis*. Springer-Verlag, Berlin, 2nd edition, 2011.

[8] T. Söderström and P. Stoica. *Instrumental variable methods for system identification*. Springer Verlag, New York, 1983.

[9] P.C. Young. Optimal IV identification and estimation of continuous-time TF models. In *15th World IFAC Congress*, Barcelona (Spain), 2002. IFAC, Elsevier.

[10] S. Victor, R. Malti, H. Garnier, and A. Oustaloup. Parameter and differentiation order estimation in fractional models. *Automatica*, 49(4), 2013.

[11] A. Oustaloup. *La dérivation non-entière*. Hermès, Paris, 1995.

[12] S.G. Samko, A.A. Kilbas, and O.I. Marichev. *Fractional integrals and derivatives: theory and applications*. Gordon and Breach Sc., 1993.

[13] M. Aoun, R. Malti, F. Levron, and A. Oustaloup. Numerical simulations of fractional systems: an overview of existing methods and improvements. *An Int. J. of Nonlinear Dynamics and Chaos in Engineering Systems. Special issue: Fractional Derivatives and Their Applications.*, 38(1-4):117–131, 2004.

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

[15] D. Matignon. Stability properties for generalized fractional differential systems. *ESAIM proceedings - Systèmes Différentiels Fractionnaires - Modèles, Méthodes et Applications*, 5, 1998.

[16] A. J. Jakeman and P. C. Young. Refined instrumental variable methods of time-series analysis: Part ii, multivariable systems. *International Journal of Control*, 29:621–644, 1979.

[17] P.C. Young, H. Garnier, and M. Gilson. Identification of continuous-time models from sampled data. In H. Garnier and L. Wang, editors, *Identification of continuous-time models from sampled data*, chapter Refined Instrumental Variable Identification of CT Hybrid Box-Jenkins Models, pages 91–131. Springer-Verlag, 2008.

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

[19] P.C. Young and A.J. Jakeman. Refined instrumental variable methods of time-series analysis: Part III, extensions. *International Journal of Control*, 31:741–764, 1980.

[20] R. Malti, S. Victor, and A. Oustaloup. Advances in system identification using fractional models. In ASME, editor, *Journal of Computational and Nonlinear Dynamics*, volume 3, pages 021401.1–021401.7, New York, USA, April 2008.

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

[22] L. Ljung. *System identification – Theory for the user*. Prentice-Hall, Upper Saddle River, N.J., USA, 2 edition, 1999.