Data-based continuous-time modelling of dynamic systems

Hugues Garnier

Data-based continuous-time modelling of dynamic systems

Hugues Garnier

***Abstract*— Data-based continuous-time model identification of continuous-time dynamic systems is a mature subject. In this contribution, we focus first on a refined instrumental variable method that yields parameter estimates with optimal statistical properties for hybrid continuous-time Box-Jenkins transfer function models. The second part of the paper describes further recent developments of this reliable estimation technique, including its extension to handle non-uniformly sampled data situation, closed-loop and nonlinear model identification. It also discusses how the recently developed methods are implemented in the CONTSID toolbox for Matlab and the advantages of these direct schemes to continuous-time model identification.**

I. INTRODUCTION

The identification of continuous-time (CT) models is a problem of considerable importance that has applications in virtually all disciplines of science. Early research on this topic focussed on identification of CT models from CT data (see e.g. [29], [30]). Subsequently, however, rapid developments in digital data acquisition and computers have resulted in attention being shifted to the identification of discrete-time (DT) models from sampled data, as documented in many books (see e.g [34], [24] and [15]). Much less attention has been devoted to CT modelling from DT data and many practitioners appear unaware that such alternative methods not only exist but may be better suited to their modelling problems.

In order to identify a continuous-time model from time-domain sampled data, two main time-domain approaches are possible. In the first, 'indirect' approach, a DT model is identified first using DT model identification methods, and this is then converted into a CT model using a standard algorithm for discrete to continuous-time conversion. In the second, 'direct' approach the CT model is identified directly from DT data. Direct data-based CT modelling is often incorrectly presented as being too complicated but, as we will see, the approaches are straightforward, reliable and have proven useful in many practical applications. These approaches have recently regained interest showing better performance than indirect approaches for both linear and nonlinear models, see *e.g.* [20], [21], [22], [12], [42]. The main motivations for identifying CT models directly from sampled data have been recently discussed in [7] (see also the Conclusions Section in this paper). Exhaustive reviews of direct estimation methods can be found in [33], [3], [5] and [22]. Amongst the available identification approaches for CT *input-output* models, the interest for *instrumental variable* (IV) methods has been growing in the last years [23], [34], [42], [28]. The main reason of this increasing interest is that IV methods offer similar performance as extended *least square* (LS) methods or other *prediction-error-minimization* (PEM) methods (see [21], [17]) and provide consistent results even for an imperfect noise structure which is the case in most practical applications. The IV schemes considered here present the major advantages over PEM methods to be much less sensitive to the initialisation stage (see [20], [21], [16]) . These IV approaches lead to optimal estimates in the linear time-invariant case if the system belongs to the model set defined. This paper concentrates on a reliable Instrumental Variable (IV)-based estimation method in particular, and presents the latest developments, including its use for closed-loop and nonlinear model identification.

H. Garnier is with Centre de Recherche en Automatique de Nancy (CRAN), Nancy-University, CNRS, BP 70239, 54506 Vandoeuvre-les-Nancy Cedex, France hugues.garnier@uhp-nancy.fr

II. RIVC FOR CT LINEAR MODELS

We focus on a statistically optimal method for the identification of continuous-time hybrid Box–Jenkins (BJ) transfer function models from discrete-time data [43]. Here, the model of the dynamic system is estimated in continuous-time, differential equation form, while the associated additive noise model is estimated as a discrete-time, autoregressive moving average (ARMA) process. This refined instrumental variable method for continuous-time systems (RIVC) was first developed in 1980 by Young and Jakeman [38] and its simplest embodiment, the simplified RIVC (SRIVC) method, has been used successfully for many years, demonstrating the advantages that this stochastic formulation of the continuous-time estimation problem provides in practical applications (see, *e.g.*, some recent such examples in [35], [41]).

However, the 'simplification' that characterises the name of the SRIVC method is the assumption, for the purposes of simplicity and algorithmic development, that the additive noise is purely white in form. Such an approach is optimal under this assumption and the inherent instrumental variable aspects of the resulting algorithm ensure that the parameter estimates are consistent and asymptotically unbiased in statistical terms, even if the noise happens to be coloured. However, the SRIVC estimates are not, in general, statistically efficient (minimum variance) in this situation because the prefilters are not designed to account for the colour in the noise process.

The hybrid RIVC estimation procedure follows logically from the refined instrumental variable (RIV) method for discrete-time models, first developed within a maximum likelihood (ML) context by Young in 1976 [32] and comprehensively evaluated by Young and Jakeman [37], [10], [34].

The RIV algorithm involves concurrent DT noise model estimation and uses this estimated noise model in the iterative-adaptive design of statistically optimal prefilters that effectively attenuate noise outside the passband of the system and prewhiten the noise remaining within the bandpass. Similarly motivated prefilters are utilised in the RIVC algorithm but they also provide a very convenient way of generating the prefiltered derivatives of the input and output variables, as required for CT model estimation.

The alternative hybrid form of the continuous-time transfer function model is considered here for two reasons. First, the approach is simple and straightforward: the theoretical and practical problems associated with the estimation of purely stochastic, continuous-time CAR or CARMA models are avoided by formulating the problem in this manner. Second, as pointed out above, one of the main functions of the noise estimation is to improve the statistical efficiency of the parameter estimation by introducing appropriately defined prefilters into the estimation procedure. And, as we shall see, this can be achieved adequately on the basis of hybrid prefilters defined by reference to discrete-time AR or ARMA noise models.

A. Problem Formulation

For simplicity of presentation, the formulation and solution of the CT estimation problem will be restricted to the case of a linear, single-input, single-output system. It is assumed that the input $u(t)$ and the noise-free output $x(t)$ are related by the following constant coefficient, differential-delay equation,

$$\begin{aligned}&x^{(n)}(t)+a_1^o x^{(n-1)}(t)+\cdots+a_n^o x(t)=\\&b_0^o u^{(m)}(t-\tau)+\cdots+b_m^o u(t-\tau)\end{aligned} \tag{1}$$

where $x^{(i)}(t)$ denotes the ith time derivative of the continuous-time signal $x(t)$ and τ is a pure time delay in time units. This is often assumed to be an integer number related to the sampling time: *i.e.*, $\tau = n_k T_s$ but this is not essential: in this CT environment, 'fractional' time delays can be introduced if required (*e.g.*, see [19], [40]). For simplicity, the time delay will not be considered in the following analysis but it can be accommodated straightforwardly if identified from the data. Equation (1) can also be written in the following compact transfer function (TF) form,

$$x(t)=G_o(p)u(t)=\frac{B_o(p)}{A_o(p)}u(t) \tag{2}$$

with

$$B_o(p)=b_0^o p^m+b_1^o p^{m-1}+\cdots+b_m^o, \tag{2a}$$

$$A_o(p)=p^n+a_1^o p^{n-1}+\cdots+a_n^o, \quad n\ge m \tag{2b}$$

where $x(t)$ is the deterministic output of the system; p is the differential operator, *i.e.*, $p^i x(t)=\frac{\mathrm{d}^i x(t)}{\mathrm{d}t^i}$; $B_o(p)$ and $A_o(p)$ are assumed to be coprime; and the system is asymptotically stable. It is assumed that the input signal $\{u(t), t_1 < t < t_N\}$ is applied to the system and this gives rise to an output signal $\{x(t), t_1 < t < t_N\}$.

In order to obtain high-quality statistical estimation results, it is vital to consider the inevitable errors that will affect the measured output signal. It is assumed here that $x(t)$ is corrupted by an additive, coloured measurement noise $\xi(t)$, so that the complete equation for the data-generating system, denoted by $\mathcal{S}$, can be written in the form,

$$\mathcal{S}: y(t)=G_o(p)u(t)+H_o(p)e_o(t) \tag{3}$$

or, in the alternative decomposed form that is more appropriate in the present context

$$\mathcal{S}\begin{cases}x(t)=G_o(p)u(t)\\ \xi(t)=H_o(p)e_o(t)\\ y(t)=x(t)+\xi(t)\end{cases} \tag{4}$$

where $H_o(p)$ is stable and invertible, while $e_o(t)$ is a zero-mean, continuous-time white noise source, which is assumed to be uncorrelated with the input $u(t)$. Finally, if the additive coloured noise $\xi(t)$ has rational spectral density, then a suitable parametric representation is the following continuous-time, autoregressive moving average (CARMA) model

$$\xi(t)=H_o(p)e_o(t)=\frac{C_o(p)}{D_o(p)}e_o(t) \tag{5}$$

where $C_o(p)$ and $D_o(p)$ are suitably defined polynomials in the p operator.

Of course, in most practical situations, the input and output signals $u(t)$ and $y(t)$ will be sampled in discrete time. In the case of uniform sampling, at a constant sampling interval T_s, these sampled signals will be denoted by $u(t_k)$ and $y(t_k)$ and the output observation equation then takes the form,

$$y(t_k)=x(t_k)+\xi(t_k) \qquad k=1,\cdots N \tag{6}$$

where $x(t_k)$ is the sampled value of the unobserved, noise-free output $x(t)$. The objective is then to identify a suitable model structure for (4) and estimate the parameters that characterise this structure, based on these sampled input and output data $Z^N=\{u(t_k); y(t_k)\}_{k=1}^N$.

Given the discrete-time, sampled nature of the data, an obvious assumption is that the discrete-time, coloured noise associated with the sampled output measurement $y(t_k)$ has rational spectral density and so can be represented by a discrete-time ARMA model. The model set to be identified and estimated, as denoted by $\mathcal{M}$ with system ($\mathcal{G}$) and noise ($\mathcal{H}$) models parameterised independently, then takes the form,

$$\mathcal{M} \quad : \{G(p,\boldsymbol{\rho}), H(q,\boldsymbol{\eta})\} \tag{7}$$

where $\boldsymbol{\rho}$ and $\boldsymbol{\eta}$ are parameter vectors that characterise the system and noise models, respectively. In particular, the system model is formulated in continuous-time terms

$$\mathcal{G}: G(p,\boldsymbol{\rho})=\frac{B(p,\boldsymbol{\rho})}{A(p,\boldsymbol{\rho})}=\frac{b_0p^m+b_1p^{m-1}+\cdots+b_m}{p^n+a_1p^{n-1}\cdots+a_n} \tag{8}$$

and the associated model parameters are stacked columnwise in the parameter vector,

$$\boldsymbol{\rho}=\begin{bmatrix}a_1 & \cdots & a_n & b_0 & \cdots & b_m\end{bmatrix}^T\in\mathbb{R}^{n+m+1} \tag{9}$$

while the noise model is in discrete-time form

$$\mathcal{H} : H(q, \boldsymbol{\eta}) = \frac{C(q^{-1}, \boldsymbol{\eta})}{D(q^{-1}, \boldsymbol{\eta})} = \frac{1 + c_1 q^{-1} + \cdots + c_q q^{-q}}{1 + d_1 q^{-1} + \cdots + d_p q^{-p}} \tag{10}$$

where q^{-r} is the backward shift operator, *i.e.*, $q^{-r} y(t_k) = y(t_{k-r})$ and the associated model parameters are stacked columnwise in the parameter vector,

$$\boldsymbol{\eta} = \begin{bmatrix} c_1 & \cdots & c_q & d_1 & \cdots & d_p \end{bmatrix}^T \in \mathbb{R}^{p+q} \tag{11}$$

Consequently, the noise TF takes the usual ARMA model form

$$\xi(t_k) = \frac{C(q^{-1}, \boldsymbol{\eta})}{D(q^{-1}, \boldsymbol{\eta})} e(t_k) \qquad e(t_k) \sim \mathcal{N}(0, \sigma^2) \tag{12}$$

where, as shown, $e(t_k)$ is a zero-mean, normally distributed, discrete-time white noise sequence.

The structure $\mathcal{S}$ does not specify any common factors in the plant (G_o) and noise (H_o) components, so that these models can be parameterised independently. More formally, there exists the following decomposition of the parameter vector $\boldsymbol{\theta}$ for the whole hybrid model,

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\rho} \\ \boldsymbol{\eta} \end{pmatrix} \tag{13}$$

such that the model equations can be written in the form

$$\mathcal{M} \begin{cases} x(t) = G(p, \boldsymbol{\rho}) u(t) \\ \xi(t_k) = H(q, \boldsymbol{\eta}) e(t_k) \\ y(t_k) = x(t_k) + \xi(t_k) \end{cases} \tag{14}$$

This model is considered as a hybrid Box–Jenkins model because of its close relationship to the DT model considered in great detail by Box and Jenkins in their seminal book on time-series analysis, forecasting and control [1] and used as the basis for the development of the original RIVC algorithm [38]. Alternatively, the model can be written in the following vector terms

$$\mathcal{M} \begin{cases} x^{(n)}(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\rho} \\ \xi(t_k) = \boldsymbol{\psi}^T(t_k) \boldsymbol{\eta} + e(t_k) \\ y(t_k) = x(t_k) + \xi(t_k) \end{cases} \tag{15}$$

where,

$$\boldsymbol{\varphi}^T(t) = \left[-x^{(n-1)}(t) \cdots - x(t) \; u^{(m)}(t) \cdots u(t) \right] \tag{15a}$$

$$\boldsymbol{\psi}^T(t_k) = [-\xi(t_{k-1}) \cdots - \xi(t_{k-p}) \; e(t_{k-1}) \cdots e(t_{k-q})] \tag{15b}$$

For the purposes of identification, the order of this single-input model (with the pure time delay τ now added for completeness) is denoted by $[n \; m \; \tau \; p \; q]$ and the complete identification problem can now be stated as follows:
Based on N uniformly sampled measurements of the input and output, $Z^N = \{u(t_k); y(t_k)\}_{k=1}^N$, identify the orders n, m, p and q of the polynomials in the system and noise TF models, as well as any pure time delay τ, and estimate the parameter vector $\boldsymbol{\theta}$ in (13) whose parameters characterise these polynomials.

B. Optimal RIVC Estimation: Theoretical Motivation

The RIVC algorithm derives from the RIV algorithm for DT systems. This was evolved by converting the maximum likelihood (ML) estimation equations to a pseudo-linear form [25] involving optimal prefilters [32], [38], [34]. A similar analysis can be utilised in the present situation because the problem is very similar, in both algebraic and statistical terms. However, to conserve space, the discussion here will be restricted to a simpler development of the RIVC algorithm and we leave the interested reader to consult with these earlier references for details of the ML analysis.

1) The Hybrid Box–Jenkins Estimation Model: Following the usual *prediction error minimisation* (PEM) approach in the present hybrid situation (which is ML estimation because of the Gaussian assumptions on $e(t_k)$), a suitable error function $\varepsilon(t_k)$, at the kth sampling instant, is given by,

$$\varepsilon(t_k) = \frac{D(q^{-1}, \boldsymbol{\eta})}{C(q^{-1}, \boldsymbol{\eta})} \left\{ y(t_k) - \frac{B(p, \boldsymbol{\rho})}{A(p, \boldsymbol{\rho})} u(t_k) \right\}$$

which can be written as,

$$\varepsilon(t_k) = \frac{D(q^{-1}, \boldsymbol{\eta})}{C(q^{-1}, \boldsymbol{\eta})} \left\{ \frac{1}{A(p, \boldsymbol{\rho})} [A(p, \boldsymbol{\rho}) y(t_k) - B(p, \boldsymbol{\rho}) u(t_k)] \right\} \tag{16}$$

where the discrete-time prefilter $D(q^{-1}, \boldsymbol{\eta})/C(q^{-1}, \boldsymbol{\eta})$ will be recognised as the inverse of the ARMA(p,q) noise model. Note that in these equations, we are mixing discrete and continuous-time operators somewhat informally in order to indicate the hydrid computational nature of the estimation problem being considered here. Thus, operations such as,

$$\frac{B(p, \boldsymbol{\rho})}{A(p, \boldsymbol{\rho})} u(t_k)$$

imply that the input variable $u(t_k)$ is interpolated in some manner. This is to allow for the inter-sample behaviour that is not available from the sampled data and so has to be inferred in order to allow for the continuous-time numerical integration of the associated differential equations.

Minimisation of a least squares criterion function in $\varepsilon(t_k)$, measured at the sampling instants, provides the basis for stochastic estimation. However, since the polynomial operators commute in this linear case, (16) can be considered in the alternative form,

$$\varepsilon(t_k) = A(p, \boldsymbol{\rho}) y_{\text{f}}(t_k) - B(p, \boldsymbol{\rho}) u_{\text{f}}(t_k) \tag{17}$$

where $y_{\text{f}}(t_k)$ and $u_{\text{f}}(t_k)$ represent the *sampled* outputs of the complete hybrid prefiltering operation involving the *continuous-time* filtering operations using the filter

$$f_{\text{c}}(p, \boldsymbol{\rho}) = \frac{1}{A(p, \boldsymbol{\rho})} \tag{18}$$

as well as *discrete-time* filtering operations, using the inverse noise model filter

$$f_{\text{d}}(q^{-1}, \boldsymbol{\eta}) = \frac{D(q^{-1}, \boldsymbol{\eta})}{C(q^{-1}, \boldsymbol{\eta})} \tag{19}$$

The associated, linear-in-the-parameters estimation model then takes the form

$$y_{\text{f}}^{(n)}(t_k) = \boldsymbol{\varphi}_{\text{f}}^T(t_k) \boldsymbol{\rho} + \eta(t_k) \tag{20}$$

where,

$$\boldsymbol{\varphi}_{\mathrm{f}}^{T}(t_k) = [-y_{\mathrm{f}}^{(n-1)}(t_k)\cdots\ -y_{\mathrm{f}}(t)\ u_{\mathrm{f}}^{(m)}(t_k)\cdots u_{\mathrm{f}}(t_k)] \tag{21}$$

and $\eta(t_k)$ is the continuous-time noise signal $\eta(t) = A(p,\boldsymbol{\rho})\xi(t)$ sampled at the kth sampling instant.

2) RIVC Estimation: Optimal methods of IV estimation (see, *e.g.*, [32], [23]) normally involve an iterative (or relaxation) algorithm in which, at each iteration, the 'auxiliary model' used to generate the instrumental variables, as well as the associated prefilters, are updated, based on the parameter estimates obtained at the previous iteration. Let us consider, therefore, the jth iteration where we have access to the estimate,

$$\hat{\boldsymbol{\theta}}^{j-1} = \begin{pmatrix}\hat{\boldsymbol{\rho}}^{j-1}\\ \hat{\boldsymbol{\eta}}^{j-1}\end{pmatrix} \tag{22}$$

obtained previously at iteration $j-1$. The most important aspect of optimal IV estimation is the definition of an optimal instrumental variable. In the present context, this is generated from the output of the continuous-time auxiliary model,

$$\hat{x}(t,\hat{\boldsymbol{\rho}}^{j-1}) = G(p,\hat{\boldsymbol{\rho}}^{j-1})u(t) \tag{23}$$

which is prefiltered in the same hybrid manner as the other variables. The associated optimal IV vector $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)$, is then an estimate of the noise-free version of the vector $\boldsymbol{\varphi}_{\mathrm{f}}(t_k)$ in (21) and is defined as follows

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k) = \left[-\hat{x}_{\mathrm{f}}^{(n-1)}(t_k)\cdots\ -\hat{x}_{\mathrm{f}}(t_k)\ u_{\mathrm{f}}^{(m)}(t_k)\ \cdots\ u_{\mathrm{f}}(t_k)\right]^{T} \tag{24}$$

where it should be noted that

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k) = \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k,\hat{\boldsymbol{\rho}}^{j-1},\hat{\boldsymbol{\eta}}^{j}) \tag{25}$$

because the instrumental variables are now prefiltered and so are a function of both the system parameter estimates at the previous iteration and the most recent noise model parameter estimates. For simplicity, however, these additional arguments will be omitted in the subsequent analysis. Note also that the noise-free version of the vector $\boldsymbol{\varphi}_{\mathrm{f}}(t_k)$ in (21), which we will define as follows,

$$\mathring{\boldsymbol{\varphi}}_{\mathrm{f}}^{T}(t_k) = \left[-x_{\mathrm{f}}^{(n-1)}(t_k)\cdots\ -x_{\mathrm{f}}(t_k)\ u_{\mathrm{f}}^{(m)}(t_k)\ \cdots\ u_{\mathrm{f}}(t_k)\right] \tag{26}$$

where $x(t) = G_o(p)u(t)$, is referred to in Section II-D when considering the statistical properties of the optimal IV parameter estimates.

The IV optimisation problem can now be stated in the form

$$\hat{\boldsymbol{\rho}}^{j}(N) = \arg\min_{\boldsymbol{\rho}} \left\| \left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)\boldsymbol{\varphi}_{\mathrm{f}}^{T}(t_k)\right]\boldsymbol{\rho} - \left[\frac{1}{N}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)y_{\mathrm{f}}^{(n)}(t_k)\right]\right\|_{\mathbf{Q}}^{2} \tag{27}$$

where $\|\mathbf{x}\|^{2} = \mathbf{x}^{T}\mathbf{Q}\mathbf{x}$ and $\mathbf{Q} = \mathbf{I}$. This results in the well-known solution of the IV estimation (IV normal) equations

$$\hat{\boldsymbol{\rho}}^{j}(N) = \left[\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)\boldsymbol{\varphi}_{\mathrm{f}}^{T}(t_k)\right]^{-1}\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)y_{\mathrm{f}}^{(n)}(t_k) \tag{28}$$

where the $\hat{\boldsymbol{\rho}}^{j}(N)$ is the IV estimate of the system model parameter vector at the jth iteration based on the appropriately prefiltered input/output data $Z^N = \{u(t_k); y(t_k)\}_{k=1}^{N}$.

As regards the hybrid prefiltering, it will be noted from (25) that this involves the inverse noise model parameters $\hat{\boldsymbol{\eta}}^{j}$ obtained at the current jth iteration. This is because, given $\hat{\boldsymbol{\rho}}^{j-1}$, an estimate of the sampled noise signal $\xi(t_k)$, at the jth iteration, is obtained by subtracting the sampled output of the auxiliary model equation (23) from the measured output $y(t_k)$, *i.e.*,

$$\hat{\xi}(t_k) = y(t_k) - \hat{x}(t,\hat{\boldsymbol{\rho}}^{j-1}) \tag{29}$$

This estimate provides the basis for the estimation of the noise model parameter vector $\boldsymbol{\eta}^{j}$, using whatever ARMA model estimation algorithm is selected for this task.

C. The RIVC and SRIVC Algorithms

The iterative RIVC and SRIVC algorithms follow directly from the RIV and SRIV algorithms for DT systems (*e.g.*, [34]). This section summarises both algorithms.

1) The RIVC Algorithm: Bearing the analysis of the previous subsection II-B.2 in mind, the main steps in the RIVC algorithm are as follows:

Step 1. Initialisation: generate an initial estimate of the TF model parameter vector $\hat{\boldsymbol{\rho}}^{o}$ using the simplified RIVC (SRIVC) algorithm (see subsection II-C.2) and use this to define the initial CT prefilter $f_{\mathrm{c}}(p,\hat{\boldsymbol{\rho}}^{o})$.

Step 2. Iterative estimation.

for $j = 1 : convergence$

(i) *Generate the IV series $\hat{x}(t,\hat{\boldsymbol{\rho}}^{j-1})$ using the auxiliary model built up from the estimated polynomials $A(p,\hat{\boldsymbol{\rho}}^{j-1})$ and $B(p,\hat{\boldsymbol{\rho}}^{j-1})$ based on $\hat{\boldsymbol{\rho}}^{j-1}$ at the previous $(j-1)$th iteration.*

(ii) *Prefilter the input $u(t_k)$, output $y(t_k)$ and instrumental variable $\hat{x}(t,\hat{\boldsymbol{\rho}}^{j-1})$ by the continuous-time filter $f_c(p,\hat{\boldsymbol{\rho}}^{j-1})$ in order to generate the filtered derivatives of these variables.*

(iii) *Obtain an optimal estimate of the noise model parameter vector $\hat{\boldsymbol{\eta}}^{j}$ based on the estimated noise sequence $\hat{\xi}(t_k)$ from* (29)*, using a selected ARMA estimation algorithm.*

(iv) *Sample the filtered derivative signals at the discrete-time sampling interval T_s and prefilter these by the discrete-time filter $f_d(q^{-1},\hat{\boldsymbol{\eta}}^{j})$, in order to define all the required elements in the data vector $\boldsymbol{\varphi}_f(t_k)$, the IV vector $\hat{\boldsymbol{\varphi}}_f(t_k)$ and the nth-order filtered derivative $y_f^{(n)}(t_k)$.*

(v) *Based on these prefiltered data, generate the latest estimate $\hat{\boldsymbol{\rho}}^{j}$ of the system model parameter vector using the en bloc IV solution* (28)*, or its recursive equivalent. Together with the estimate $\hat{\boldsymbol{\eta}}^{j}$ of the noise model parameter estimate from (iii), this provides the estimate $\hat{\boldsymbol{\theta}}^{j}$ of the composite parameter vector at the jth iteration.*

end

Step 3. After the convergence of the iterations is complete, compute the estimated parametric error covariance matrix $\hat{\mathbf{P}}_{\rho}$, associated with the converged estimate

$\hat{\boldsymbol{\rho}}$ of the system model parameter vector, from the expression (see Section II-D),

$$\hat{\mathbf{P}}_\rho \;=\; \hat{\sigma}^2\left[\sum_{k=1}^{N}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^T(t_k)\right]^{-1} \tag{30}$$

where $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t_k)$ is the IV vector obtained at convergence and $\hat{\sigma}^2$ is the estimated residual variance.

2) The SRIVC Algorithm: It will be noted that the above formulation of the RIVC estimation problem is considerably simplified if it is assumed that the additive noise is white, *i.e.*, $C(q^{-1},\boldsymbol{\eta}) = D(q^{-1},\boldsymbol{\eta}) = 1$. In this case, simplified RIVC (SRIVC) estimation involves only the parameters in the $A(p,\boldsymbol{\rho})$ and $B(p,\boldsymbol{\rho})$ polynomials and the prefiltering only involves the continuous-time prefilter $f_{\mathrm{c}}(p,\boldsymbol{\rho}) = 1/A(p,\boldsymbol{\rho})$. Consequently, the main steps in the SRIVC algorithm are the same as those in the RIVC algorithm, except that the noise model estimation and subsequent discrete-time prefiltering in steps (ii) and (iii) of the iterative procedure are no longer required and are omitted.

It is worth noting that the RIVC algorithm has a much longer computation time than the SRIVC algorithm. As a result, it is advantageous to use the SRIVC algorithm for initial model order identification and only employ the full RIVC algorithm in those situations where the theoretical assumptions are satisfied and it is essential to have the most efficient parameter estimates and better estimates of the uncertainty on the parameters. For day-to-day usage, the SRIVC algorithm provides a quick and reliable approach to continuous-time model identification and estimation.

D. Theoretical Background and Statistical Properties of the RIVC Estimates

The motivational arguments presented in Section II-B suggest that, upon convergence, the RIVC parameter estimates will possess the optimal statistical properties of consistency and asymptotic efficiency when the additive noise has a Gaussian normal probability distribution and rational spectral density. This section presents more formal analysis to verify further the optimality of the estimates and confirm the asymptotic independence of the system and noise model parameter estimates.

1) Optimality of RIVC Estimation: In the control and systems literature, optimal IV estimation is usually considered in relation to the so-called 'extended IV' approach to estimation, as developed for the DT case [23]. A similar approach can be applied in the present CT case by re-writing the IV optimisation equation (27) in the following alternative form that explicitly reveals a continuous-time prefilter $f(p)$

$$\hat{\boldsymbol{\rho}}(N) = \arg\min_{\rho} \|[\,\frac{1}{N}\sum_{k=1}^{N}\boldsymbol{\zeta}_{\mathrm{f}}(t_k)f(p)\boldsymbol{\varphi}^T(t_k)\hat{\boldsymbol{\rho}} - \left[\frac{1}{N}\sum_{k=1}^{N}\boldsymbol{\zeta}_{\mathrm{f}}(t_k)f(p)y^{(n)}(t_k)\right]_Q^2 \tag{31}$$

where $f(p)$ is the stable prefilter, $\boldsymbol{\zeta}_{\mathrm{f}}(t_k)$ is the prefiltered instrumental vector $\boldsymbol{\zeta}_{\mathrm{f}}(t_k) = f(p)\boldsymbol{\zeta}(t_k)$ and $\mathbf{Q}$ is a positive-definite matrix. By definition, when $G_o \in \mathcal{G}$, the extended IV estimate provides a consistent estimate under the following two conditions

$$\begin{cases}\bar{\mathsf{E}}\{\boldsymbol{\zeta}_{\mathrm{f}}(t_k)f(p)\boldsymbol{\varphi}^T(t_k)\} \text{ is non-singular,}\\ \bar{\mathsf{E}}\{\boldsymbol{\zeta}_{\mathrm{f}}(t_k)f(p)\boldsymbol{\xi}(t_k)\} = 0\end{cases} \tag{32}$$

Clearly, the selection of the instrumental variable vector $\boldsymbol{\zeta}_{\mathrm{f}}(t_k)$, the weighting matrix $\mathbf{Q}$ and the prefilter $f(p)$ may have a considerable effect on the covariance matrix $\mathbf{P}_\theta$ produced by the IV estimation algorithm.

In the open-loop situation, the Cramér–Rao lower bound on P_θ for any unbiased identification method (*e.g.*, [23], [15]) defines the optimal solution. In this regard, it has been shown that the minimum value of the covariance matrix $\mathbf{P}_\theta$, as a function of the design variables $\boldsymbol{\zeta}_{\mathrm{f}}(t_k)$, $f(p)$ and $\mathbf{Q}$, exists and is given by

$$\mathbf{P}_\theta \geq \mathbf{P}_\theta^{opt}$$

with

$$\mathbf{P}_\theta^{opt} = [\bar{\mathsf{E}}\{\mathring{\boldsymbol{\zeta}}_{\mathrm{f}}(t_k)\mathring{\boldsymbol{\zeta}}_{\mathrm{f}}^T(t_k)\}]^{-1} \tag{33}$$

where $\mathring{\boldsymbol{\zeta}}_{\mathrm{f}}(t_k)$ is the optimally prefiltered IV vector, with the associated design variables defined as

$$\mathbf{Q} = \mathbf{I}, \tag{34a}$$

$$f(p) = \frac{1}{H_o(p)A_o(p)} = \frac{D_o(p)}{C_o(p)A_o(p)}, \tag{34b}$$

$$\mathring{\boldsymbol{\zeta}}(t_k) = \left[-x^{(n-1)}(t_k)\ \cdots\ -x(t_k)\ u^{(m)}(t_k)\ \cdots\ u(t_k)\right]^T \tag{34c}$$

so that,

$$\mathring{\boldsymbol{\zeta}}_{\mathrm{f}}(t_k) = f(p)\mathring{\boldsymbol{\zeta}}(t_k) \tag{35}$$

which will be recognised as the noise-free, prefiltered vector $\mathring{\boldsymbol{\varphi}}_{\mathrm{f}}^T(t_k)$ defined earlier in (26).

2) Comments:

- Not surprisingly, the above analysis justifies the RIVC algorithmic design that iteratively updates those aspects of the theoretical solution that are not known *a priori*: in this case, the unknown model polynomials and the noise-free output of the system that is, of course, the source of the instrumental variables. If it is assumed that, in all identifiable situations, the RIVC algorithm converges in the sense that $\hat{\boldsymbol{\rho}} \Rightarrow \boldsymbol{\rho}$ and $\hat{\boldsymbol{\eta}} \Rightarrow \boldsymbol{\eta}$, then the RIVC estimates will be consistent and asymptotically efficient.
- The optimal filter $f(p)$ in (34b) is formulated in CT terms. In the proposed RIVC algorithm, this filter takes a hybrid form, as discussed in the previous sections. One very important aspect of TF modelling is the identification of the model structure: *i.e.*, the degrees n, m, p, and q of the model polynomials and any associated pure time delay τ. A model order selection method associated to the SRIVC model estimation method allows the user to automatically search over a whole range of different model orders. Two statistical measures are then used to help to user choose the best model structure (see Subsection II.E and [43].

- Both RIVC/SRIVC routines are avalaible in the CONTSID (see Section IV below) and CAPTAIN[1] Toolboxes for MATLAB®.

E. Model Order Identification

One very important aspect of TF modelling is the identification of the model structure: *i.e.*, the degrees n, m, p, and q of the model polynomials and any associated pure time delay τ. One statistical measure that is useful in this regard is the coefficient of determination R_T^2, defined as follows

$$R_T^2 = 1 - \frac{\sigma_{\hat{\xi}}^2}{\sigma_y^2} \tag{36}$$

where $\sigma_{\hat{\xi}}^2$ is the variance of the estimated noise $\hat{\xi}(t_k)$ and σ_y^2 is the variance of the measured output $y(t_k)$. R_T^2 is clearly a normalised measure of how much of the output variance is explained by the deterministic system part of the estimated model. However, it is well known that this measure, on its own, is not sufficient to avoid over-parametrisation and identify a parsimonious model, so that other model order identification statistics are required. In this regard, because the SRIVC and RIVC methods exploit optimal instrumental variable methodology, they are able to utilise the special properties of the instrumental product matrix (IPM) [39]; in particular, the YIC statistic [34] which is defined as follows

$$\text{YIC} = \log_e \frac{\hat{\sigma}^2}{\sigma_y^2} \; + \; \log_e\{\text{NEVN}\}; \quad \text{NEVN} = \frac{1}{n_\theta} \sum_{i=1}^{n_\theta} \frac{\hat{p}_{ii}}{\hat{\theta}_i^2} \tag{37}$$

Here, $n_\theta = n + m + p + q + 1$ is the number of estimated parameters; $\hat{p}_{ii}$ is the ith diagonal element of the block-diagonal covariance matrix $\mathbf{P}_\theta$, where,

$$\mathbf{P}_\theta = \begin{pmatrix} \mathbf{P}_\rho & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_\eta \end{pmatrix} \tag{38}$$

and so is an estimate of the variance of the estimated uncertainty on the ith parameter estimate. $\hat{\theta}_i^2$ is the square of the ith parameter estimate in the $\boldsymbol{\theta}$ vector, so that ratio $\hat{p}_{ii}/\hat{\theta}_i^2$ is a normalised measure of the uncertainty on the ith parameter estimate.

From the definition of R_T^2, we see that the first term in the YIC is simply a relative measure of how well the model explains the data: the smaller the model residuals the more negative the term becomes. The normalised error variance norm (NEVN) term, on the other hand, provides a measure of the conditioning of the IPM, which needs to be inverted when the IV normal equations are solved (see, *e.g.*, [34]): if the model is overparameterised, then it can be shown that the IPM will tend to singularity and, because of its ill-conditioning, the elements of its inverse (in the form here of the covariance matrix $\mathbf{P}_\theta$) will increase in value, often by several orders of magnitude. When this happens, the second term in the YIC tends to dominate the criterion function, indicating over-parametrisation.

[1] `http://www.es.lancs.ac.uk/cres/captain/`

It is important to note that, based on practical experience, the YIC is normally best considered during SRIVC identification, which is much less computationally intensive than RIVC identification, so allowing for much faster investigation of the model order range selected by the user. In this situation, n_θ is replaced by $n_\rho = n + m + 1$ the $\hat{p}_{ii}$ are obtained by reference to the covariance matrix $\mathbf{P}_\rho$.

Although heuristic, the YIC has proven very useful in practical identification terms. It should not, however, be used as a sole arbiter of model order: rather the combination of R_T^2 and YIC provides an indication of the best parsimonious models that can be evaluated by other standard statistical measures (*e.g.*, the auto and partial autocorrelation of the model residuals, the cross-correlation of the residuals with the input signal $u(t_k)$, *etc.*). Also, the physical interpretation of the model can often provide valuable information on the model adequacy: for instance, a model with complex eigenvalues caused by overparametrisation may prove incompatible with the non-oscillatory nature of the physical system under study.

III. Latest developments for the RIVC method

Recent developments aimed at extending the RIVC method to handle wider practical situations in order to enhance the application field of direct CT model identification.

A. Multiple-input Systems

It is clearly straightforward to extend the RIVC/SRIVC methods to the multiple-input situation if the TF denominator is common to all input channels. The situation is not so straightforward in the case where there are different denominator polynomials for each input channel. However, following the RIV approach for DT systems [9], the algorithms can be extended to handle this situation [4]: indeed, the current version of RIVC in the CONTSID Toolbox provides this option.

B. Non-uniformly Sampled Data

One advantage of the SRIVC approach to continuous-time modelling is that it can be based on irregularly sampled data and can handle 'fractional' pure time delays. The current implementation of the SRIVC algorithm in the CONTSID Toolbox can handle irregularly sampled data. However, the RIVC algorithm has not yet been upgraded in this regard because it requires additional interpolation and re-sampling in order to generate a regularly sampled series for the ARMA noise model estimation parts of the algorithm.

C. Closed-loop Model Identification

Provided there is an external command input signal, the identification and estimation of a system within a closed automatic control loop has always been straightforward when using IV estimation methodology [31], [8]. In the case of the RIVC/SRIVC algorithms, a two-stage approach, such as that used by Van den Hof [27] for discrete-time systems, is the most effective, since it does not require prior knowledge of the control system. Recent research [44] has shown that a modification of this approach employing the SRIVC algorithm (rather than the FIR model estimation used by Van den Hof) for estimating the control input signal, followed by full RIVC estimation of the system, based on this estimated control input, works extremely well.

D. Hammerstein and LPV Model Identification

Direct identification of CT nonlinear models is still an unmature subject. This section discusses briefly the extension of the RIVC method for the identification of Hammerstein and LPV CT Box–Jenkins models. In the case of Hammerstein hybrid BJ model, the nonlinear function $f(.)$ is assumed to be a sum of known basis functions $\gamma_1, \gamma_2, \ldots, \gamma_l$ given as:

$$\bar{u}(t) = \sum_{i=1}^{l} \alpha_i \gamma_i(u(t)) \text{ with } \alpha_1 = 1. \quad (39)$$

The hybrid CT BJ Hammerstein model is described by the following input-output relationship:

$$\begin{cases} x(t) = G(p)\bar{u}(t) \\ \xi(t_k) = H(q^{-1})e(t_k), \\ y(t_k) = x(t_k) + \xi(t_k), \end{cases} \quad (40)$$

where

$$G(p) = \frac{B(p)}{A(p)}. \quad (41)$$

where the coloured noise associated with the sampled output measurement $y(t_k)$ has rational spectral density and can be represented by a discrete-time autoregressive moving average ARMA model:

$$\xi(t_k) = H(q^{-1})e(t_k) = \frac{C(q^{-1})}{D(q^{-1})}e(t_k) \quad (42)$$

The RIVC method has very recently been extended to estimate the parameters of such CT hybrid BJ Hammerstein models [12], [13].

So called Linear Parameter Varying (LPV) models have been the subject of recent interest. The RIVC approach has recently been extended to estimate CT LPV input/output models [14].

IV. SOFTWARE ASPECTS - THE CONTSID TOOLBOX

The field of system identification is an extensive and versatile area. It is easy to get confused by the vast number of approaches and variants of methods available. We have seen so far that direct continuous-time model identification from sampled data is now a mature subject and it is important to package the identification tools in a user-friendly way. An attempt to do that was carried out with the CONTinuous-time System IDentification (CONTSID) toolbox for MATLAB®. The CONTSID toolbox was first released in 1999 [2]. It has gone through several updates. The key features of the CONTSID toolbox are [6]:

- it supports most of the time-domain methods developed over the last thirty years [3] for identifying linear dynamic continuous-time parametric models from measured input/output sampled data;
- it provides transfer function and state-space model identification methods for single-input single-output (SISO) and multiple-input multiple-output (MIMO) systems, including both traditional and more recent approaches;
- it can handle mild irregularly sampled data in a straightforward way;
- it may be seen as an add-on to the system identification (SID) toolbox for MATLAB®. To facilitate its use, it has been given a similar setup to the SID toolbox;
- it provides a flexible graphical user interface (GUI) that lets the user analyse the experimental data, identify and evaluate models in an easy way;
- It can be freely downloaded from `http://www.cran.uhp-nancy.fr/contsid/`

The latest version of the CONTSID toolbox has the following three major additions:

- it supports errors-in-variables CT transfer function model identification [18], [26];
- it provides routines to estimate linear CT transfer function model in closed loop [8], [44];
- it includes methods to identify nonlinear CT Hammerstein models [12], [13].

V. CONCLUSIONS

This paper has first described the full RIVC algorithm for identifying hybrid Box–Jenkins transfer function models for linear, continuous-time systems from discrete-time, sampled data. The latest developments of the RIVC approach for nonuniformly sampled data, closed-loop identification as well as for nonlinear Hammerstein and LPV model identification have also been briefly discussed.

It is felt that continuous-time model identification, based on a stochastic formulation of the transfer function estimation problem, provides a theoretically elegant and practically useful approach to the modelling of stochastic dynamic systems from sampled data.

It is an approach that has many advantages in scientific terms since it provides differential equation models that conform with models used in most scientific research, where conservation equations are normally formulated in terms of differential equations. It is also a model defined by a unique set of parameter values that are not dependent on the sampling interval, so eliminating the need for conversion from discrete to continuous time that is an essential element of indirect approaches to estimation based on discrete-time model estimation. These direct continuous-time model identification methods have proven to be particularly well suited in the case of mild non-uniformly sampled data, dominant system modes with widely different natural frequencies (stiff systems), fast sampled data, or when the input does not respect the zero-order hold assumption. Finally but not the least, these direct data-based CT modelling methods have proven successful in many practical applications and are available as user-friendly and computationally efficient algorithms in the CONTSID toolbox for Matlab™.

VI. ACKNOWLEDGMENTS

I am extremely grateful to Professor Peter Young for teaching me the basics of the Refined Instrumental Variable estimation concept. I would like also gratefully acknowledge the contributions of all of my students in the developments presented in the paper: Michel Mensler, Marion Gilson, Eric Huselstein, Stephane Thil, Damien Kuss and Vincent Laurain.

REFERENCES

[1] G.E.P. Box and G.M. Jenkins. *Time Series Analysis Forecasting and Control*. Holden-Day: San Francisco, 1970.

[2] H. Garnier and M. Mensler. CONTSID: a continuous-time system identification toolbox for Matlab. *5th European Control Conference (ECC'99)*, Karlsruhe (Germany), September 1999.

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

[4] H. Garnier, M. Gilson, P.C. Young, and E. Huselstein. An optimal IV technique for identifying continuous-time transfer function model of multiple input systems. *Control Engineering Practice*, 15(4):471–486, 2007.

[5] H. Garnier and L. Wang (Eds). *Identification of Continuous-time Models from Sampled Data*. Springer-Verlag, London, 2008.

[6] H. Garnier, M. Gilson, T. Bastogne, and M. Mensler. CONTSID toolbox: a software support for continuous-time data-based modelling. *In Identification of continuous-time models from sampled data*, H. Garnier and L. Wang (Eds.), Springer, London, pages 249–290, 2008.

[7] H. Garnier, L. Wang, and P.C. Young. Direct Identification of Continuous-time Models from Sampled Data: Issues, Basic Solutions and Relevance. *In Identification of continuous-time models from sampled data*, H. Garnier and L. Wang (Eds.), Springer, London, pages 1–29, 2008.

[8] M. Gilson, H. Garnier, P.C. Young, and P. Van den Hof. Instrumental variable methods for continuous-time closed-loop model identification. *In Identification of continuous-time models from sampled data*, H. Garnier and L. Wang (Eds.), Springer-Verlag, London, pp. 133-160, 2008.

[9] A.J. Jakeman, L.P. Steele, and P.C. Young. Instrumental variable algorithms for multiple input systems described by multiple transfer functions. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-10:593–602, 1980.

[10] 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.

[11] A.J. Jakeman and P.C. Young. Advanced methods of recursive time-series analysis. *International Journal of Control*, 37:1291–1310, 1983.

[12] V. Laurain, M. Gilson, H. Garnier, and P.C. Young. Refined instrumental variable methods for identification of Hammerstein continuous-time Box-Jenkins models. *47th IEEE Conference on Decision and Control (CDC'2008)*, Cancun (Mexico), December 2008.

[13] V. Laurain, M. Gilson, and H. Garnier. Refined instrumental variable methods for Hammerstein Box-Jenkins models. *In System Identification, Environmetric Modelling and Control System Design*, L. Wang, H. Garnier and T. Jakeman (Eds.), Springer, London, 2011.

[14] V. Laurain, M. Gilson, R. Toth, and H. Garnier. Direct identification of continuous-time LPV input/ouput models. *IET Control Theory and Applications*, special issue "Continuous-time model identification", 2011.

[15] L. Ljung. *System Identification. Theory for the User*. Prentice Hall, Upper Saddle River, 2nd edition, 1999.

[16] L. Ljung. Initialisation aspects for subspace and output-error identification methods. *European Control Conference*, Cambridge, UK, 2003.

[17] L. Ljung. Experiments with identification of continuous-time models. In *15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009.

[18] K. Mahata and H. Garnier. Identification of continuous-time errors-in-variables models. *Automatica*, 46(9):1477–1490, 2006.

[19] K. Mahata and H. Garnier. Identification of continuous-time Box-Jenkins models with arbitrary time-delay. *46th Conference on Decision and Control (CDC'2007)*, New Orleans, LA, USA, December 2007.

[20] G.P. Rao and H. Garnier. Numerical illustrations of the relevance of direct continuous-time model identification. *15th IFAC World Congress*, Barcelona, Spain, July 2002.

[21] G.P. Rao and H. Garnier. Identification of continuous-time systems: direct or indirect? *Systems Science*, 30(3):25–50, 2004.

[22] G. P. Rao and H. Unbehauen. Identification of continuous-time systems. *IEE Proceedings Control Theory & Appl.*, 153(2), March 2006.

[23] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer Verlag, New York, 1983.

[24] T. Söderström and P. Stoica, *System Identification*, Series in Systems and Control Engineering. Prentice Hall, 1989.

[25] V. Solo. *Time Series Recursions and Stochastic Approximation*. PhD thesis, Australian National University, Canberra, Australia, 1978.

[26] S. Thil, H. Garnier and M. Gilson Third-order cumulants based methods for continuous-time errors-in-variables model identification. *Automatica*, 44(3), 2008.

[27] P. Van den Hof. Closed-loop issues in system identification. *Annual Reviews in Control*, 22:173–186, 1998.

[28] L. Wang, H. Garnier and T. Jakeman (Eds). *System Identification, Environmetric Modelling and Control System Design*. Springer-Verlag, 2011.

[29] P.C. Young. In flight dynamic checkout - a discussion. *IEEE Transactions on Aerospace*, AS2(3):1106–1111, 1964.

[30] P.C. Young. The determination of the parameters of a dynamic process. *Radio and Electronic Engineering (Journal of IERE)*, 29:345–361, 1965.

[31] P.C. Young. An instrumental variable method for real-time identification of a noisy process. *Automatica*, 6:271–287, 1970.

[32] P.C. Young. Some observations on instrumental variable methods of time-series analysis. *International Journal of Control*, 23:593–612, 1976.

[33] P.C. Young. Parameter estimation for continuous-time models - a survey. *Automatica*, 17(1):23–39, 1981.

[34] P.C. Young. *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin, 1984.

[35] P.C. Young. Data-based mechanistic modeling of engineering systems. *Journal of Vibration and Control*, 4:5–28, 1998.

[36] P.C. Young. Data-based mechanistic modeling of environmental, ecological, economic and engineering systems. *Journal of Environmental Modelling and Software*, 13:105–122, 1998.

[37] P.C. Young and A.J. Jakeman. Refined instrumental variable methods of time-series analysis: Part I, SISO systems. *International Journal of Control*, 29:1–30, 1979.

[38] 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.

[39] P.C. Young, A.J. Jakeman, and R. McMurtrie. An instrumental variable method for model order identification. *Automatica*, 16:281–296, 1980.

[40] P.C. Young. The data-based mechanistic approach to the modelling, forecasting and control of environmental systems. *Annual Reviews in Control*, 30:169–182, 2006.

[41] P.C. Young and H. Garnier. Identification and estimation of continuous-time, data-based mechanistic models for environmental systems. *Environmental Modelling & Software*, 21:1055–1072, 2006.

[42] P.C. Young. The refined instrumental variable method: unified estimation of discrete and continuous-time transfer function models. *Journal Européen des Systèmes Automatisés*, 2008.

[43] P.C. Young, H. Garnier, and M. Gilson. Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models. In *Identification of continuous-time models from sampled data*, H. Garnier and L. Wang (Eds.), pages 91–132. Springer-Verlag, London, 2008.

[44] P.C. Young, H. Garnier, and M. Gilson. Simple refined IV methods of closed-loop system identification. *15th IFAC Symposium on System Identification (SYSID'2009), Saint-Malo (France), July 2009.*