

AN OPTIMAL INSTRUMENTAL VARIABLE APPROACH FOR IDENTIFYING HYBRID CONTINUOUS-TIME BOX-JENKINS MODELS

Peter Young*Hugues Garnier** Marion Gilson**

* *Centre for Research on Environmental Systems and
Statistics, Lancaster University, UK; and Integrated
Catchment Assessment and Management Centre, School of
Resources, Environment and Society, Australian National
University, Canberra, p.young@lancaster.ac.uk*

** *Centre de Recherche en Automatique de Nancy, CRAN
UMR 7039 CNRS-UHP-INPL, Université Henri Poincaré,
Nancy 1, BP 239, F-54506 Vandœuvre-les-Nancy Cedex,
France, hugues.garnier@cran.uhp-nancy.fr,
marion.gilson@cran.uhp-nancy.fr*

Abstract: The paper describes and evaluates an optimal instrumental variable method for identifying hybrid continuous-time transfer function models of the Box-Jenkins form from discrete-time sampled data, where the relationship between the measured input and output is a continuous-time transfer function, while the noise is represented as a discrete-time AR or ARMA process. The performance of the proposed hybrid parameter estimation scheme is evaluated by Monte Carlo simulation analysis. *Copyright ©2006 IFAC*

Keywords: Continuous-time models, hybrid models, instrumental variable, sampled data, parameter estimation

1. INTRODUCTION

This paper presents a method for identifying hybrid continuous-time ‘Box-Jenkins’ transfer function models from discrete-time data. This is a direct development of the Simplified Refined Instrumental Variable method for Continuous-time systems (SRIVC: Young and Jakeman, 1980) which has been used successfully for many years and has shown the advantages of the stochastic formulation of the Continuous-Time (CT) estimation problem. However, the ‘simplification’ that characterises the name of the SRIVC method is the assumption, for the purposes of algorithmic development, that the additive noise is purely white in form. Although 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, the estimates are not statistically efficient in this situation because the prefilters are not designed to account for the colour in the noise process. The proposed hybrid estimation procedure is the first time-domain implementation of an optimal CT estimation scheme for a hybrid BJ transfer function model¹. It follows logically from the Refined Instrumental Variable (RIV) estimation method for Discrete-Time (DT) models (Young, 1976; Young and Jakeman, 1979) by including concurrent DT noise model estimation and the

¹ An optimal frequency domain approach for the same model is presented in Pintelon *et al.* (2000).

use of this estimated noise model in the iterative-adaptive design and implementation of the pre-filters. For this reason, we refer to the new method as the **R**efined **I**nstrumental **V**ariable method for **C**ontinuous-time systems (RIVC). The hybrid form of the BJTF model is considered for two reasons. First, the theoretical and practical problems associated with the estimation of purely stochastic continuous-time AR or ARMA 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 in the paper, this can be achieved adequately on the basis of prefilters defined by reference to DT AR or ARMA noise models.

2. PROBLEM FORMULATION

Consider a linear, single-input, single-output, CT system² whose input $u(t)$ and output $y(t)$ are related by a constant coefficient differential equation

$$x^{(n)}(t) + a_1x^{(n-1)}(t) + \dots + a_nx^{(0)}(t) = b_0u^{(m)}(t) + b_1u^{(m-1)}(t) + \dots + b_mu^{(0)}(t) \quad (1)$$

where $x^{(i)}(t)$ denotes the i -th time-derivative of the continuous-time signal $x(t)$. Equation (1) can also be written in the transfer function (TF) form:

$$x(t) = \frac{B(s)}{A(s)}u(t), \quad (2)$$

with

$$B(s) = b_0s^m + b_1s^{m-1} + \dots + b_m, \\ A(s) = s^n + a_1s^{n-1} + \dots + a_n,$$

where s is the differential operator, *i.e.* $s^p x(t) = \frac{d^p x(t)}{dt^p}$. It is assumed that the input signal $\{u(t), t_1 < t < t_N\}$ is applied to the system and that the output $x(t)$ is sampled at discrete times t_1, \dots, t_N , not necessarily uniformly spaced. In order to obtain high quality statistical estimation results, it is vital to consider the inevitable errors that will affect the measured output. In the case of uniformly sampled data, the measured output $y_k = y(t_k)$, where $t_k = k\Delta t$, is assumed to be corrupted by an additive measurement noise $\xi(t_k)$,

$$y(t_k) = x(t_k) + \xi(t_k). \quad (3)$$

where $x(t_k)$ is the hypothetical noise-free, deterministic output of the system. The identification problem can now be stated as follows: estimate the parameters of the differential equation model (1) from N sampled measurements of the input and output $Z^N = \{u(t_k); y(t_k)\}_{k=1}^N$.

² A time-delay on the system input is not considered for simplicity here but is easy to accommodate.

3. OPTIMAL IV METHODS FOR CT MODELS

Most of the direct CT model identification techniques are deterministic in the sense that they do not consider the estimation of a noise model (see *e.g.* Garnier *et al.*, 2003). Disregarding the noise properties can lead to statistical inefficiency (increased variance of the estimates) and does not provide information on the estimated variance-covariance properties of the parameter estimates.

3.1 The SRIVC Method

One particularly successful stochastic identification method is the iterative **S**implified **R**efined **I**nstrumental **V**ariable method for **C**ontinuous-time systems (SRIVC: see Young and Jakeman, 1980; Young, 2002a,b) where, for simplicity, the noise is assumed to be discrete-time, white noise process so that no explicit noise modelling is necessary, except in relation to the estimation of the variance of the white noise process. This approach involves a method of iterative-adaptive prefiltering based on a quasi-optimal³ statistical solution to the problem in this white noise case. SRIVC is a logical extension of the more heuristically defined ‘state variable filter’ methods used previously and follows from the optimal **R**efined **I**nstrumental **V**ariable (RIV) and SRIV algorithms for DT identification (Young, 1984). This technique also has the advantage that the prefilters are automatically optimized in the iterative algorithm.

The SRIVC algorithm is based on the assumption that the additive noise $\xi(t_k)$ in (3) is simply a zero mean, normally distributed, white noise process $e(t_k)$: *i.e.* $e(t_k) = N(0, \sigma^2)$. In this situation, following the usual Prediction Error Minimization (PEM) approach, a suitable error function $\varepsilon(t)$ in continuous-time is given by the output error (OE),

$$\varepsilon(t) = y(t) - \frac{B(s)}{A(s)}u(t).$$

which can be written as,

$$\varepsilon(t) = \frac{1}{A(s)} [A(s)y(t) - B(s)u(t)]. \quad (4)$$

Minimization 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

³ The method is quasi-optimal because true optimality would require optimal interpolation of the input signal $u(t)$ over the sampling interval, whereas only simple interpolation is used in the present SRIVC implementation. However, this normally produces very good, near optimal estimation results.

linear case, the filter $F(s) = 1/A(s)$ can be taken inside the square brackets to yield:

$$\varepsilon(t) = A(s)y_f(t) - B(s)u_f(t) \quad (5)$$

or,

$$\begin{aligned} \varepsilon(t) = & y_f^{(n)}(t) + a_1 y_f^{(n-1)}(t) + \dots + a_n y_f^{(0)}(t) \\ & - b_0 u_f^{(m)}(t) - \dots - b_m u_f^{(0)}(t) \end{aligned} \quad (6)$$

where

$$\begin{cases} y_f^{(i)}(t) = f_i(t) * y(t), & i = 0, \dots, n \\ u_f^{(i)}(t) = f_i(t) * u(t), & i = 0, \dots, m, \end{cases} \quad (7)$$

Here, the $*$ symbol denotes the convolution operator and the set of filters takes the form,

$$f_i(t) = \mathcal{L}^{-1}\left(\frac{s^i}{A(s)}\right). \quad (8)$$

where \mathcal{L}^{-1} is the inverse Laplace transform. So that, at the k^{th} sampling instant $t = t_k$, the associated estimation model can be written in the form,

$$y_f^{(n)}(t_k) = \phi_f^T(t_k)\theta + \varepsilon(t_k) \quad (9)$$

where $\varepsilon(t_k)$ is the equation error, and

$$\begin{aligned} \phi_f^T(t_k) = & [-y_f^{(n-1)}(t_k) \dots - y_f^{(0)}(t_k) \\ & u_f^{(m)}(t_k) \dots u_f^{(0)}(t_k)] \end{aligned} \quad (10)$$

$$\theta = [a_1 \dots a_n \ b_0 \dots b_m]^T. \quad (11)$$

Thus, provided we assume that $A(s)$ in (8) is known *a priori*, the estimation model (9) forms a basis for the definition of a likelihood function and maximum likelihood estimation.

There are two problems with this formulation. The most obvious one is, of course, that the $A(s)$ polynomial required in (8) is not known *a priori*. The less obvious one is that, in practical applications, we cannot assume that the noise $\xi(t_k)$ will always be white: in most applications it will be coloured. Both of these problems can be solved by employing a similar approach to that used first in the RIV algorithm for DT system identification and estimation (see Young (1984) and the prior references therein). Here, a ‘relaxation’ optimization procedure is devised that iteratively adjusts the unknown polynomials in the TF model (2), as well as the estimate of the instrumental variable, until they converge. Here the instrumental variable at each iteration of the algorithm is generated by the following auxiliary model

$$\hat{x}(t) = \frac{B(s, \hat{\theta}_N^{IV})}{A(s, \hat{\theta}_N^{IV})} u(t). \quad (12)$$

where $\hat{\theta}_N^{IV}$ is the estimated parameter vector obtained at the previous iteration of the algorithm. In this iterative algorithm, the selection of the initial $\hat{A}_0(s)$ polynomial does not have to be particularly accurate provided the prefilter based on it does not seriously attenuate any signals

within the pass-band of the system being modelled (Young, 2002b). It is also assumed that studies based on model identification criteria have defined the orders n and m of the polynomials that characterize the TF model, as well as any pure time delays, the effects of which are removed prior to estimation.

3.2 The RIVC method for hybrid BJ models

If the noise $\xi(t_k)$ is coloured and described by the ARMA(p, q) model, as shown below in (13), then the SRIVC method has no claim to optimality in statistical terms. However, because of the inherent IV mechanism, the estimates are consistent and asymptotically unbiased. Moreover, experience has shown that the algorithm is very robust in practical application and normally yields estimates with reasonable statistical efficiency (*i.e.* low but not minimum variance). Nevertheless, it is worthwhile extending the algorithm to the coloured noise case by employing a hybrid approach where the noise modelling, as well as the noise-derived parts of the prefiltering, are carried out in discrete-time terms.

In general, following the normal approach to the estimation of stochastic TF models, it makes sense to assume that the discrete-time noise process $\xi(t_k)$ in the observation equation (3) has rational spectral density and so can be described by an ARMA(p, q) process, *i.e.*,

$$\xi(t_k) = \frac{D(z^{-1})}{C(z^{-1})} e(t_k) \quad e(t_k) = N(0, \sigma^2) \quad (13)$$

where,

$$\begin{aligned} C(z^{-1}) &= 1 + c_1 z^{-1} + \dots + c_p z^{-p} \\ D(z^{-1}) &= 1 + d_1 z^{-1} + \dots + d_q z^{-q} \end{aligned} \quad (14)$$

are, respectively, the Autoregressive (AR) and Moving Average (MA) polynomials in the backward shift operator, denoted here by z^{-r} ; *i.e.* $z^{-r} y_k = y_{k-r}$. Note that there are only a few direct continuous-time model identification methods that explicitly model the unknown noise transfer function in this manner (Young and Jakeman, 1980; Johansson, 1994; Pintelon *et al.*, 2000). Following the approach used in the SRIVC case, we now need to consider a hybrid prefiltering operation involving both continuous and discrete-time prefilters. In particular, since the discrete-time noise (13) is only associated with the definition of the observation equation (3) and we are dealing with a linear model, the prefiltering can be considered in these two stages: first, the continuous-time prefiltering already described in the SRIVC algorithm above; and second, discrete-time prefiltering based on the inverse of the ARMA(p, q) noise model.

In order to explain this hybrid prefiltering, consider the discrete-time sampled value $\varepsilon(t_k)$ of $\varepsilon(t)$ in (4). Introducing the ARMA(p, q) noise process (13), the RIVC error function $\varepsilon(t_k)$, now defined in discrete-time, can be written as,

$$\varepsilon(t_k) = \frac{C(z^{-1})}{D(z^{-1})}\varepsilon(t_k) \quad (15)$$

where the discrete-time prefilter will be recognised as the inverse of the ARMA(p, q) noise model. With this definition, the estimation model takes a similar form to (9), *i.e.* :

$$y_f^{(n)}(t_k) = \phi_f^T(t_k)\theta + \varepsilon(t_k) \quad (16)$$

but the prefilters required in the definition of the estimation model are obviously more complex, requiring the practical implementation of the above hybrid filtering operations.

Like the SRIVC algorithm, the full RIVC algorithm that exploits the above prefiltering concepts involves an iterative relaxation approach to optimization. The main steps in the algorithm are as follows:

The RIVC Algorithm.

Step 1. Apply the SRIVC algorithm in the standard manner based on an initial $\hat{A}_0(s)$, chosen as discussed previously, and generate initial estimates of the $\hat{A}_1(s)$ and $\hat{B}_1(s)$ polynomials.

Step 2. Iterative estimation.

for $j = 2 : \text{convergence}$

(1) (i) Obtain estimates of the polynomials $C_j(z^{-1})$ and $D_j(z^{-1})$ in an ARMA(p, q) noise model based on the estimated noise sequence $\hat{\xi}(t_k) = y(t_k) - \hat{x}(t_k)$, using a standard discrete-time AR or ARMA model estimation algorithm.

(2) (ii) Prefilter the input and output signals by the inverse noise filter $\hat{D}_j(z^{-1})/\hat{C}_j(z^{-1})$ and obtain updated estimates of the polynomials $\hat{A}_j(s)$ and $\hat{B}_j(s)$ using the SRIVC algorithm based on these prefiltered data.

end

4. SIMULATION EXAMPLES

This section considers a simulation example that illustrates the performance of the SRIVC algorithm in the noise free case and the proposed full RIVC method in case of coloured output noise. The simulation model used in this analysis provides a very good test for CT and DT estimation methods: it was first suggested by Rao and Garnier (2002) (see also Garnier *et al.* 2003); and has been commented upon by Ljung (2003). Its transfer function is given by (Rao and Garnier, 2002):

$$G_o(s) = \frac{-6400s + 1600}{s^4 + 5s^3 + 408s^2 + 416s + 1600}. \quad (17)$$

4.1 Noise-free case

The noise-free case is considered first in order to investigate the effect of systematic interpolation errors that are introduced into the computation of the prefiltered time-derivatives. This is due to the fact that, in general, the inter-sample behavior of the prefilter inputs is not known, so that time-domain approaches to CT model identification often require a high sampling frequency compared with the bandwidth of the system to be identified in order to minimize these systematic interpolation errors. As we shall see, this is no longer the case with the SRIVC method, which is not sensitive to the sampling interval.

The observation time is fixed at $T_f = 20$ s and the number of data points available N is inversely proportional to the sampling interval Δt . The input $u(t)$ is chosen as a binary ± 1 signal. Note that in the case of the chosen piece-wise constant excitation signal, the system response can be calculated exactly at the sampling instances via appropriate Zero-Order Hold (ZOH) discretization of the CT system. The estimation methods considered are the simple LSSVF method (see *e.g.* Garnier *et al.*, 2003); the iterative SRIVC method; and, finally, an indirect method based on the estimation of a DT ARX model converted to a CT model by using the appropriate `d2cm` conversion tool in Matlab with the ZOH option. Note that the SRIVC routine is initialized here with the estimates obtained from the LSSVF method. The cut-off frequency of the SVF filter is set to 10 rad/sec.

As the SRIVC method is an iterative procedure, it is stopped either when a specified number of iterations has elapsed or when the relative percentage error in the parameter estimate vector is sufficiently small. In order to compare the accuracy of the different approaches, the sum of the normalised estimation error (NEE) of the parameters is considered

$$\text{NEE}(\hat{\theta}) = \left\| \frac{\theta^o - \hat{\theta}}{\theta^o} \right\|_2 = \sqrt{\sum_{j=1}^{n_p} \left(\frac{\theta_j^o - \hat{\theta}_j}{\theta_j^o} \right)^2}, \quad (18)$$

where n_p is the number of parameters to be estimated, $\hat{\theta}$ denotes the estimated parameter vector, $\hat{\theta}_j$ represents the j -th element of $\hat{\theta}$, while the superscript 'o' is used to denote the true value. The normalised estimation error versus the sampling period for the three methods is plotted in Figure 1(a). From this figure, the LSSVF method yields relatively poor results. This is caused by the small numerical errors introduced into the simulation of the CT filtering operations required to generate the filtered time-derivatives. This conventional approach requires a relatively high sampling frequency compared to the bandwidth of the

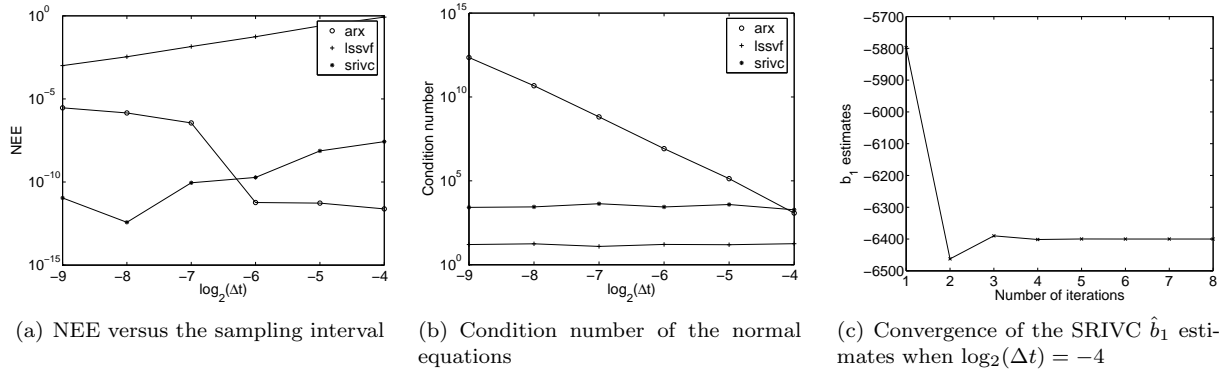


Fig. 1. Estimation results in the noise-free case

system in order to be identified well with accurate parameter estimates. This result is in agreement with the statements often found in the literature pointing out that direct continuous-time model identification requires a high sampling frequency to minimize the systematic errors. The use of a more sophisticated interpolation technique than the `lsim` routine used here would make it possible to obtain better results. However, this is not necessary in the case of the SRIVC method because its iterative IV based procedure is able to counteract the presence of the simulation-induced errors, which are effectively treated as additive noise on the filtered time-derivatives.

Although the SRIVC and the indirect estimation approaches perform similarly overall in this noise-free situation, it is clear that, at small sampling intervals, the SRIVC estimation results are better; while at the coarser sampling intervals, the position is reversed. This is an expected result, since it is well known that DT estimation schemes suffer from numerical issues at small sampling intervals because the eigenvalues of the DT model migrate ever closer to the unit circle in the complex domain as the sampling frequency increases.

Figure 1(b) shows how the condition number of the normal equations for the three methods varies with the sampling interval. This reveals a significant numerical advantage of the direct approach for all chosen sampling periods. The indirect approach also suffers from additional numerical problems that are encountered in the conversion from the DT to the CT domain as the sampling interval decreases.

The SRIVC algorithm needs between 2 to 9 iterations to converge from the LSSVF parameter estimates, as illustrated for \hat{b}_1 in Figure 1(c). In the case of coarse sampling periods, an alternative to LSSVF initiation is the discrete-time RIV algorithm from the CAPTAIN Toolbox⁴.

⁴ This is also now available in the latest version 4.1 of the CONTSID Toolbox.

4.2 Noisy output case

In these experiments, A PRBS (Pseudo-Random Binary Signal) of maximum length is used to excite the system over its whole bandwidth. The characteristics of the signal, whose amplitude switches between -1 and $+1$, are the following: the number of stages of the shift register is set to $n_s = 9$, the clock period is set to $n_p = 3$, which results in the number of samples $N = 1533$. The sampling interval is chosen as $\Delta t = 50$ ms. Based on the CT process (17), the following hybrid system is considered

$$\begin{aligned} x(t) &= G_0(s)u(t) \\ y(t_k) &= x(t_k) + \xi(t_k) \end{aligned} \quad (19)$$

Here, $\xi(t_k)$ is coloured noise derived from an AR(2) process $1/C(z^{-1})$, where,

$$C(z^{-1}) = 1 + 1.9628z^{-1} + 0.9851z^{-2} \quad (20)$$

The variance of the additive noise $\xi(t_k)$ is adjusted to obtain a signal-to-noise ratio of 10dB.

The MCS results obtained from the Continuous-time Output Error (COE), SRIVC and full RIVC algorithms are shown in Table 1. These are based on $N_{exp} = 200$ random realizations, with the white noise input to the AR noise model being selected randomly for each realization. In order to compare the statistical performance of the different approaches, the computed mean $\bar{\theta}_j$ and standard deviation $\sigma_{\hat{\theta}_j}$ of the estimated parameters are presented, as well as the empirical normalised mean square error (NMSE) which is defined as

$$\text{NMSE}(\hat{\theta}_j) = \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} \left(\frac{\theta_j^o - \hat{\theta}_j(i)}{\theta_j^o} \right)^2, \quad (21)$$

where $\hat{\theta}_j(i)$ is the j -th element of the estimated parameter vector, while the superscript ‘o’ denotes the true value of the parameter. The number of iterations required for the iterative methods to converge is also considered in the analysis of the estimation results.

From Table 1, we see that both COE and SRIVC methods deliver the same good results, although

method	True value	b_0	b_1	a_1	a_2	a_3	a_4	c_1	c_2	$N_{iter} \pm \sigma_{N_{iter}}$
COE (MCS)	$\hat{\theta}_j$	-6410.8	1578.6	5.0062	408.1	415.9	1605.0			10.17 \pm 0.83
	$\sigma_{\hat{\theta}_j}$	241.1	337.5	0.1875	1.29	20.10	38.1			
	NMSE(θ_j)	1.4e-3	4.4e-2	1.4e-3	1.0e-5	2.3e-3	5.7e-4			
SRIVC (MCS)	$\hat{\theta}_j$	-6410.8	1578.7	5.0063	408.1	415.9	1605.0			6.55 \pm 0.55
	$\sigma_{\hat{\theta}_j}$	241.1	337.5	0.1876	1.29	20.10	38.1			
	NMSE(θ_j)	1.4e-3	4.4e-2	1.4e-3	1.0e-5	2.3e-3	5.7e-4			
RIVC (MCS)	$\hat{\theta}_j$	-6399.2	1594.4	4.9974	408.0	415.3	1603.3	-1.9612	0.9835	7.64 \pm 0.56
	$\sigma_{\hat{\theta}_j}$	8.6	146.0	0.0211	0.16	7.55	23.1	0.0045	0.0044	
	NMSE(θ_j)	1.8e-6	8.3e-3	1.8e-5	1.5e-7	3.3e-4	2.1e-4			
RIVC (SR)	$\hat{\theta}_j$	-6409.0	1472.9	5.038	408.2	429.6	1637.1	-1.9615	0.9841	
	$\sigma_{\hat{\theta}_j}$	8.3	138.0	0.0226	0.16	8.29	21.5			

Table 1. Monte Carlo simulation results with additive coloured noise. SR denotes single run results and MCS denotes from MCS analysis

the COE requires more iterations, on average to converge. However, as expected, the hybrid RIVC method clearly improves the statistical efficiency and also provides consistent estimates of the AR model parameters.

5. CONCLUSIONS

This paper has described how the full RIVC method for identifying hybrid Box-Jenkins transfer function models for single input, single output systems has been developed from the earlier SRIVC algorithm. The extension to multiple-input, single-output systems is straightforward and the RIVC algorithm will shortly become available in the CAPTAIN and CONTSID Toolboxes⁵ (these already include the SRIVC algorithm).

Based on simulation results obtained in the noise-free case, it is clear that the performance of the SRIVC method is not sensitive to the sampling frequency, in contrast to the conventional LSSVF method. Also, as expected, it performs somewhat better than alternative indirect approaches, based on DT estimation followed by conversion to CT, particularly at higher sampling frequencies. In the case of noisy data, the results obtained in the Monte Carlo Simulation studies, using a difficult fourth order model, demonstrate the improvement in statistical efficiency that accrues from the explicit inclusion of noise model effects in the RIVC algorithm. In particular, the empirical parametric error variance is reduced and agrees with that obtained from the algorithm's estimate of the parametric error covariance matrix, as required.

⁵ SRIVC and RIVC algorithms are available in the CONTSID (<http://www.cran.uhp-nancy.fr/contsid/>) and CAPTAIN (<http://www.es.lanacs.ac.uk/cres/captain/>) Toolboxes for Matlab. CONTSID contains the major identification tools for continuous-time systems, including SRIVC/RIVC; CAPTAIN has tools for discrete (SRIV/RIV) and continuous (SRIVC/RIVC) time systems; time variable and state dependent parameter (non-linear) models; and nonstationary signal extraction and adaptive forecasting.

REFERENCES

- 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.
- Jakeman, A.J. and P.C. Young (1979). Refined instrumental variable methods of time-series analysis: Part II, multivariable systems. *International Journal of Control* **29**, 621–644.
- Johansson, R. (1994). Identification of continuous-time models. *IEEE Transactions on Signal Processing* **42**(4), 887–896.
- Ljung, L. (2003). Initialisation aspects for subspace and output-error identification methods. In: *European Control Conference (ECC'2003)*. Cambridge (U.K.).
- Pintelon, R., J. Schoukens and Y. Rolain (2000). Box-Jenkins continuous-time modeling. *Automatica* **36**, 983–991.
- Rao, G.P. and H. Garnier (2002). Numerical illustrations of the relevance of direct continuous-time model identification. In: *15th Triennial IFAC World Congress on Automatic Control*. Barcelona (Spain).
- Young, P.C. (1976). Some observations on instrumental variable methods of time-series analysis. *International Journal of Control* **23**, 593–612.
- Young, P.C. (1984). *Recursive Estimation and Time-Series Analysis*. Springer-Verlag. Berlin.
- Young, P.C. (2002a). Comments on "On the estimation of continuous-time transfer functions". *International Journal of Control* **75**(9), 693–697.
- Young, P.C. (2002b). Optimal IV identification and estimation of continuous-time TF models. In: *15th Triennial IFAC World Congress on Automatic Control*. Barcelona (Spain).
- 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.