15th Triennial World Congress, Barcelona, Spain

# OPTIMAL IV IDENTIFICATION AND ESTIMATION OF CONTINUOUS-TIME TF MODELS

**Peter C. Young**

*Systems and Control Group, Centre for Research on Environmental Systems and Statistics, Lancaster University, Lancaster LA1 4YQ, U.K.,*
*and*
*CRES, Australian National University, Canberra, Australia*

Abstract: The paper outlines an optimal, maximum likelihood formulation of the continuous-time transfer function estimation problem and shows how the Refined Instrumental Variable (RIVC) algorithm provides an iterative solution to this problem. With the help of a simulation example, it then compares and contrasts the performance of the RIVC algorithm with that of two other, sub-optimal, continuous-time transfer function estimation procedures that have been suggested more recently. *Copyright © 2002 IFAC*

Keywords: continuous-time, transfer function, optimal instrumental variables, maximum likelihood, optimal estimation, comparative study.

## 1. INTRODUCTION

Since the early 1960's, numerous different approaches have been suggested for the identification and estimation [2] of continuous-time, linear Transfer Function (TF) models from normal operating data (see e.g. the reviews by Young, 1981, and Unbehauen and Rao, 1997). The present paper revisits one of the algorithms first suggested many years ago in this context, the iterative or recursive-iterative [3] *Refined Instrumental Variable* (RIVC) algorithm (Young and Jakeman, 1980). The paper shows how, by virtue of its special exploitation of adaptive prefilters on the input-output signals, the RIVC method can be interpreted in optimal, maximum likelihood terms. This revisitation appears justified since, despite its many advantages, the RIVC method still appears to be relatively unknown. Moreover, the algorithm is now widely accessible in the CAPTAIN Matlab Toolbox (see http://www.es.lancs.ac.uk/cres/captain/) and has recently been incorporated in the CONTSID Matlab Toolbox (see http://www.cran.uhp-nancy.fr/ cran /i2s/contsid/contsid.html).

The RIVC algorithm has been used for many years in practical applications (e.g. most recently, Price *et al*., 1999). In the present paper, however, its advantages in comparison with other algorithms are demonstrated on the simulation example used in the recent paper by Wang and Gawthrop (2001: WG from here on). The RIVC estimation results obtained in this manner are compared with those obtained by WG, as well as those obtained using another IV algorithm: the IVGPMF algorithm of Garnier *et al*. (1995). This third algorithm uses the so-called *Poisson Moment Functional* (PMF) implementation of the *State Variable Filter* (SVF) concept and it also relates very closely to much earlier work by the present author (Young, 1970a and the prior references therein), who referred to the PMF filter chain as the 'Method of Multiple Filters' (MMF).

In contrast to the other two algorithms, the RIVC approach does not require the user to specify any as-

[1] The author is most grateful to Professor Hughes Garnier for providing the IVGPMF algorithm (as part of their CONTSID Matlab®Toolbox) and for his advice on the use of the algorithm.

[2] Here I use the statistical meanings of these words: 'identification' is the definition of the most appropriate model order; and 'estimation' is the estimation of the parameters that characterize this identified model.

[3] An on-line, recursive form of the algorithm has also been developed (see e.g. Young, 1984).

pect of the prefilters (SVFs) other than their dynamic order. Rather these prefilters, which are so important in continuous-time TF estimation, are adjusted in an iterative fashion, so that they can perform two simultaneous functions: first to optimally filter the data and so make the estimation more statistically efficient (i.e. lower and, in the Gaussian normal case, minimum variance parameter estimates); and secondly, to generate the filtered derivatives of the input and output signals. In this second role, they perform a similar function to the prefilters used by WG and IVGPMF for this same purpose but their prefilters are normally restricted to having real eigenvalues [4] . In addition, the iterative, adaptive mode of solution used by the RIVC algorithm not only ensures that, on convergence, the estimates have statistically optimum properties, it also generates information on the parametric error covariance matrix. This information is useful for subsequent *Monte Carlo Simulation* (MCS) analysis, as well as providing the standard error bounds on the parameter estimates.

## 2. THEORETICAL OVERVIEW

### 2.1 *The RIVC Algorithm*

The RIVC approach to continuous-time linear model identification and estimation suggested by Young and Jakeman (1980) derives from the equivalent RIV approach for discrete-time systems (see Young, 1984 and the prior references therein). It is also a logical development of the earlier, more heuristic methods developed by the author (Young, 1970a and the prior references therein). The theoretical basis for the method can be outlined by considering the following SISO system, although MISO and MIMO extensions are straightforward:

$$\begin{aligned} x(t) &= \frac{B(s)}{A(s)}u(t-\delta) \\ y(t) &= x(t) + e(t) \end{aligned} \qquad (1)$$

Here $A(s)$ and $B(s)$ are polynomials in the derivative operator $s = d/dt$ of the form:

$$\begin{aligned} A(s) &= s^n + a_1 s^{n-1} + ... + a_{n-1}s + a_n \\ B(s) &= b_0 s^m + b_1 s^{m-1} + ... + b_{m-1}s + b_m \end{aligned}$$

and $\delta$ is any pure time delay in time units. This model structure is denoted by the triad $[n, m, \delta]$. In (1), $u(t)$ is the input signal, $x(t)$ is the 'noise free' output signal and $y(t)$ is the noisy output signal. Initially, the noise $e(t)$ is considered as zero mean, white noise with Gaussian amplitude distribution and variance $\sigma^2$, although we will see later that this assumption is not restrictive.

[4] The MMF/PMF implementation of prefilters is constrained to prefilters with real eigenvalues. The WG prefilters are not so restricted but, for simplicity, WG impose such a constraint in their numerical examples.

Following the usual *Maximum Likelihood* (ML) approach, a suitable error function that defines the likelihood is given by,

$$\begin{aligned} \varepsilon(t) &= y(t) - \frac{B(s)}{A(s)}u(t-\delta) \\ &= \frac{1}{A(s)}\left\{A(s)y(t) - B(s)u(t-\delta)\right\} \end{aligned}$$

which is the basis for the *response* or *output error* estimation methods. However, since the operators commute in this case, the $1/A(s)$ filter can be taken inside the brackets to yield the expression,

$$\varepsilon(t) = A(s)y^*(t) - B(s)u^*(t-\delta) \qquad (2)$$

or,

$$\begin{aligned} \varepsilon(t) = s^n y^*(t) + a_1 s^{n-1} y^*(t) + ... + a_n y^*(t) \\ -b_0 s^m u^*(t-\delta) - ... - b_m u^*(t-\delta) \end{aligned}$$

where the $^*$ superscript indicates that the associated variable has been '*prefiltered*' by $1/A(s)$. The advantage of this transformation is that (2) is now linear in the unknown parameters $a_i, i = 1, ..., n; b_j, j = 0, ..., m$, so that the associated estimation model can be written in the form:

$$s^n y^*(t) = \mathbf{z}^*(t)^T \mathbf{a} + e(t) \qquad (3)$$

where,

$$\mathbf{z}^*(t) = \left[-s^{n-1}y^*(t) ... - y^*(t)\; s^m u^*(t-\delta) ... u^*(t-\delta)\right]^T$$

$$\mathbf{a} = [a_1 \; ... \; a_n \; b_0 \; ... \; b_m]^T$$

As a result, all of the prefiltered derivatives appearing as variables in this estimation model are measurable as the inputs of the integrators that appear in the realization of the prefilter $1/A(s)$. Thus, provided we assume that $A(s)$ is known, the estimation model (3) forms a basis for the definition of a likelihood function and ML estimation.

There are two problems with this formulation. The obvious one is, of course, that $A(s)$ is not known *a priori*. The less obvious one is that, in practical applications, we cannot assume that the noise $e(t)$ will have the nice white noise properties assumed above: it is likely that the noise will be a coloured noise process, say $\xi(t)$. Both of these problems can be solved by employing a similar approach to that used in the RIV algorithm for discrete-time (backward shift operator TF) system identification and estimation (see Young, 1984 and the prior references therein). Here, a 'relaxation' optimization procedure is devised that adaptively adjusts an initial estimate $\hat{A}_0(s)$ of $A(s)$ iteratively until it converges on an optimal estimate of $A(s)$. And the coloured noise problem is solved conveniently by exploiting IV estimation within this iterative optimization algorithm.

The continuous-time version RIVC of this RIV algorithm is described fully in Young and Jakeman (1980). Like most IV methods, it exploits an IV variable $\hat{x}(t)$ generated from the following 'auxiliary model' (Young, 1970a):

$$\hat{x}(t) = \frac{\hat{B}(s)}{\hat{A}(s)} u(t-\delta) \qquad (4)$$

and an associated IV vector defined as,

$$\hat{x}^*(t) = \left[-s^{n-1}\hat{x}^*(t) ... - \hat{x}^*(t) \; s^m u^*(t-\delta) ... u^*(t-\delta)\right]^T$$

The iterative RIVC algorithm [5] can then be summarized as follows:

**The RIVC Algorithm**

(1) Select the initial $\hat{A}_0(s)$ polynomial either automatically or manually (see explanation below)

(2) Use $\hat{A}_0(s)$ to generate the pre-filtered variables $y^*(t)$ and $u^*(t)$ and obtain estimates $\hat{A}_1(s)$ and $\hat{B}_1(s)$ using linear least squares.

(3) Iterate: $k = 2 : ni$ (default $ni = 4$)
(i) Using $\hat{A}_{k-1}(s)$ and $\hat{B}_{k-1}(s)$ to replace $A(s)$ and $B(s)$, respectively, generate both the pre-filtered data vector $z^*(t)$ and the prefiltered instrumental variable vector $\hat{x}^*(t)$, the latter using the IV variable $\hat{x}(t)$ generated by the auxiliary model (4).
(ii) Calculate the IV estimate

$$\hat{\mathbf{a}} = \mathbf{C}^{-1}\mathbf{b}$$

$$\mathbf{C} = \sum_{i=1}^{N} \hat{\mathbf{x}}^*(t_i)\mathbf{z}^*(t_i)^T$$

$$\mathbf{b} = \sum_{i=1}^{N} \hat{\mathbf{x}}^*(t_i)s^n y^*(t_i)$$

where $t_i$ denotes the $i^{th}$ of $N$ sampling instants.

(4) Generate an estimate of the parametric covariance matrix $\mathbf{P}$ using the symmetric version of the IV algorithm (Young, 1970b, 1984), i.e.

$$\mathbf{P} = \hat{\sigma}^2 \left[\sum_{i=1}^{N} \hat{\mathbf{x}}^*(t_i)\hat{\mathbf{x}}^*(t_i)^T\right]^{-1} \qquad (5)$$

where $\hat{\sigma}^2$ is the variance of the model residuals. The square root of the diagonal elements of $\mathbf{P}$ provides the standard errors on $\hat{\mathbf{a}}$ (see theorem below).

[5] recursive-iterative and on-line recursive versions of this algorithm are easily implemented: see Young and Jakeman (1980).

*Comments*

1. The initiation of the above RIVC algorithm involves the selection of a suitable $\hat{A}_0(s)$ prefilter polynomial. This is not a difficult task and there are several alternatives available to the user in the CAPTAIN Toolbox. An automatic option is based on the user specifying a suitable sub-sampling interval for initial discrete-time estimation (this can often be the actual sampling interval, since the discrete-time RIV algorithm is quite robust and works well even with rapidly sampled data). The discrete-time model obtained in this manner is then automatically transformed to continuous-time form using the Matlab *d2cm* tool and provides the $\hat{A}_0(s)$ polynomial. Manual options include the user specification of a suitable single pole value to generate an all-pole MMF/PMF prefilter or a full order prefilter. The latter two options can be specified without iterative adaption, in which case the former is equivalent to IVGPMF.

Of course, the final estimates are quite robust to the specification of $\hat{A}_0(s)$ since the prefilter is thereafter adaptively adjusted by the algorithm: $\hat{A}_0(s)$ is simply required as a device to allow for the initial least squares estimation step. In the example below, for instance, automatic $\hat{A}_0(s)$ selection via a discrete-time model based on the actual sampling interval is clearly the simplest user-option, but *identical* results are obtained if the all-pole filter option (MMF/PMF) is selected for *any* specified pole value in the range -0.007 to -6. In effect, therefore, this latter option is simple to use and virtually automatic. Based on experience, it is the author's preferred option.

2. An associated RIVCID order identification algorithm allows the user to automatically search over a whole range of different model orders (see example below); a very useful option in practice.

3. The RIVC algorithm is computationally efficient: in the example below, for instance, it is 5.3 times faster than the IVGPMF algorithm (cpu time 0.17 sec. compared with 0.9 sec.). And the RIVCID algorithm takes only 9.9 sec. to search over all models in the range [1,1,0] to [5,5,3].

### 2.2 *Theoretical Justification for the IV Variable in RIVC*

The following theorem is a generalization of a similar theorem for discrete-time TF models proven rigorously by Pierce (1972) and later by Young (1984, p.213-215) using simpler and somewhat less rigorous, prediction error analysis.

*Theorem*

(i) If the $e(t)$ in (1) is a zero mean, Gaussian white noise process;

(ii) the parameter values are admissible (i.e. the model is stable and identifiable); and

(iii) $u(t)$ is persistently exciting, then the ML estimate $\hat{\mathbf{a}}_N$, obtained from the data set of $N$ samples, possesses a limiting normal distribution, such that the asymptotic covariance matrix of the estimation errors associated with the estimate $\hat{\mathbf{a}}_N$ is of the form:

$$\mathbf{P} = \hat{\sigma}^2 \left[ \sum_{i=1}^{N} \hat{\mathbf{x}}^*(t_i)\hat{\mathbf{x}}^*(t_i)^T \right]^{-1} \qquad (6)$$

*Proof* Modification of the proof in Pierce (1972) and Young (1984) to the continuous-time case.

Although there is no formal proof, simple arguments based on the nature of the IV iterations, comprehensive MCS studies, and all experience over the last 20 years, has shown that the algorithm is strongly convergent. Provided the above conditions are satisfied, therefore, the converged RIVC estimates are optimal in a ML sense.

## 3. THE WG EXAMPLE

The WG example concerns the identification and estimation of the following continuous-time TF model:

$$y(t) = \frac{-2s+1}{s^3+1.6s^2+1.6s+1}u(t) + e(t) \qquad (7)$$

using a total of 3000 samples obtained from an experiment in which the system (7) is enclosed within a feedback loop in series with a switch and a relay with hysteresis (see WG for details). The noisy data set generated in this manner is shown in figure 1, which has the same general form and noise level as that used by WG.

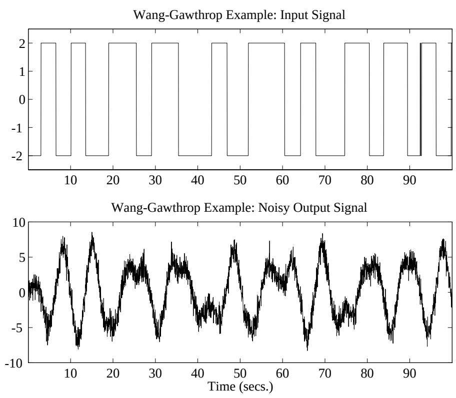


Fig. 1. Input (upper panel) and output (lower panel) data obtained from the WG 'relay experiment'.

In the identification stage of the analysis, two statistical measures are used to choose between a range of model orders. These are the $R_T^2$ and $YIC$ criteria (e.g. Young, 1989), which are defined as follows, where $np = n + m + 1$:

$$R_T^2 = 1 - \frac{\hat{\sigma}^2}{\sigma_y^2}$$

$$YIC = \log_e \left\{ \frac{\hat{\sigma}^2}{\sigma_y^2} \right\} + \log_e \frac{1}{np} \sum_{i=1}^{i=np} \frac{\hat{\sigma}^2 p_{ii}}{\hat{\theta}_i^2}$$

Here, $\sigma_y^2$ is the variance of the output signal; $\hat{\theta}_i^2$ is the squared value of the $i^{th}$ estimated parameter; and $p_{ii}$ is the $i^{th}$ diagonal element of the RIVC estimated parametric error covariance matrix $\mathbf{P}$. The $R_T^2$ criterion will be recognized as a *Coefficient of Determination* based on the simulated model errors (note: *not* the one-step-ahead prediction errors). The YIC criterion is more complex and provides a measure of how well the parameters are defined statistically: the more negative the YIC, the better the definition.

### 3.1 *RIVC and WG Analysis*

In this case, RIVCID applied over the range [1,1,0] to [5,5,3] identifies two models as strong contenders: the $[2, 2, 0]$ model has the most negative $YIC$ (-9.08), with an associated $R_T^2 = 0.91$. However, the correct $[3, 2, 0]$ model has the second best $YIC = -8.54$, which is still very negative implying very well defined estimates, and its $R_T^2 = 0.94$ is significantly better. So the two criteria, taken together, suggest that the correct $[3, 2, 0]$ model is superior in these terms. All other possibilities are rejected: the $[3, 3, 0]$ model estimated by WG has almost the same $R_T^2$ but its $YIC = -2.25$ suggests over-parameterization; while $[4, 4, 0]$ and $[5, 5, 0]$ models have similar $R_T^2$ values to the $[3, 2, 0]$ model but their $YIC$s are $0.60$ and $-1.44$, respectively, so they are decisively rejected.

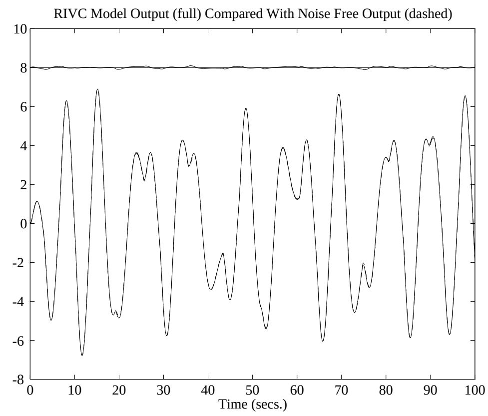


Fig. 2. Comparison of $[3, 2, 0]$ model output (full line) and the deterministic output of the true simulated system (7), shown dashed. The error (+8) is shown above.

The estimated parameters in the $[3, 2, 0]$ model are as follows:

$$\hat{a}_1 = 1.5917(0.0442); \hat{a}_2 = 1.6037(0.0229);$$
$$\hat{a}_3 = 1.0066(0.0319); \hat{b}_0 = -2.0013(0.0599);$$
$$\hat{b}_1 = 0.9923(0.0176);$$

where the figures in parentheses are the estimated standard errors. These results are very good and a significant improvement on those obtained by WG (see Young, 2001). The parameter estimates are very close to the actual parameter values from (7) and, not surprisingly, the model output matches the deterministic output of the true simulated system very well, as shown in figure 2.

Finally, Table 1 shows the results of an MCS exercise similar to that carried out by WG. Here, the WG simulation was repeated 50 times, each time with different and independent additive random noise. The data from each such realization were then used to estimate the model parameters and Table 1 compares the standard errors on the estimated parameters obtained from the MCS analysis (second row) with those obtained from the RIVC estimated $\mathbf{P}$ matrix (first row). The results demonstrate clearly that the RIVC algorithm is performing optimally and verify the theory behind the optimal RIVC methodology.

Table 1: *RIVC Standard Error Comparison*

| Param. | $\hat{a}_1$ | $\hat{a}_2$ | $\hat{a}_3$ | $\hat{b}_0$ | $\hat{b}_1$ |
|---|---|---|---|---|---|
| RIVC | 0.044 | 0.023 | 0.032 | 0.060 | 0.018 |
| MCS | 0.044 | 0.023 | 0.030 | 0.056 | 0.021 |

### 3.2 *RIVC and IVGPMF Analysis*

This second comparative analysis is based on the 1000 sample data set shown in figure 3, which was generated as a random realization, again using the WG simulation. This shorter data set provides a greater estimation challenge than the longer data set in figure 1, particularly since the input excitation is rather poor.

RIVCID identification again suggests the correct [3,2,0] model order[6] and automatic initial prefilter selection was used in the RIVC algorithm. In contrast, the prefilter parameter $\alpha$ (see Garnier *et al*., 1995) in the IVGPMF has to be specified manually by the user. In order to ensure that the algorithm performed as well as possible, therefore, this parameter was selected as the value that produced parameter estimates as close to the *true* values as possible (i.e. a prefilter parameter $\alpha = 2$ and an associated prefilter $1/C(s)$ with $C(s) = (s+2)^4$). This means that the IVGPMF results obtained here are the best achievable in practice.

The comparative estimation results obtained in the above manner are given in Table 2[7]. As expected from theory, the RIVC results are superior, even though the IVGPMF prefilter parameter is set at its best possible value. However, the author has been informed (H. Garnier *pers. comm.*) that an experienced user will be able to choose a prefilter parameter that achieves similar performance. This being so, the IVGPMF results are perfectly acceptable in most

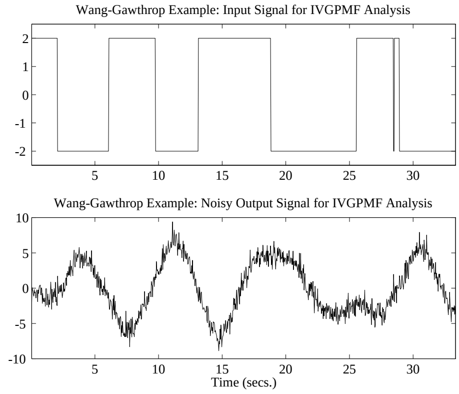


Fig. 3. Input (upper panel) and output (lower panel) data used in the IVGPMF analysis.

practical terms. Even in this more difficult, low sample size situation, with rather poor input excitation, only one of the IVGPMF parameter estimates ($\hat{a}_1 = 1.552$) appears not so good, but it still lies within the bounds computed by the RIVC algorithm. And in relation to the true system, the associated model step and frequency responses are not too much worse than those of the RIVC estimated model.

Table 2: *RIVC and IVGPMF Estimation Results*

| TRUE | 1.6 | 1.6 | 1.0 | -2.0 | 1.0 |
|---|---|---|---|---|---|
| RIVC | 1.605 | 1.605 | 0.993 | -1.990 | 1.033 |
| SE(**P**) | 0.083 | 0.042 | 0.059 | 0.109 | 0.032 |
| IVGPMF | 1.552 | 1.597 | 0.957 | -1.948 | 1.030 |

Of course, single estimation simulations such as this are not a reliable way to evaluate the overall estimation performance of an algorithm. This is better accomplished by MCS analysis, the results of which are shown in Table 3 for an ensemble of 100 random realizations. We see that both the mean parameter values and the mean standard errors of the RIVC estimates are better than the IVGPMF estimates, although the difference is not great because the best possible MMF/PMF prefilter parameter was utilized for all realizations in the MCS analysis (again being extra specially fair to the algorithm). Also the RIVC standard errors conform with those predicted by the RIVC estimated covariance matrix $\mathbf{P}$, once more confirming the efficacy of the background theory.

Table 3: *RIVC and IVGPMF MCS Results*

| SE (**P**) | 0.076 | 0.039 | 0.056 | 0.100 | 0.031 |
|---|---|---|---|---|---|
| RIVC | 1.605 | 1.605 | 1.004 | -2.009 | 1.003 |
| MCS SE | 0.068 | 0.031 | 0.048 | 0.083 | 0.028 |
| IVGPMF | 1.611 | 1.607 | 1.008 | -2.014 | 1.005 |
| MCS SE | 0.102 | 0.050 | 0.075 | 0.138 | 0.036 |

## 4. A COMMENT ON PREFILTERS

Obviously, one of the main differences between the RIVC and both the WG and IVGPMF estimation algorithms is the way in which the prefilters are chosen. In particular, the optimal RIVC prefilter is generated

[6] The IVGPMF algorithm uses a somewhat different order convention and the model is designated [3,1,0] in these terms.

[7] The IVGPMF algorithm does not provide standard error estimates.

within the algorithm via an iteratively updated, adaptive optimization procedure. In the WG example, this yields a prefilter $1/A(s)$, with $A(s) = s^3 + 1.5917s^2 + 1.6037s + 1.0066$ upon convergence of this iterative procedure (compared with the optimal $s^3 + 1.6s^2 + 1.6s + 1.0$). In the WG and IVGPMF approaches, on the other hand, the prefilter is fixed and defined as $1/C(s)$ using either (i) a much more computationally intensive optimization procedure (see WG paper), or (ii) manual selection (IVGPMF). In this example, WG find that $C(s) = (s + 2)^3$ following optimization. The best manual selection for the MMF/PMF filter chain is $C(s) = (s + 2)^4$.

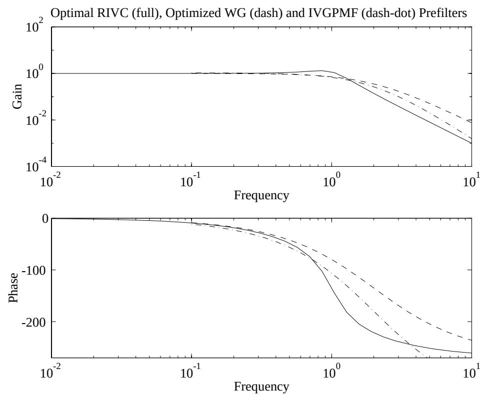


Fig. 4. Comparison of Bode plots for the optimal RIVC (full), optimized WG (dashed) and IVGPMF (dash-dot) prefilters

The Bode plots for these prefilters are compared in figure 4. We see that, although the characteristics are broadly similar, there are differences, with the IVGPMF filter more closely approximating $1/A(s)$ over the most important part of the passband. In particular, the pass-band of the RIVC filter gain characteristic (upper panel, full line) is somewhat narrower and there is a small resonant peak (with related differences in the phase characteristics, as shown in the lower panel). This is the main reason for the differences in the estimation results. The RIVC filter is more precisely defined in relation to the passband of the system and so its noise attenuation properties are better than the WG and IVGPMF prefilters, with the result that the statistical efficiency is correspondingly higher.

## 5. CONCLUSIONS

This paper outlines the optimal RIVC algorithm for the identification and estimation of continuous-time TF models and compares its performance with that of two other, alternative algorithms, WG and IVGPMF. It is shown to be both more statistically and computationally efficient than these other algorithms. The IVGPMF alternative performs quite well in the comparative study provided one assumes prior knowledge is available to help choose the prefilter parameter (not required by RIVC). It is interesting to note that, despite these obvious advantages of the RIVC algorithm, it is not widely used. Perhaps this is because RIVC exploits iteratively adaptive prefilters and so it is incorrectly perceived as too complicated and not robust enough for practical usage. The present paper demonstrates that this is clearly not the case. The RIVC algorithm involves a simple, rapidly convergent, iterative optimization procedure and appears to be very robust in practical terms. Indeed, despite its optimality, it is rather easier to use than the IVGPMF algorithm, since it does not require manual specification of the prefilter parameters and, as we have seen, it is computationally quite a lot faster.

## 6. REFERENCES


Garnier H., P. Sibille and A. Richard (1995) Continuous-time canonical state-space model identification via Poisson moment functionals, *Proc. 34th IEEE Conference on Decision and Control (CDC95), New Orleans - USA*, 3004-3009.

Pierce, D. A. (1972) Least squares estimation in dynamic disturbance time-series models, *Biometrika*, **59**, 73-78.

Price, L., P. C. Young, D. Berckmans, K. Janssens and J. Taylor (1999) Data-based mechanistic modelling and control of mass and energy transfer in agricultural buildings, *Annual Reviews in Control*, **23**, 71-82.

Unbehauen, H. and G. P. Rao (1997) Identification of continuous-time systems: a tutorial, *Proc. 11th IFAC Symposium on System Identification, Kitakyushu, Japan*, 1023-1049.

Wang, L and P. Gawthrop (2001) On the estimation of continuous-time TFs, *Int. Jnl. of Control*, **74**, 889-904.

Young, P. C. (1970a) An instrumental variable method for real-time identification of a noisy process, *Automatica*, **6**, 271-287.

Young, P. C. (1970b) The Differential Equation Error Method of Real-Time Process identification, Ph.D Thesis, University of Cambridge, U.K.

Young, P. C. (1981) Parameter estimation for continuous-time models - a survey, *Automatica*, **17**, 23-39.

Young, P. C. (1984) *Recursive Estimation and Time Series Analysis*, Springer-Verlag: Berlin.

Young, P. C. (1989) Recursive estimation, forecasting and adaptive control, chapter in C.T. Leondes (ed.) *Control and Dynamic Systems: Advances in Theory and Applications*, **30**, Academic Press: San Diego, 119-166.

Young, P. C. (2002) Comments on "On the estimation of continuous-time transfer functions" by L. Wang and P. Gawthrop, *Int. Jnl. of Control*. In press.

Young, P. C. and A. J. Jakeman (1980) Refined instrumental variable methods of recursive time-series analysis: part III, extensions, *Int. Jnl. of Control*, **31**, 741-764.