LUND UNIVERSITY

Continuous-Time Identification of SISO Systems using Laguerre Functions

Chou, C. T.; Verhaegen, Michel; Johansson, Rolf

Published in:
IEEE Transactions on Signal Processing

1999

Link to publication

Citation for published version (APA):
Chou, C. T., Verhaegen, M., & Johansson, R. (1999). Continuous-Time Identification of SISO Systems using Laguerre Functions. *IEEE Transactions on Signal Processing*, *47*(2), 349-362.

Total number of authors:
3

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Continuous-Time Identification of SISO Systems Using Laguerre Functions

Chun Tung Chou, Michel Verhaegen, and Rolf Johansson

***Abstract*— This paper looks at the problem of estimating the coefficients of a continuous-time transfer function given samples of its input and output data. We first prove that any nth-order continuous-time transfer function can be written as a fraction of the form $\sum_{k=0}^{n} \bar{b}_k L_k(s) / \sum_{k=0}^{n} \bar{a}_k L_k(s)$, where $L_k(s)$ denotes the continuous-time Laguerre basis functions. Based on this model, we derive an asymptotically consistent parameter estimation scheme that consists of the following two steps: 1) Filter both the input and output data by $L_k(s)$, and 2) estimate $\{\bar{a}_k, \bar{b}_k\}$ and relate them to the coefficients of the transfer function. For practical implementation, we require the discrete-time approximation of $L_k(s)$ since only sampled data is available. We propose a scheme that is based on higher order Padé approximations, and we prove that this scheme produces discrete-time filters that are approximately orthogonal and, consequently, a well-conditioned numerical problem. Some other features of this new algorithm include the possibility to implement it as either an off-line or a quasi-on-line algorithm and the incorporation of constraints on the transfer function coefficients. A simple example will be given to illustrate the properties of the proposed algorithm.**

***Index Terms*— Asymptotic consistency, continuous-time systems, Laguerre basis functions, Padé approximation, parameter estimation, system identification, total least-squares.**

I. Introduction

THIS PAPER looks at the problem of estimating the coefficients of single-input single-output (SISO) dynamic models specified by a linear differential equation with constant coefficients.

The approach taken here can be classified as the state variable filter approach according to an earlier survey paper on continuous-time identification by Young [22] or as the "linear dynamics operations" approach according to a later survey paper on the same topic by Unbenhauen and Rao [16]. This particular approach consists of two steps. In the first step, the input/output data is passed through a bank of filters. This filtering action is chosen in such a way that the filtered input and output are related by an equation whose unknown coefficients are related to those of the differential equation. The second step then consists of estimating these unknown coefficients and relating them to those of the differential equation. The prime motivation of this approach is that the undesirable action of differentiating noisy data is replaced by the action of this bank of filters. Various choices of filters have been suggested, for example, the use of integrators has been suggested by Sagara and Zhou [13] and Schoukens [14], and the use of low-pass filters by Johansson [5] and Moonen *et al.* [10]. In this paper, we have chosen these filters to be continuous-time Laguerre functions, and we shall justify this choice later.

This paper looks at two different scenarios of the continuous-time identification problem. We shall first study the noise-free case, and then, we shall look at the case where the output is corrupted by white measurement noise.

The noise-free case for continuous-time identification is not as straightforward as its discrete-time counterpart. Since only the sampled versions of the continuous-time signals are available, the output of the state variable filters can only be computed from discrete approximations of these filters. We find that discrete-time filters obtained from first-order approximation methods [for example, first-order-hold (FOH) or Tustin transform] have two main drawbacks. First, these approximations are only accurate when the sampling frequency is very high compared with the bandwidth of the system to be identified. Second, the estimated continuous-time model may be unstable even if the given continuous system is a stable one.

In order to overcome these problems, we propose a scheme to compute discrete-time approximation of continuous-time Laguerre filters by applying second- or third-order Padé approximations in a particular way. Besides the fact that these discrete-time filters give excellent approximation over a large bandwidth, there is an added advantage that is due to the use of Laguerre filters with higher order Padé approximation. The resulting discrete-time filters are approximately orthogonal, and the numerical conditioning of the associated parameter estimation problem is therefore improved. However, the above added advantage comes with a price. These higher order discrete-time filters are unstable. Although unstable filters can be implemented as causal/anticausal filter, there is the problem of unknown end conditions.

Our main contribution is that we have derived an accurate and asymptotically consistent parameter estimation scheme—based on using higher order approximations of continuous-time Laguerre filters—which can deal with both the unknown initial conditions of the system and the unknown end conditions of these anticausal filters. Furthermore, we have derived an explicit transformation that maps the coefficients of the equation relating the filtered signals to those of

Manuscript received November 13, 1996; revised June 13, 1998. The associate editor coordinating the review of this paper and approving it for publication was Dr. Yingbo Hua.

C. T. Chou and M. Verhaegen are with the Systems and Control Laboratory, Department of Electrical Engineering, Delft University of Technology, Delft, The Netherlands (e-mail: M.verhaegen@et.tudelft.nl).

R. Johansson is with the Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Publisher Item Identifier S 1053-587X(99)00762-X.

1053–587X/99$10.00 © 1999 IEEE

the differential equation. This transformation also allows us to incorporate *a priori* knowledge such as the relative degree of the transfer function. Finally, this algorithm can be implemented in either an off-line or quasi-on-line fashion.

This paper is organized in the following way. We first define the parameter estimation problem in Section II, and then, we present the outline of the solution for the noise-free case in Section III. In Section IV, we suggest a scheme to obtain discrete-time approximations of continuous-time filters and compare its accuracy with higher order hold circuits. We also prove, in this section, that the discrete-time filters obtained by applying our proposed approximation scheme to continuous-time Laguerre filters are approximately orthogonal. The parameter estimation algorithm for the noise-free case is then presented in Section V and for the noisy case in Section IV. Section VII gives a simple example that illustrates the properties of the proposed algorithm. Finally, the conclusions are given in Section VIII.

II. Problem Statement

We consider a continuous-time SISO system whose input $u(t)$ and output $y(t)$ are related by a linear constant coefficient differential equation of order n

$$\begin{aligned} &y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_0 y(t) \\ &\quad = b_n u^{(n)}(t) + b_{n-1}u^{(n-1)}(t) + \cdots + b_0 u(t) \end{aligned} \tag{1}$$

where $y^{(n)}(t)$ denotes the nth derivative of the continuous-time signal $y(t)$. The system is assumed to be subjected to an arbitrary set of initial conditions

$$u_0 = [u(0) \quad u^{(1)}(0) \quad \cdots \quad u^{(n-1)}(0)] \tag{2}$$

$$y_0 = [y(0) \quad y^{(1)}(0) \quad \cdots \quad y^{(n-1)}(0)]. \tag{3}$$

Furthermore, we make two assumptions on the polynomials $A(s)$ and $B(s)$, which are defined as

$$A(s) = s^n + a_{n-1}s^{n-1} + \cdots + a_0 \tag{4}$$

$$B(s) = b_n s^n + b_{n-1}s^{n-1} + \cdots + b_0. \tag{5}$$

First, these two polynomials are assumed to be relatively prime. Second, the roots of the polynomial $A(s)$ are assumed to have negative real parts; in other words, the system defined by (1) is assumed to be asymptotically stable.

For the identification problem, it is assumed that the continuous-time signals $u(t)$ and $y(t)$ are sampled at regular time interval T. The measured output $z(t)$ is assumed to be corrupted by an additive white measurement noise $e(t)$ with variance $(1/T)\lambda^2$; in other words,

$$z(t) = y(t) + e(t). \tag{6}$$

The data available for identification is $\{u(kT), z(kT)\}$ (for $k = 0, \cdots, N$), and they will be denoted as $\{u(k), z(k)\}$ when the context makes it clear that they denote samples of the continuous-time signals.

The identification problem can now be stated as follows: Given the sampled data $\{u(k), z(k)\}$, determine the coefficients $a_{n-1}, \cdots, a_0$ and $b_n, \cdots, b_0$ of the differential equation model.

III. Outline of Solution

In this section, we outline the solution to the identification problem in the noise-free case. This consists of two steps. In the first step, the measured input and output are filtered through a bank of Laguerre functions. It can be shown that the filtered input, filtered output, and the impulse responses of the Laguerre filters are related by a linear equation with constant coefficients and that these coefficients can be computed from the null space of a specific matrix. Once these coefficients are determined, the coefficients of the differential equation can be computed via a linear transformation, and this is the second step.

We first define a few notation that will be used in this paper. Let a be any positive real number and w be the allpass filter

$$w = \frac{s-a}{s+a}. \tag{7}$$

In addition, let $\{L_k(s)\}$ for $k = 0, 1, 2, \cdots$ be the set of Laplace domain Laguerre functions where

$$L_k(s) = \frac{\sqrt{2a}}{s+a}\left(\frac{s-a}{s+a}\right)^k \tag{8}$$

and $\{\ell_k(t)\}$ be their corresponding functions in the time domain.

Consider the Laplace transform of the differential equation defined by (1)

$$A(s)Y(s) = B(s)U(s) + C(s) \tag{9}$$

where $U(s)$ and $Y(s)$ are, respectively, the Laplace transforms of the input $u(t)$ and the output $y(t)$. The polynomial $C(s)$ is defined as

$$C(s) = c_{n-1}s^{n-1} + \cdots + c_0 \tag{10}$$

and it contains the part of system response due to the unknown initial conditions. Note that the coefficients c_i depend on the unknown parameters a_i and b_i as well as the unknown initial conditions.

By substituting

$$s = a\,\frac{1+w}{1-w} \tag{11}$$

into (9) and multiplying both sides of the resulting equation by $(1-w)^n$, we arrive at a model of the form

$$\overline{A}(w)Y(s) = \overline{B}(w)U(s) + \overline{C}(w) \tag{12}$$

where

$$\overline{A}(w) = (1-w)^n A\left(a\,\frac{1+w}{1-w}\right) \tag{13}$$

and the polynomials $\overline{B}(w)$ and $\overline{C}(w)$ are similarly defined. Furthermore, we shall denote the coefficients of these three polynomials by

$$\overline{A}(w) = \overline{a}_n w^n + \cdots + \overline{a}_0 \tag{14}$$

$$\overline{B}(w) = \overline{b}_n w^n + \cdots + \overline{b}_0 \tag{15}$$

$$\overline{C}(w) = \overline{c}_n w^n + \cdots + \overline{c}_0. \tag{16}$$

Note that an identical transformation technique is also used by Johansson [5], except that w is chosen to be $(a/(s+a))$. In general, this type of transformation is referred to as a linear dynamic operation in the survey paper on continuous-time identification by Unbenhauen and Rao [16].

By multiplying both sides of (12) by the lowpass filter $\sqrt{2a}/(s+a)$, we arrive at the Laguerre model structure that will be used in this paper.

$$\begin{aligned}&(\overline{a}_n L_n(s) + \cdots + \overline{a}_0 L_1(s))Y(s)\\ &\quad = (\overline{b}_n L_n(s) + \cdots + \overline{b}_0 L_0(s))U(s)\\ &\qquad + (\overline{c}_n L_n(s) + \cdots + \overline{c}_0 L_0(s)).\end{aligned} \tag{17}$$

Note that in the above derivation, we have effectively shown that an nth-order transfer function can be written as a ratio of two linear functions of Laguerre basis functions, and this representation is coined as the generalized ARX model by Wahlberg [20]. Note also that the above derivation is in sharp contrast to the present literatures on using Laguerre functions and other orthonormal basis functions for system identification [8], [17], [20], [21], where a transfer function is approximated by a truncated Laguerre expansion or, more generally, a finite sum of orthonormal basis functions.

In terms of time domain signals, (17) can be written as

$$\begin{aligned}&\overline{a}_n[L_n y](t) + \overline{a}_{n-1}[L_{n-1}y](t) + \cdots + \overline{a}_0[L_0 y](t)\\ &\quad = \overline{b}_n[L_n u](t) + \overline{b}_{n-1}[L_{n-1}u](t) + \cdots + \overline{b}_0[L_0 u](t)\\ &\qquad + \overline{c}_n \ell_n(t) + \overline{c}_{n-1}\ell_{n-1}(t) + \cdots + \overline{c}_0 \ell_0(t)\end{aligned} \tag{18}$$

where

$$[L_i y](t) = \ell_i(t) * y(t) \tag{19}$$

$$[L_i u](t) = \ell_i(t) * u(t) \tag{20}$$

and $*$ denotes the convolution operator.

Define the matrices

$$\overline{W} = -\begin{bmatrix} \ell_n(0) & \ell_{n-1}(0) & \cdots & \ell_0(0)\\ \vdots & \vdots & \vdots & \vdots\\ \ell_n(N) & \ell_{n-1}(N) & \cdots & \ell_0(N)\end{bmatrix} \tag{21}$$

$$\overline{U} = -\begin{bmatrix} [L_n u](0) & [L_{n-1}u](0) & \cdots & [L_0 u](0)\\ \vdots & \vdots & \vdots & \vdots\\ [L_n u](N) & [L_{n-1}u](N) & \cdots & [L_0 u](N)\end{bmatrix} \tag{22}$$

$$\overline{Y} = \begin{bmatrix} [L_n y](0) & [L_{n-1}y](0) & \cdots & [L_0 u](0)\\ \vdots & \vdots & \vdots & \vdots\\ [L_n y](N) & [L_{n-1}y](N) & \cdots & [L_0 y](N)\end{bmatrix}. \tag{23}$$

Under the assumption that the input signal $u(t)$ is persistently exciting, then we can recover the coefficients $\overline{a}_i$, $\overline{b}_i$, $\overline{c}_i$ from the null space of the following matrix.

$$M = [\overline{W} \quad \overline{U} \quad \overline{Y}]. \tag{24}$$

We shall show in Section V how we can compute the coefficients of the differential equation (1) given $\overline{a}_i$ and $\overline{b}_i$.

Remark 3.1: Note that it is also possible to write an order n discrete-time transfer function as a ratio of two linear functions of discrete-time Laguerre functions. The steps are completely analogous to those used above with $(s-a)/(s+a)$ and $\sqrt{2a}/(s+a)$ replaced, respectively, by $(1-\tilde{a}z)/(z-\tilde{a})$ and $\sqrt{(1-\tilde{a}^2)T}/(z-\tilde{a})$, where $-1 < \tilde{a} < 1$.

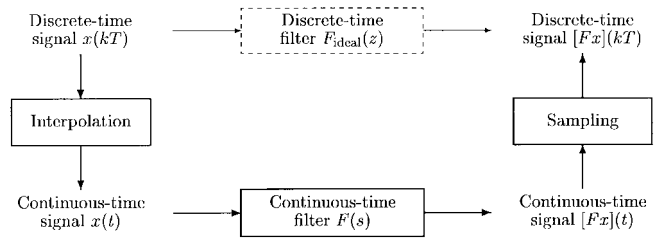


Fig. 1. Two approaches to Problem 4.1.

IV. Discrete Approximation of Filters

A Note on Notation: Both continuous and discrete-time signals and filters will appear in this section. We use $x(t)$ and $x(kT)$ to denote the continuous-time signal and its sample. The notation $[Fx]$ is used to denote convolution in both time domains; the nature of filter F should make it clear. Finally, $[Fx](kT)$ may mean either "samples of $F(s)$ convoluted with $x(t)$" or "$F(z)$ convoluted with $x(kT)$"; this again can be decided from what F is.

A. Approximation Scheme

Recall from the last section that the first identification step requires the value of $[L_i u](t)$ and $[L_i y](t)$ at $t = 0, T, \cdots$. Since only samples of $u(t)$ and $y(t)$ are available, this gives the motivation to study the problem:

Problem 4.1: Let $x(t)$ be a unknown continuous-time signal whose bandwidth is below the Nyquist frequency, and $F(s)$ is a given continuous-time filter. The problem is, given sampled data $x(kT)$, compute $[Fx](kT)$.

There are two ideal solutions to this problem. The first one is to use Shannon reconstruction to compute $x(t)$ from $x(kT)$. The signal $[Fx](kT)$ can then be obtained via filtering and sampling. This approach is depicted in Fig. 1 by the thick solid line path. Practically, the Shannon reconstruction is approximated by some higher order hold (HOH) circuits.

The second ideal solution is a more direct one: Is there a discrete-time filter whose output is $[Fx](kT)$ when its input is $x(kT)$? This solution path is depicted by the "dashbox" in Fig. 1.

Theorem 4.1: Based on problem formulation 4.1, we maintain that the discrete-time filter

$$F_{\text{ideal}}(z) = F\left(\frac{1}{T}\log(z)\right) \tag{25}$$

has the property $[F_{\text{ideal}}x](kT) = [Fx](kT)$.

Proof: Consider the discrete Fourier transform (DFT) of $[F_{\text{ideal}}x](kT)$, which is equal to $F_{\text{ideal}}(e^{j\omega T})X(e^{j\omega T})$, where $X(e^{j\omega T})$ is the DFT of $x(kT)$. By (25), we have $F_{\text{ideal}}(e^{j\omega T}) = F(j\omega)$, and the bandlimited assumption of $x(t)$ implies that $X(e^{j\omega T}) = X(j\omega)$. This means that the signals $[F_{\text{ideal}}x](kT)$ and $[Fx](kT)$ have the same spectrum, and this completes our argument. △

Equation (25) indicates that this ideal discrete-time filter cannot be realized as a finite-dimensional rational transfer function. In the next paragraph, we suggest a scheme to approximate this filter based on Padé approximation, and we

shall examine the properties of these approximate filters when F is a Laguerre function in Section IV-B.

Let $\Psi_p(\varphi(z))$ denote a pth-order Padé approximation of $\varphi(z)$. Define the function

$$\rho_p(z) = \frac{1}{T}\,\Psi_p(\log(1+x))|_{x=z-1} \tag{26}$$

and we propose

$$F_{\text{ideal}}(z) \approx F(\rho_p(z)). \tag{27}$$

Note that this approximation is not of Padé type, but we shall show in Theorem 4.2 that it has, in a certain sense, the same accuracy as Padé type of approximations. Bearing this in mind and in order to avoid expressions such as "discrete-time approximation of $F(s)$ based on pth-order Padé approximation of ...," we shall simply refer to it as "pth Padé approximation of $F(s)$."

Some of the discussion later on will be limited to Padé approximation of orders one to three. For completeness, expressions for $\rho_p(z)$ ($p = 1, \cdots, 3$) are given below.

$$\rho_1(z) = \frac{2}{T}\frac{z-1}{z+1} \tag{28}$$

$$\rho_2(z) = \frac{3}{T}\frac{z^2-1}{z^2+4z+1} \tag{29}$$

$$\rho_3(z) = \frac{1}{3T}\frac{11z^3+27z^2-27z-11}{z^3+9z^2+9z+1}. \tag{30}$$

Note that $\rho_1(z)$ coincides with Tustin transformation.

B. Properties of the Approximate Filters

This section examines the properties of the approximate filters obtained from Padé approximation. Note that the section on accuracy is applicable to any rational filter $F(s)$, whereas the rest applies only to Laguerre filters.

1) Accuracy: This aim of this section is to examine the accuracy of the Padé approximation scheme. In order to derive an expression for the error, we first argue that this scheme can also be viewed as a filter-dependent interpolation scheme, c.f., the thick-lined path in Fig. 1. It can be shown, using the same type of argument in Theorem 4.1, that if we first interpolate $x[kT]$ by $F(s)^{-1}F(\rho_p(e^{sT}))$ and then follow by filtering by $F(s)$ and sampling, then the resulting discrete-time signal is given by $[F(\rho_p(z))x](kT)$. This interpretation implies that the approximation error of Padé method is identical to that of $1 - F(s)^{-1}F(\rho_p(e^{sT}))$.

Theorem 4.2: Let ρ_p be as defined in (26); then

$$F(s) - F(\rho_p(e^{sT})) = O(T^{2p+1}). \tag{31}$$

Consequently, let $\epsilon_{\text{Padé},\,p}$ denote the error in estimating $[Fx](kT)$ using pth-order Padé approximation; then

$$\epsilon_{\text{Padé},\,p} = O(T^{2p+1}). \tag{32}$$

Proof: See Appendix B.

This theorem shows that the approximation error is of the same order as one that is based on the true Padé approximation. As a comparison, under the assumption that the continuous-time signal $x(t)$ is sufficiently smooth, a pth-order hold circuit gives an approximation error of $O(T^{p+1})$, which is larger than that of Padé approximation of the same order.

A more empirical approach would be to compare the frequency domain error measure

$$\epsilon(e^{j\omega T}) = |\hat{F}_d(e^{j\omega T}) - F_{\text{ideal}}(e^{j\omega T})| \tag{33}$$

where $\hat{F}_d(z)$ an approximation to F_{ideal}. Here, we assume that $F(s)$ is $L_1(s)$ with $a = 1$ and $T = 0.01s$. The error in approximating $L_1(s)$ using first-, second-, and third-order holds and the scheme based on Padé approximations orders 1 to 3 are plotted in Fig. 2. It can be seen that drastic improvement can be obtained by using higher order Padé approximations, as second-order approximation introduces nearly zero error up to $1/1000T$, whereas third-order approximation gives almost zero error up to $1/100T$. This shows that we can get a good interpolation property without too much oversampling.

2) Tapped Delay Line Structure: An advantage to using Laguerre filters is that they can be implemented efficiently in a manner similar to a tapped delay line. Recall the definitions of w in (7); we have

$$L_k(\rho_p(z)) = L_0(\rho_p(z))w(\rho_p(z))^k. \tag{34}$$

This shows that these approximate filters can again be implemented like a tapped delay line; note that this does not apply to HOH's. Furthermore, $w(\rho_p(z))$ for $p = 1, 2, 3$ are allpass filters, as shown in (35)–(37) at the bottom of the page. This allpass property will be made use of in Section IV-B3.

3) Approximate Orthogonality: It is a well-known fact that the Laguerre functions $L_k(s)$ are orthonormal functions. We shall show that if we apply our proposed approximation scheme to $L_k(s)$, the resulting discrete-time filters are approximately orthogonal in a sense that will be defined later. Based on this result and a result proved by Wahlberg [20], which states that the use of discrete-time Laguerre functions in linear regression gives rise to a numerically well-conditioned problem, we can therefore justify our choice from a numerical

$$w(\rho_1(z)) = \frac{\alpha z - 1}{z - \alpha} \quad \text{where} \quad \alpha = \frac{2-aT}{2+aT} \tag{35}$$

$$w(\rho_2(z)) = -\frac{(aT-3)z^2 + 4aTz + (aT+3)}{(aT+3)z^2 + 4aTz + (aT-3)} \tag{36}$$

$$w(\rho_3(z)) = -\frac{(3aT-11)z^3 + 27(aT-1)z^2 + 27(aT+1)z + (3aT+11)}{(3aT+11)z^3 + 27(aT+1)z^2 + 27(aT-1)z + (3aT-11)} \tag{37}$$

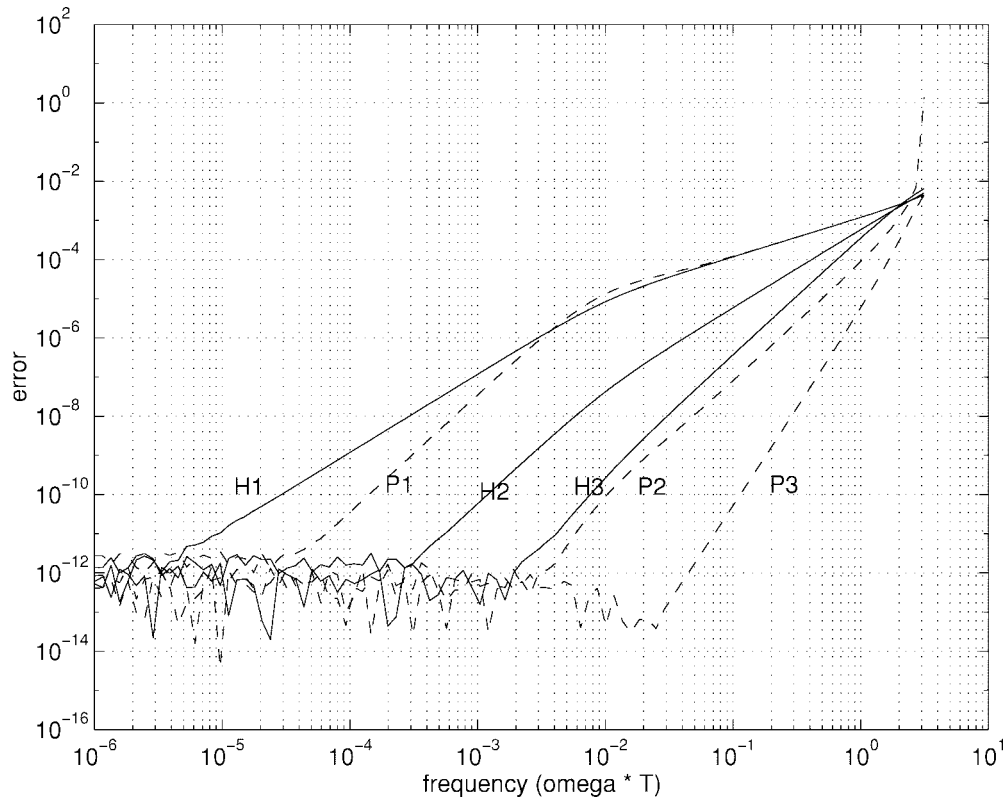


Fig. 2. Error [see (33)] in discrete filters obtained by using different orders of hold circuits and Padé approximation. Keys: solid lines for hold circuits (labeled as "H") and dashed lines for Padé approximations (labeled as "P"); the number after the letter indicates the order of approximation.

point of view. The use of Laguerre filters was, in fact, implicitly suggested by Johansson [5] as well for improvement on numerical properties. The suggested scheme orthogonalizes the family of filters $\{(a/(s+a))^k\}$, and it can be shown that it effectively expresses these lowpass filters as a sum of Laguerre functions.

Let $e(t)$ be a discrete-time zero-mean white noise sequence with variance $1/T$ (T is the sampling interval), and let $L_{k,d}(z)$ denote the discrete-time approximation of the continuous-time function $L_k(s)$. We shall investigate the property of the following covariance matrix

$$\Pi = \lim_{N\to\infty} \frac{1}{N} \overline{E}^T \overline{E} \tag{38}$$

where

$$\overline{E} = \begin{bmatrix} [L_{n,d}e](0) & [L_{n-1,d}e](0) & \cdots & [L_{0,d}e](0) \\ \vdots & \vdots & \vdots & \vdots \\ [L_{n,d}e](N) & [L_{n-1,d}e](N) & \cdots & [L_{0,d}e](N) \end{bmatrix}. \tag{39}$$

Due to the preservation in allpass structure in $L_{k,d}(z)$, it can easily be shown, for example, by considering the cross spectrum, that for a nonzero T

$$\begin{aligned} &\mathrm{corr}([L_{i,d}e](t), [L_{j,d}e](t)) \\ &\quad = \mathrm{corr}([L_{|i-j|+1,d}e](t), [L_{1,d}e](t)) \end{aligned} \tag{40}$$

where corr denotes the cross correlation between two stochastic sequences. This implies that the covariance matrix Π has the structure of a symmetric Toeplitz matrix. (Note that identical results are proved by Wahlberg [20] when these discrete-time filters are, in fact, discrete-time Laguerre functions.)

The following theorem asserts that filters $\{L_{k,d}(z)\}$ are approximately orthogonal.

Theorem 4.3: If $L_{k,d}(z)$ is derived from $L_k(s)$ using Padé approximation of order 1, 2, or 3, then

$$\lim_{T\to 0} \Pi = \kappa I \tag{41}$$

where I denotes the identity matrix $\kappa = 1$ for first- and third-order approximations, and $\kappa = 4/3$ for second-order approximation.

Proof: See Appendix B.

4) Stability: By using Jury's stability test [6] or root locus, it can be shown that the filters obtained by applying second- and third-order Padé approximations to $L_k(s)$ are unstable for all positive T. Straightforward implementation of these unstable filters will certainly cause problems in parameter estimation, but this can be overcome by implementing them as causal/anticausal filters. Appendix A shows how any unstable filter can be converted into a causal/anticausal filter, and it also contains an example to illustrate the concept.

5) Discussions: An alternative way to obtain an approximation to $F_{\mathrm{ideal}}(z)$ is to solve the model reduction problem

$$\min_{Q(z)\in\Sigma(p)} \oint_{|z|=1} |W(z)(F_{\mathrm{ideal}}(z) - Q(z))|_2^2 \frac{dz}{z} \tag{42}$$

where $\Sigma(p)$ is the set of all rational functions of order p, and $W(z)$ is a frequency weighting that can used to take

into account the system bandwidth. This method may produce filters with better approximation property than those obtained from Padé approximation. A drawback to this method is that a model reduction problem has to be solved for each Laguerre filter used. Furthermore, the approximate filters may not have a tapped delayed line structure and, therefore, cannot be efficiently implemented.

To conclude, we have shown in this section that the approximate filters obtained from applying the Padé scheme to Laguerre filters have certain desirable properties. We therefore recommend this approximation scheme instead of HOH. Further remarks on HOH can be found in Remarks 5.1 and 5.2 and the comparison study in Section VII.

V. Parameter Estimation: Noise-Free Case

Recall from Section III that the coefficients $\overline{a}_i$ and $\overline{b}_i$ can be recovered from the null space of the matrix M in (24). Since only samples of $u(t)$ and $y(t)$ are available, M will be estimated by using approximate filters, as discussed in Section IV. If the approximation scheme is FOH or Tustin transformation, then under the conditions that the input signal is persistently exciting and that approximation error is insignificant, then the coefficients $\overline{a}_i$ and $\overline{b}_i$ can be computed from the null space of

$$M_1 = [\overline{W} \quad \overline{U}_d \quad \overline{Y}_d] \tag{43}$$

where the matrices $\overline{U}_d$ and $\overline{Y}_d$ are defined in the same way as $\overline{U}$ and $\overline{Y}$ [refer to (22) and (23)], except that their elements are computed from discrete approximation of $L_k(s)$.

However, for the case where either a second- or third-order Padé approximation scheme is used, we need to compensate for the effect of unknown end conditions of these mixed causal/anticausal filters. We shall show how we can deal with these unknown end conditions in Section V-A. After that, in Section V-B, we will derive a transformation that enables us to calculate the differential equation coefficients a_i and b_i from $\overline{a}_i$ and $\overline{b}_i$. Finally, in Section V-C, we will present a computation scheme to calculate $\overline{a}_i$ and $\overline{b}_i$.

A. Corrections for Unknown End Conditions

Note that the initial conditions of the discrete filters can be chosen to be zero without causing any problems. However, this is not the case for the end conditions if causal/anticausal filters are used since these end conditions are generally nonzero. There are two possible solutions here: We may either choose to estimate these end conditions as in [19] or to compensate for their effect in some way. The latter approach will be taken here.

Note that when we use causal/anticausal filters, the correct output is given by the sum of response due to the input with zero end condition and the response due to zero input with nonzero end conditions. (This fact is illustrated in the example in Appendix A.) As an example, consider the case where the signal $u(t)$ is filtered by a discrete approximation of $L_k(s)$, $L_{k,d}(z)$. The correct output is given by

$$[L_k u](t) \approx [L_{k,d} u](t) = [L_{k,d,0} u](t) + \zeta_k(t)^T \eta_{u,k} \tag{44}$$

where

$[L_{k,d,0}u](t)$	response due to zero end conditions;
$\eta_{u,k}$	vector of end conditions;
$\zeta_k(t)$	vector whose ith element is the response at time t due to the end conditions e_i, where e_i is the unit vector whose ith element is unity.

Similarly, we have

$$[L_k y](t) \approx [L_{k,d} y](t) = [L_{k,d,0} y](t) + \zeta_k(t)^T \eta_{y,k} \tag{45}$$

for $k = 0, 1, 2, \cdots$.

Substituting (44) and (45) into (18), we have

$$\begin{aligned} &\overline{a}_n [L_{n,d,0} y](t) + \overline{a}_{n-1} [L_{n-1,d,0} y](t) + \cdots \\ &\quad + \overline{a}_0 [L_{0,d,0} y](t) + \sum_{i=0}^{n} \overline{a}_i \zeta_i(t)^T \eta_{y,i} \\ &\approx \overline{b}_n [L_{n,d,0} u](t) + \overline{b}_{n-1} [L_{n-1,d,0} u](t) + \cdots \\ &\quad + \overline{b}_0 [L_{0,d,0} u](t) + \sum_{i=0}^{n} \overline{b}_i \zeta_i(t)^T \eta_{u,i} \\ &\quad + \overline{c}_n \ell_n(t) + \overline{c}_{n-1} \ell_{n-1}(t) + \cdots + \overline{c}_0 \ell_0(t). \end{aligned} \tag{46}$$

Define the matrices

$$\overline{F} = \begin{bmatrix} \zeta_0(0)^T & \zeta_1(0)^T & \cdots & \zeta_n(0)^T \\ \vdots & \vdots & \vdots & \\ \zeta_0(N)^T & \zeta_1(N)^T & \cdots & \zeta_n(N)^T \end{bmatrix} \tag{47}$$

and

$$M_2 = [\overline{W} \quad \overline{F} \quad \overline{U}_d \quad \overline{Y}_d] \tag{48}$$

where $\overline{U}_d$ and $\overline{Y}_d$ are defined in a similar way as $\overline{U}$ and $\overline{Y}$ in (22) and (23), except their elements are given by $[L_{k,d,0}u](t)$ and $[L_{n,d,0}y](t)$. Under the conditions that the input is persistently exciting and that the error in approximating the continuous-time filter $L_k(s)$ is negligible, then the coefficients $\overline{a}_i$ and $\overline{b}_i$ can be retrieved from the null space of the matrix M_2.

Since both the second- and third-order Padé approximations of $L_k(s)$ have $(k+1)$ unstable poles, the number of columns in the matrix $\overline{F}$ as defined above contains $((n+1)(n+2))/2$ columns, of which many are identical when $n \geq 2$. This is due to both the structure of the Laguerre functions and the preservation of allpass property by Padé approximations. By exploiting these structures, we can reduce the number of columns in $\overline{F}$ to $(2n+1)$.

B. Parameter Transformation

In this section, we derive a transformation that relates the coefficients $\overline{a}_i$ and $\overline{b}_i$ in (18) to those in the differential equation (1): a_i and b_i. For convenience, we define a_n to be 1.

From (11) and (13), we have

$$A(s) = \left(\frac{s+a}{2a}\right)^n \overline{A}\left(\frac{s-a}{s+a}\right) \tag{49}$$

$$= \frac{1}{(2a)^n} \sum_{j=0}^{n} \overline{a}_i (s-a)^j (s+a)^{n-j}. \tag{50}$$

By comparing the coefficients on both sides of the above equation, we arrive at the relation

$$\alpha = \Phi \times \frac{1}{(2a)^n} \times \overline{\alpha} \tag{51}$$

where

$$\alpha = [a_n \quad \cdots \quad a_0]^T \quad \overline{\alpha} = [\overline{a}_n \quad \cdots \quad \overline{a}_0]^T \tag{52}$$

and Φ is a $(n+1)$ by $(n+1)$ matrix whose $(n+1-i, n+1-j)$ element is the coefficient of s^i in the polynomial $(s-a)^j(s+a)^{n-j}$. In addition, the matrix Φ is invertible if and only if $a \neq 0$.

Recall from Section III that the polynomial $\overline{B}(w)$ is related to $B(s)$ in exactly the same manner; hence, with β and $\overline{\beta}$ defined similarly as α and $\overline{\alpha}$, respectively, we have

$$\beta = \Phi \times \frac{1}{(2a)^n} \times \overline{\beta}. \tag{53}$$

C. Computing the Coefficients $\overline{a}_i$ and $\overline{b}_i$

At this stage, we assumed that we have the matrix M computed according to (48) if filtering schemes such as ZOH, FOH, or the Tustin transform is used or, according to (43), if the filter used is a mixed causal/anticausal one. Since we are only interested in computing the coefficients a_i and b_i, we shall first perform a QR factorization on the matrix M. This operation also reduces the amount of data that we have to handle later. It is reasonable to assume that the matrix M has more rows than columns, and we have

$$M = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{bmatrix}. \tag{54}$$

Here, the number of rows in the matrix R_{12} is the same as the number of columns in either $\overline{W}$ or $[\overline{W} \quad \overline{F}]$ [see (24) and (48)], whichever is appropriate. Furthermore, R_{22} is square matrix of dimension $2(n+2)$ and

$$R_{22}\overline{p} \approx 0 \quad \text{where} \quad \overline{p} = \begin{bmatrix} \overline{\beta} \\ \overline{\alpha} \end{bmatrix} \tag{55}$$

where the degree of approximation depends on the integration error introduced by the filter approximation scheme. Note that the solution to the above problem is not unique as the length of the vector $\overline{p}$ is unknown. The extra constraint required to guarantee this uniqueness is obtained from the assumption that a_n is equal to 1. From the first row of (51), we arrive at the following constraint on the parameter vector $\overline{p}$.

$$[\Phi_1^T \quad 0_{1\times(n+1)}]\overline{p} = 1 \tag{56}$$

where Φ_i^T is the ith row of the matrix Φ.

Note that (51) and (53) may be used to introduce further constraints on the coefficients a_i and b_i. For example, if it is known that the system is strictly proper, i.e., $b_n = 0$, and b_{n-1} equals to the constant χ, then we have the following constraints:

$$\begin{bmatrix} 0_{1\times(n+1)} & \Phi_1^T \\ 0_{1\times(n+1)} & \Phi_2^T \end{bmatrix}\overline{p} = \begin{bmatrix} 0 \\ \chi \end{bmatrix}. \tag{57}$$

Furthermore, we can also impose constraints that are linear in the coefficients a_i and b_i. Finally, we would like to remark that similar methods to impose constraints (such as the relative degree or the number of integrators) are also mentioned in [9] and [12].

Let us assume that all the constraints imposed on $\overline{p}$ are summarized in the matrix equation

$$C\overline{p} = d \tag{58}$$

and the total number of constraints is q. We can now obtain the parameter estimate $\hat{\overline{p}}$ by solving the following total least-squares problem:

$$\min \|\Delta\|_F^2 \text{ such that } (R_{22}+\Delta)\hat{\overline{p}} = 0 \quad \text{and} \quad C\hat{\overline{p}} = d.$$

The solution to the above estimation problem is given in [3], or it can also be reformulated as a restricted total least squares problem [18]. The key step in solving the above problem is to reformulate it as follows: Find the smallest Δ such that the equation below is consistent.

$$\underbrace{\begin{bmatrix} C & -d \\ R_{22}+\Delta & 0 \end{bmatrix}}_{H} \begin{bmatrix} \hat{\overline{p}} \\ 1 \end{bmatrix} = 0. \tag{59}$$

If $\Delta = 0$, the matrix H, which is of dimension $(2n+2+q)\times(2n+2+1)$ (where $q \geq 1$), is generically of rank $(2n+2+1)$. In order to obtain a consistent solution to the above linear equation, H must be modified such that its rank becomes $(2n+2)$. For example, consider the simple case where $q = 1$ and $d \neq 0$; we have

$$\text{rank}(H) = \text{rank}([C \quad -d]) + \text{rank}(R_{22}+\Delta) \tag{60}$$

and therefore the original problem becomes one of finding the Δ, which has the smallest Frobenius norm such that the rank of $(R_{22}+\Delta)$ is $(2n+2)$. Let $U\,\text{diag}(\sigma_1, \cdots, \sigma_{2n+2})V^T$ be the SVD of R_{22}; then, the required Δ is given by $-U\,\text{diag}(0, \cdots, 0, \sigma_{2n+2})V^T$.

Remark 5.1: Although the above algorithmic description assumes that discrete filters obtained from the same approximation scheme are applied to both $u(t)$ and $y(t)$, the above algorithm can be modified such that filters from two different approximation schemes are applied, respectively, to $u(t)$ and $y(t)$. This extra degree of freedom is useful in some circumstances. For example, if it is known that the input $u(t)$ is a piecewise constant signal over the sampling interval, then ZOH approximation should be used as the Padé schemes will try to smooth the signal out.

Remark 5.2: Since the Padé approximations of Laguerre filters can be implemented as a tapped delay line alike structure (see Section IV-B2), this means the $(n+1)$ columns of $\overline{U}_d$ can be obtained from $(n+1)$ pth-order filters, where p is the order of approximation. For a n_x-state, n_i-input, n_o-output state space system, it takes $(n_x+n_i)(n_x+n_o)$ multiplications to simulate one time step. Thus, the total number of operations for all these filters per time step is $(n+1)(p+1)^2$. On the other hand, as the higher order hold discrete approximations of continuous-time Laguerre filters do not take this special structure, we need to use a first-order, a second-order, ... and

an $(n+1)$th-order filters with $(p+1)$ inputs, where p is the order of the hold circuit. The operation count for hold circuit is therefore $(1+p)(1+1)+(2+p)(2+1)+\cdots+(n+1+p)(n+1+1) \approx (n^3/3)+(p+1)(n^2/2)$ and is likely to be more than that needed by Padé filters.

Remark 5.3: Due to the use of anticausal filtering, the above algorithm cannot be implemented on line. However, since it can cope with both unknown initial and end conditions, it can be used in a quasi-on-line manner to cope with a large data set or to combine data obtained from different experiments. For example, assume that data $\{u(k), y(k)\}$ for $k=1,\cdots,2N$ is available, and we first divide it into two batches with the first batch consists of data $k=1,\cdots,N$ and the rest in the second one. From the first data batch, we arrive at the intermediate result $R_{22}^{[1]}\overline{p} \approx 0$ [cf. (55)] and note that this relation is independent of the unknown initial and end conditions. We then process the second data batch as though it is a fresh set of data, and we obtain the intermediate result $R_{22}^{[2]}\overline{p} \approx 0$. Finally, $\overline{p}$ can be estimated by solving these two sets of linear equations together with $R_{22}^{[1]}$ multiplied by a forgetting factor if needed. Note that there is a lower limit to the number of data points in each data batch to be processed this way; it is bounded from below by the number of columns in M_1 (43) or M_2 (48), whichever is appropriate.

VI. Output Corrupted by White Noise

In this section, we will show how to compute the coefficients of the differential equation (1) when the measured output $z(t)$ is corrupted by discrete-time white noise $e(t)$ of zero mean and variance λ^2/T. In other words, the measured output is given by

$$z(t) = y(t) + e(t) \tag{61}$$

where $y(t)$ is the noise-free output given by (1). Our parameter estimates will be analyzed based on the concept of *asymptotic consistency* [1]. Let $\hat{p}$ be an estimate of the true parameter vector p; then, $\hat{p}$ is said to be asymptotic consistent if

$$\lim_{T\to 0}\lim_{N\to\infty} \hat{p} = p \tag{62}$$

where N denotes the number of samples.

In the first step of our algorithm, we filter $u(t)$ and $z(t)$ through some discrete approximations of Laguerre filters $L_k(s)$. We shall, as before, assemble these filtered quantities in a matrix

$$M = [\overline{W} \quad \overline{U}_d \quad \overline{Z}_d] \tag{63}$$

where $\overline{W}$ contains all the terms used for compensating the effect of unknown initial and end conditions. The matrix $\overline{Z}_d$ contains the filtered measured output and can, in fact, be written as the sum of the filtered noise-free signal $\overline{Y}_d$ and the filtered noise $\overline{E}_d$, i.e.,

$$\overline{Z}_d = \overline{Y}_d + \overline{E}_d. \tag{64}$$

Let $\overline{\alpha}$ and $\overline{\beta}$ be the vectors as defined in Section V-B; then, the noise-free continuous-time signals satisfy the relation

$$[\overline{W} \quad \overline{U} \quad \overline{Y}]\begin{bmatrix}\overline{\xi}\\ \overline{\beta}\\ \overline{\alpha}\end{bmatrix} = 0. \tag{65}$$

Here, $\overline{\xi}$ contains all the coefficients that are due to unknown initial and end conditions, but they are not of interest to us. In order to compute $\overline{\alpha}$ and $\overline{\beta}$, we need to modify the matrix $\overline{Z}_d$ by a matrix Δ, which depends on the covariance matrix of $\overline{E}_d$ such that the equation

$$[\overline{W} \quad \overline{U}_d \quad \overline{Z}_d + \Delta]\begin{bmatrix}\overline{\xi}\\ \overline{\beta}\\ \overline{\alpha}\end{bmatrix} = 0 \tag{66}$$

has a solution. We shall now state the theorem in this section.

Theorem 6.1: Given the QR factorization

$$[\overline{W} \quad \overline{U}_d \quad \overline{Z}_d] = Q \times \begin{bmatrix} R_{11} & R_{12} & R_{13}\\ 0 & R_{22} & R_{23}\\ 0 & 0 & R_{33}\\ 0 & 0 & 0\end{bmatrix} \tag{67}$$

and under the condition that the matrices $\overline{W}$ and $\overline{U}_d$ have full rank, the limits

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{\sqrt{N}}\left(R_{11}\overline{\xi} + R_{12}\overline{\beta} + R_{13}\overline{\alpha}\right) = 0 \tag{68}$$

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{\sqrt{N}}\left(R_{22}\overline{\beta} + R_{23}\overline{\alpha}\right) = 0 \tag{69}$$

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{N} R_{33}^T R_{33}\overline{\alpha} = \kappa\lambda^2\overline{\alpha} \tag{70}$$

hold, where κ is given in Theorem 4.3.

Proof: From (65) and Theorem 4.3, it can be shown that

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{N}\begin{bmatrix}\overline{W}^T\\ \overline{U}_d^T\\ \overline{Z}_d^T\end{bmatrix}[\overline{W} \quad \overline{U}_d \quad \overline{Z}_d]\begin{bmatrix}\overline{\xi}\\ \overline{\beta}\\ \overline{\alpha}\end{bmatrix} = \lim_{T\to 0}\begin{bmatrix}0\\ 0\\ \lambda^2\Pi\overline{\alpha}\end{bmatrix} = \begin{bmatrix}0\\ 0\\ \kappa\lambda^2\overline{\alpha}\end{bmatrix}. \tag{71}$$

After substituting the given QR factorization into the above equation, the first block row gives

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{N} R_{11}^T(R_{11}\overline{\xi} + R_{12}\overline{\beta} + R_{13}\overline{\alpha}) = 0. \tag{72}$$

Under the assumption that $\overline{W}$ has full rank, the matrix

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{\sqrt{N}} R_{11}^T \tag{73}$$

is therefore nonsingular, and hence, we have (68).

The second block row of (71) gives

$$\lim_{T\to 0}\lim_{N\to\infty} \frac{1}{N}[R_{12}^T(R_{11}\overline{\xi} + R_{12}\overline{\beta} + R_{13}\overline{\alpha}) + R_{22}^T(R_{22}\overline{\beta} + R_{23}\overline{\alpha})] = 0. \tag{74}$$

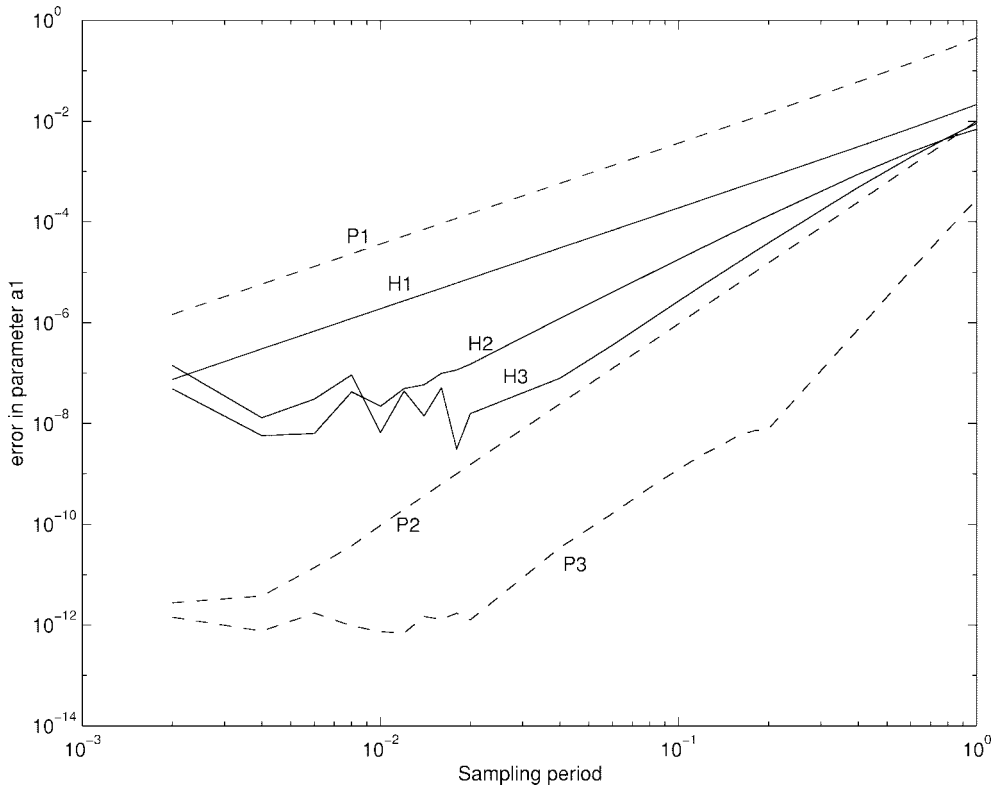


Fig. 3. Error in estimating parameter a_1 using different filter approximation schemes. Keys: solid lines for hold circuits (labeled as "H") and dashed lines for Padé approximations (labeled as "P"); the number after the letter indicates the order of approximation.

By using the assumption that $\overline{U}_d$ is of full rank as well as (68), we arrive at (69).

Finally, we consider the third block row of (71), which is

$$\lim_{T\to 0}\lim_{N\to\infty}\frac{1}{N}[R_{12}^T(R_{11}\overline{\xi}+R_{12}\overline{\beta}+R_{13}\overline{\alpha}) + R_{22}^T(R_{22}\overline{\beta}+R_{23}\overline{\alpha})+R_{33}^T R_{33}\overline{\alpha}]=\kappa\lambda^2\overline{\alpha}. \tag{75}$$

We then obtain (70) as a consequence of (68) and (69). QED

This theorem suggests that we can solve for $\overline{\alpha}$ and $\overline{\beta}$ in two separate steps. For a sufficiently large N and a nonzero T, (70) can be written as

$$\frac{1}{N}R_{33}^T R_{33}\overline{\alpha}\approx\lambda^2\Pi\overline{\alpha} \tag{76}$$

where Π is defined in (38). We have shown that Π becomes κI as T tends to zero, and our experience shows that this is a good approximation if T is small enough. However, we found that better results are obtained if we compute Π for a nonzero T, and we shall show how this can be done in Appendix C. The above equation says that $\overline{\alpha}$ is given by the solution of a generalized eigenvalue problem. Alternatively, the problem can be solved by using generalized singular value decomposition [4], and this is given in Appendix D. Finally, we can solve for $\overline{\beta}$ by using (69) as

$$R_{22}\overline{\beta}+R_{23}\overline{\alpha}=0. \tag{77}$$

Similar to the algorithm for the noise-free case given in Section V-C, we can also impose linear constraints on the parameters of the differential equation. This can be done by first removing the noise in R_{33} by setting the smallest generalized singular value to zero, and then, the rest proceeds as before.

Remark 6.1: Following Remark 3.1, the above parameter estimation algorithm can be applied *mutatis mutandis* in the discrete-time case, when the measured output is corrupted by discrete-time white noise.

VII. Simulation Example

In this section, we present a simple simulation example to illustrate the properties of our proposed algorithm. This example is based on the following second-order continuous-time system:

$$G(s)=\frac{s-1}{(s+1)^2}. \tag{78}$$

The bandwidth of this system is approximately 0.16 Hz, and it is excited by a chirp signal that lasts for 200 s with starting and end frequencies at 0.02 and 0.14 Hz, respectively.

Study 1—Noise-Free Case: The aim of this study is to look at the effect of sampling interval on the accuracy of parameter estimates. In this experiment, the time window for observation is fixed at 200 s; hence, the number of data points available is inversely proportional to the sampling period. The discretization schemes used are first-, second-, and third-

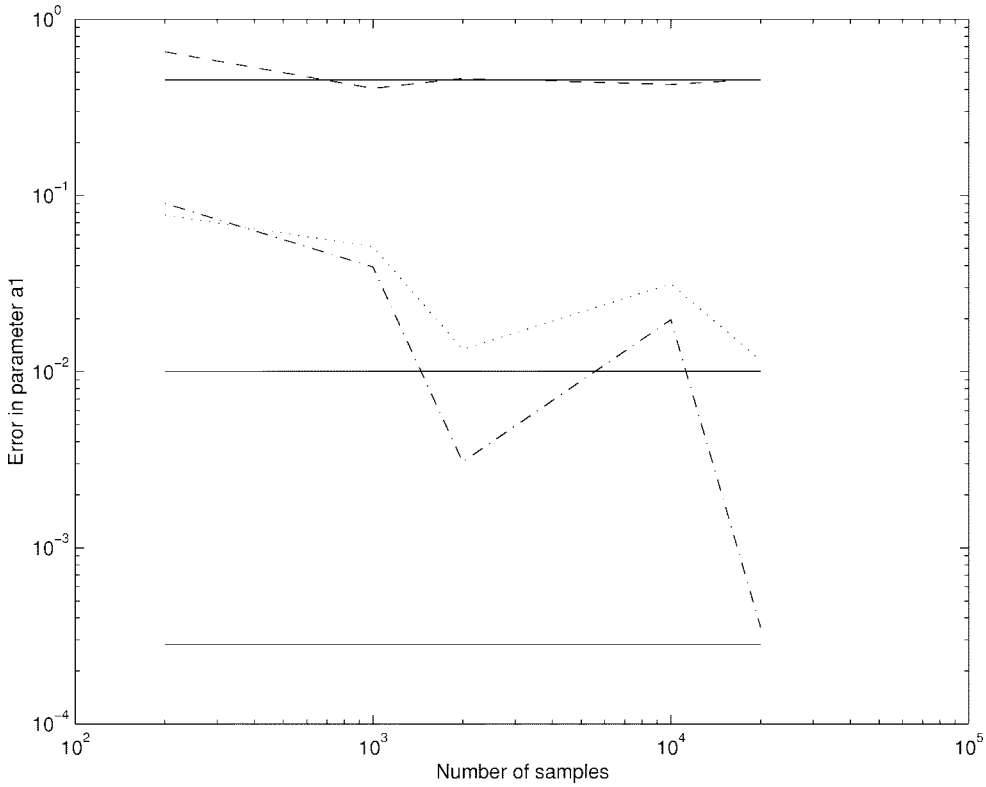


Fig. 4. Average error on the estimates of a_1 over 50 simulations. The solid lines indicate the error in the noise-free case and they are, from top to bottom Tustin, Padé second-order, Padé third-order, respectively. The other curves show the average error; keys: Tustin (dashed). Padé second-order (dotted). Padé third-order (dash-dot).

order hold circuits with $m = p - 1$ (see Appendix E for the meaning of m, p and how these HOH's are implemented) as well as Padé approximations of orders 1 to 3. Fig. 3 shows the estimation error for the parameter a_1, and the estimation errors for the other coefficients show a similar pattern. Note that there is almost a one-to-one correspondence between how well a parameter is estimated (Fig. 3) and how well a discretization scheme works (Fig. 2); this means that Padé approximation generally gives a smaller estimation error than higher order hold circuit of the same order. It can also be seen from the figure that if the schemes based on first- and second-order Padé approximations are to give the same accuracy, then we needs to sample 10 to 100 times faster in the former case.

Study 2—Noisy Case: In this experiment, the sampling frequency is 1 Hz, and white measurement noise that corresponds to a signal-to-noise ratio of 36 dB is added to the output. Our identification algorithm based on Padé approximations is used to estimate the parameters. Altogether, 50 different noise realizations are used, and the mean parameter estimates are computed. In Fig. 4, we plot the mean of the error on the parameter a_1 against the number of samples used in identification. There are two sources of error: variance error due to noise and integration error due to finite sampling interval, which is given by the error in noise-free case. For the Tustin transform, the two sources of error are of a similar order of magnitude. For both second- and third-order approximations, the variance error dominates for small sample sizes, and the error converges to the noise-free level as the sample size grows.

VIII. Conclusions

In this paper, we have proposed a parameter estimation scheme for the identification of continuous-time transfer function based on an "ARX-like" Laguerre model and discrete-time filters that are obtained from applying higher order Padé approximation to continuous-time Laguerre functions. If the measured output is corrupted by white noise, the proposed scheme is shown to give an asymptotically consistent estimate. The identification scheme based on Padé approximation is compared with one that is based on higher order hold circuits, and it is found that the former algorithm gives more accurate parameter estimates. A comparison between the proposed scheme and an indirect identification method based on noise-free data generated by a simulation model of an induction motor shows that the proposed scheme gives superior parameter estimates [7].

Appendix A
Causal/Anticausal Filtering

The aim of this Appendix is to show how an unstable filter can be transformed into a form that is suitable for causal/anticausal filtering. We will start with a particular state space equivalent of the unstable filter where the A matrix is in real Schur form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \tag{79}$$

where the block A_{11} contains all the stable poles, whereas the

block A_{22} contains all the unstable poles. We can partition the state vector and the other state-space matrices conformally as A, as follows:

$$\begin{bmatrix} x_{k+1}^c \\ x_{k+1}^{ac} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_k^c \\ x_k^{ac} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u_k \tag{80}$$

$$y_k = [C_1 \quad C_2] \begin{bmatrix} x_k^c \\ x_k^{ac} \end{bmatrix} + D u_k. \tag{81}$$

First, perform the substitutions $z_k^c \leftarrow x_k^c$ and $z_k^{ac} \leftarrow x_{k+1}^{ac}$. By expressing z_{k+1}^c, z_{k-1}^{ac}, and y_k as the subject of the equations, in terms of z_k^c, z_k^{ac}, and u_k, we arrive at the following form after some algebraic manipulations:

$$\begin{cases} z_{k+1}^c = \tilde{A}_{11} z_k^c + \tilde{A}_{12} z_k^{ac} + \tilde{B}_1 u_k \\ z_{k-1}^{ac} = \tilde{A}_{22} z_k^{ac} + \tilde{B}_2 u_k \\ y_k = \tilde{C}_1 z_k^c + \tilde{C}_2 z_k^{ac} + \tilde{D} u_k \end{cases} \tag{82}$$

where

$$\tilde{A}_{11} = A_{11} \tag{83}$$
$$\tilde{A}_{12} = A_{12} A_{22}^{-1} \tag{84}$$
$$\tilde{A}_{22} = A_{22}^{-1} \tag{85}$$
$$\tilde{B}_1 = B_1 - A_{12} A_{22}^{-1} B_2 \tag{86}$$
$$\tilde{B}_2 = -A_{22}^{-1} B_2 \tag{87}$$
$$\tilde{C}_1 = C_1 \tag{88}$$
$$\tilde{C}_2 = C_2 A_{22}^{-1} \tag{89}$$
$$\tilde{D} = D - C_2 A_{22}^{-1} B_2. \tag{90}$$

The set of equations given in (82) can now be used for simulation. The procedure is to first solve for z_k^{ac} using the second state transition equation; the computed z_k^{ac} can then be used to solve for z_k^c using the first state transition equation.

Example 8.1: As an illustration, consider the following state space system with both stable and unstable modes:

$$x_{k+1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_k \tag{91}$$

$$y_k = [1 \quad 1] x_k. \tag{92}$$

Let us assume that the initial condition $x_0 = 0$, and the input is $\delta(k-5)$, where δ is the discrete-time impulse function. Then, the system response is

$$y_k = \begin{cases} 0, & \text{for } k < 6 \\ \left(\frac{1}{2}\right)^{k-6} + 3^{k-6}, & \text{for } k \geq 6. \end{cases} \tag{93}$$

The system transformation given above will lead us to the following causal/anticausal system:

$$\begin{cases} z_{k+1}^c = \frac{1}{2} z_k^c + u_k \\ z_{k-1}^{ac} = \frac{1}{3} z_k^{ac} - \frac{1}{3} u_k \\ y_k = z_k^c + \frac{1}{3} z_k^{ac} - \frac{1}{3} u_k. \end{cases} \tag{94}$$

Since $A_{12} = 0$, the two state transition equations are independent of each other, and they may therefore be solved independently. We begin by solving the first state transition equation with $z_0^c = 0$ and $u(k) = \delta(k-5)$, and its solution is

$$z_k^c = \begin{cases} 0, & \text{for } k < 6 \\ \left(\frac{1}{2}\right)^{k-6}, & \text{for } k \geq 6. \end{cases} \tag{95}$$

With the final condition $z_N^{ac} = \eta$ (where $N \geq 5$) and $u(k) = \delta(k-5)$, the second state transition equation gives

$$z_k^{ac} = \begin{cases} \left(\frac{1}{3}\right)^{N-k} \eta, & \text{for } N \geq k \geq 5 \\ \left(\frac{1}{3}\right)^{4-k} \left(\left(\frac{1}{3}\right)^{N-4} \eta - \frac{1}{3} \right), & \text{for } k \leq 4. \end{cases} \tag{96}$$

The output of the causal/anticausal system can now be computed and is equal to

$$y_k = \begin{cases} \left(\frac{1}{2}\right)^{k-6} + \left(\frac{1}{3}\right)^{N-k+1} \eta, & \text{for } k \geq 6 \\ \left(\frac{1}{3}\right)^{N-4} \eta - \frac{1}{3}, & \text{for } k = 5 \\ \left(\frac{1}{3}\right)^{5-k} \left(\left(\frac{1}{3}\right)^{N-4} \eta - \frac{1}{3} \right), & \text{for } k \leq 4. \end{cases} \tag{97}$$

By comparing this with the output of the original system, it can readily be seen that these two representations give identical output if the end condition η is chosen to be 3^{N-5}. Let $H(k)$ denote the discrete-time unit step; then, the output of the causal/anticausal system may be written compactly as

$$y_k = \left(\frac{1}{2}\right)^{k-6} H(k-6) - \left(\frac{1}{3}\right)^{6-k} H(6-k) + \left(\frac{1}{3}\right)^{N-k+1} \eta. \tag{98}$$

This illustrates that if the initial condition is zero, the output can be written as the sum of the response due to the input with zero final condition and the response due zero input with an end condition. Furthermore, the response due to zero input with an end condition is linear function of the end condition. These facts are used in Section V-A to derive the system identification algorithm.

APPENDIX B
PROOFS

Proof of Theorem 4.2: From (26) and the fact that the pth-order Padé approximation has error of order $(2p+1)$ [11], we have

$$\frac{1}{T} \log(1+x) = \rho_p(1+x) + O(x^{2p+1}). \tag{99}$$

Substituting $x = e^{sT} - 1 = sT + O((sT)^2)$, we have

$$\rho_p(e^{sT}) = s + O(T^{2p+1}). \tag{100}$$

By writing $F(s)$ as a ratio of two polynomials $(p(s)/q(s))$, we have

$$F(s) - F(\rho_p(e^{sT})) = \frac{p(s)q(e^{sT}) - q(s)p(e^{sT})}{q(s)q(e^{sT})}. \tag{101}$$

By substituting (100) in (101), we arrive at the result stated in Theorem. △

Proof of Theorem 4.3: Let us define the transfer function

$$H_{ij}(z) = \frac{1}{T} L_{|i-j|+1,\,d}(z) L_{1,\,d}\left(\frac{1}{z}\right) \tag{102}$$

and let its two-sided Z-transform or Laurent series expansion [15] be given by

$$\begin{aligned} H_{ij}(z) = \cdots + h_{i,\,j;\,-2} z^{-2} + h_{i,\,j;\,-1} z^{-1} + h_{i,\,j;\,0} \\ + h_{i,\,j,\,1} z + h_{i,\,j;\,2} z^2 + \cdots. \end{aligned} \tag{103}$$

It can be shown that the $(i,\,j)$ element of the covariance matrix Π (38) is given by the constant term of the Laurent series expansion, i.e., $h_{i,\,j;\,0}$ of $H_{ij}(z)$. Therefore, Theorem 4.3 can be restated as

$$\lim_{T\to 0} h_{i,\,j;\,0} = \begin{cases} \kappa, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \tag{104}$$

where κ takes a value that depends on the order of approximation.

We will use $p_{s,\,i}$ and $p_{us,\,i}$ to denote, respectively, the stable and unstable poles of the transfer function $H_{ij}(z)$; then, the Laurent series expansion converges for all complex z within the annulus $A = \{z \in C\colon \max|p_{s,\,i}| < |z| < \min|p_{us,\,i}|\}$. It can be shown, by using the expressions of the transfer functions, that for any $aT > 0$, the poles of $H_{ij}(z)$ do not lie on the unit circle. Therefore, the set A is nonempty. The coefficients $h_{i,\,j;\,0}$ can be computed from the integral

$$h_{i,\,j;\,0} = \frac{1}{2\pi j} \oint_{\gamma} \frac{H_{ij}(z)}{z}\, dz \tag{105}$$

where γ is any contour within the annulus A, and without loss of generality, we shall choose γ to be the unit circle. Let $\text{res}(f,\,z)$ denote the residue of the function f at the isolated singularity z; then, by the Cauchy's residue theorem, the coefficient $h_{i,\,j;\,0}$ is given by

$$h_{i,\,j;\,0} = \text{res}\left(\frac{H_{ij}(z)}{z},\,0\right) + \sum_i \text{res}\left(\frac{H_{ij}(z)}{z},\,p_{s,\,i}\right) \tag{106}$$

where the second sum is over the different isolated stable poles of $H_{ij}(z)$.

The proof will be divided into three parts according to the type of approximation scheme used.

Case 1—Tustin Transformation: We first consider the situation where $i = j$. We have

$$\begin{aligned} H_{ii}(z) &= L_{1,\,d}(z) L_{1,\,d}\left(\frac{1}{z}\right) \\ &= \frac{2aT}{(2+aT)^2} \frac{(z+1)^2}{(z-\alpha)(z+\alpha)} \end{aligned} \tag{107}$$

where α is $(2-aT)/(2+aT)$. By computing the appropriate residues, it can be shown that

$$h_{i,\,i;\,0} = \frac{2}{2+aT} \tag{108}$$

and therefore

$$\lim_{T\to 0} h_{i,\,i;\,0} = 1. \tag{109}$$

For the case where $i \neq j$, let $k = |i-j|$, and we have

$$\begin{aligned} H_{ij}(z) &= L_{k,\,d}(z) L_{1,\,d}\left(\frac{1}{z}\right) \\ &= \frac{-2aT}{(2+aT)^2} \frac{(z+1)^2(\alpha z - 1)^{k-1}}{(z-\alpha)^{k+1}}. \end{aligned} \tag{110}$$

For $\alpha \neq T/2$, the coefficient $h_{i,\,j;\,0}$ is given by

$$h_{i,\,j;\,0} = \text{res}\left(\frac{H_{ij}(z)}{z},\,0\right) + \text{res}\left(\frac{H_{ij}(z)}{z},\,\alpha\right). \tag{111}$$

By using the expression of $H_{ij}(z)$ given above, it can be shown that

$$\text{res}\left(\frac{H_{ij}(z)}{z},\,0\right) = \frac{-2aT}{(2+aT)^2} \frac{1}{\alpha^{k+1}} \tag{112}$$

$$\text{res}\left(\frac{H_{ij}(z)}{z},\,\alpha\right) = \frac{-2aT}{(2+aT)^2} \times \lim_{z\to\alpha} \frac{1}{k!} \frac{d^k}{dz^k}\left(\frac{(z+1)^2(\alpha z-1)^{k-1}}{z}\right) \tag{113}$$

$$= \frac{-2aT}{(2+aT)^2} \frac{1}{k} \frac{\alpha^{k-1}(\alpha^2-1)}{\alpha^2}. \tag{114}$$

Since $\alpha \to 1$ as $T \to 0$, we have

$$\lim_{T\to 0} h_{i,\,j;\,0} = 0. \tag{115}$$

Case 2—Second-Order Padé Approximations: The discrete approximation of the filter $\sqrt{2a}/(s+a)$ using second-order Padé approximation can be written as

$$\frac{\sqrt{2a}T}{3+aT} \frac{z^2+4z+1}{(z-p_1)(z-p_2)} \tag{116}$$

where p_1 and p_2 are the roots of the equation $(aT+3)z^2 + 4aTz + (aT-3) = 0$. It can be shown by using root locus that for any positive T, this equation always has two distinct real roots, and one of them is always stable, whereas the other one is always unstable. We shall assume here that p_1 is the stable pole and that p_2 is the unstable pole. It can further be shown that

$$\lim_{T\to 0} p_1 = 1 \tag{117}$$

$$\lim_{T\to 0} p_2 = -1. \tag{118}$$

We first consider the case where $i = j$. We have

$$\begin{aligned} H_{ii}(z) = &\frac{2aT}{(3+aT)(aT-3)} \\ &\times \frac{(z^2+4z+1)^2}{(z-p_1)(z-p_2)\left(z-\frac{1}{p_1}\right)\left(z-\frac{1}{p_2}\right)}. \end{aligned} \tag{119}$$

Provided that $T \neq 3/a$ so that $p_1 \neq 0$, we have

$$\begin{aligned} h_{i,\,i;\,0} = &\text{res}\left(\frac{H_{ii}(z)}{z},\,0\right) + \text{res}\left(\frac{H_{ii}(z)}{z},\,p_1\right) \\ &+ \text{res}\left(\frac{H_{ii}(z)}{z},\,\frac{1}{p_2}\right). \end{aligned} \tag{120}$$

By evaluating these residues, it can be shown that

$$\begin{aligned}&\operatorname{res}\left(\frac{H_{ii}(z)}{z}, 0\right)\\&\quad=\frac{2aT}{(3+aT)(aT-3)}\end{aligned}\tag{121}$$

$$\begin{aligned}&\operatorname{res}\left(\frac{H_{ii}(z)}{z}, p_1\right)+\operatorname{res}\left(\frac{H_{ii}(z)}{z}, \frac{1}{p_2}\right)\\&\quad=\frac{-1}{aT-3}\,\frac{2(p_1p_2-1)^2}{p_1(p_1-p_2)}.\end{aligned}\tag{122}$$

Finally, by taking the limit

$$\lim_{T\to 0} h_{i,i;0}=\tfrac{4}{3}.\tag{123}$$

The proof for the case where $i \neq j$ is similar to that of Case 1 and is omitted. See [2] for details.

Case 3—Third-Order Padé Approximations: The proof for this case is similar to those of the previous two cases and is omitted. The details can be found in [2].

Appendix C
Computing Π

We see in Appendix B that the (i, j) element of Π is given by the constant term in the Laurent series expansion of the causal/anticausal transfer function

$$H_{ij}(z)=L_{|i-j|+1,\,d}(z)L_{1,\,d}\left(\frac{1}{z}\right).\tag{124}$$

Let us assume that $H_{ij}(z)$ has already been converted to the form given in (82); then, the required constant term is also given by the unit impulse response of $H_{ij}(z)$ at time $t=0$ subject to the initial condition $z^c_{-\infty}=0$ and end condition $z^{ac}_{\infty}=0$.

Using the state transition equation for the anticausal state in (82), it is straightforward to show that

$$z^{ac}_k=\begin{cases}0, & \text{for } k=0,1,\cdots\\ \tilde{A}_{22}^{-k-1}\tilde{B}_2, & \text{for } k=-1,-2,\cdots\end{cases}\tag{125}$$

By using this result and the state transition equation for the causal state, we have

$$z^c_0=\sum_{k=0}^{\infty}\tilde{A}^k_{11}\tilde{A}_{12}z^{ac}_{-1-k}\tag{126}$$

$$=\underbrace{\left(\sum_{k=0}^{\infty}\tilde{A}^k_{11}\tilde{A}_{12}\tilde{A}^k_{22}\right)}_{P}\tilde{B}_2.\tag{127}$$

The matrix P can be shown to satisfy the Sylvester equation

$$P-\tilde{A}_{11}P\tilde{A}_{22}=\tilde{A}_{12}.\tag{128}$$

Finally, the required answer is given by

$$\Pi_{ij}=\tilde{C}_1P\tilde{B}_2+\tilde{D}.\tag{129}$$

Appendix D
Solving for $\overline{\alpha}$

Recall from Section VI that $\overline{\alpha}$ is given by the solution of the equation

$$\frac{1}{N}R^T_{33}R_{33}\overline{\alpha}=\lambda^2\Pi\overline{\alpha}.\tag{130}$$

Let L be a Cholesky factor of Π, i.e., $\Pi=LL^T$. The above equation can be solved by using the generalized SVD [4] of the matrices $(1/\sqrt{N})R_{33}$ and L^T as

$$\frac{1}{\sqrt{N}}R_{33}=V_RAP^{-1}\tag{131}$$

$$L^T=V_LBP^{-1}\tag{132}$$

such that both V_R and V_L are orthogonal matrices, and P is invertible. Both A and B are diagonal matrices with nonnegative elements

$$A=\operatorname{diag}(\alpha_1,\cdots,\alpha_{n+1})\tag{133}$$

$$B=\operatorname{diag}(\beta_1,\cdots,\beta_{n+1}).\tag{134}$$

Furthermore, the generalized singular values are defined as

$$s_i=\frac{\alpha_i}{\beta_i}.\tag{135}$$

By substituting (131) and (132) into (130), we have

$$(A^2-\lambda^2B^2)P^{-1}\overline{\alpha}=0.\tag{136}$$

This implies that an estimate for λ is given by the smallest generalized singular value, say, s_j; then, an estimate of $\overline{\alpha}$ is then given by the jth column of P.

Appendix E
Implementation of Higher Order Holds

This section discusses how the HOH's used in the simulation example in Section VII are implemented. These HOH's are specified by two parameters p and m, where p is the order, and m determines the part of the data that is used for interpolation. We assume the continuous-time filter is given in state space form

$$\dot{x}(t)=Ax(t)+Bu(t)\tag{137}$$

$$y(t)=Cx(t)+Du(t).\tag{138}$$

This gives the following discrete-time state equation

$$x((k+1)T)=e^{AT}x(kT)+\int_0^T e^{A\tau}Bu((k+1)T-\tau)\,d\tau.\tag{139}$$

The idea of pth (where $p>0$) order hold is to approximate $u((k+1)T-\tau)$ for $\tau\in[0,T]$ by a pth-order polynomial

$$\begin{aligned}u((k+1)T-\tau)&\approx g(k,\tau)\\&=g_0(k)+g_1(k)\tau+\cdots+g_p(k)\tau^p.\end{aligned}\tag{140}$$

The coefficients $g_i(k)$ are determined from the $(p+1)$ inputs $u((k-m)T),\cdots,u((k-m+p)T)$, where m is a prespecified

integer in $[0, p-1]$. Specifically, $g_i(k)$'s are determined by the interpolation conditions

$$u((k-r)T) = g(k, (r+1)T), \quad \text{for} \quad r = m, \cdots, m-p \tag{141}$$

and they have to be solved for each k. By replacing $u((k+1)T - \tau)$ by its approximation, (139) can be written as

$$x((k+1)T) = e^{AT}x(kT) + \sum_{i=0}^{p} g_i(k) \int_0^T e^{A\tau}\tau^i \, d\tau B. \tag{142}$$

Let $\mathcal{A}(i) = \int_0^T e^{A\tau}\tau^i \, d\tau$; these matrices can be computed by using the recursion

$$\mathcal{A}(0) = A^{-1}(e^{AT} - I) \tag{143}$$

$$\mathcal{A}(i+1) = A^{-1}(e^{AT}T^{i+1} - (i+1)\mathcal{A}(i)) \quad \text{for } i = 0, 1, \cdots \tag{144}$$

REFERENCES

[1] S. Bigi, T. Söderström, and B. Carlsson, "An IV-scheme for estimating continuous-time stochastic models from discrete-time data," in *Proc. 10th IFAC Symp. Syst. Identification*, Copenhagen, Denmark, 1994, vol. 3, pp. 645–650.

[2] C. T. Chou, M. Verhaegen, and R. Johansson, "Continuous-time identification of SISO systems using Laguerre functions," Elect. Eng. Dept., Delft Univ. Technol., Tech. Rep. TUD/ET/SCE96.003, 1996.

[3] J. W. Demmel, "The smallest perturbation of a submatrix which lowers the rank and constrained total least squares problems," *SIAM J. Numer. Anal.*, vol. 24, no. 1, pp. 199–206, 1987.

[4] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Baltimore, MD: Johns Hopkins Univ. Press, 1983.

[5] R. Johnasson, "Identification of continuous-time models," *IEEE Trans. Signal Processing*, vol. 42, pp. 887–897, 1994.

[6] E. W. Kamen, *Introduction to Signals and Systems*, 2nd ed. New York: Macmillan, 1990.

[7] R. F. F. Koning, C. T. Chou, M. H. G. Verhaegen, J. B. Klaassens, and J. R. Uittenbogaart, "A novel approach on parameter identification for inverter driven induction machines," *IEEE Trans. Contr. Syst. Technol.*, to be published.

[8] P. Lindskog, "Methods, algorithms and tools for system identification based on prior knowledge," Ph.D. dissertation, Linköping Univ., Linköping, Sweden, 1996.

[9] R. H. Middleton and G. C. Goodwin, *Digital Control and Estimation.* Englewood Cliffs, NJ: Prentice-Hall, 1990.

[10] M. Moonen, B. De Moor, and J. Vandewalle, "SVD-based subspace methods for multivariable continuous-time systems identification," in *Identification of Continuous-time Systems*, N. K. Sinha and G. P. Rao, Eds. Boston, MA: Kluwer, 1991, pp. 473–488.

[11] P. P. Petrushev and V. A. Popov, *Rational Approximation of Real Functions.* Cambridge, U.K.: Cambridge Univ. Press, 1987.

[12] R. Piche, personal communication.

[13] S. Sagara and Z.-Y. Zhao, "Numerical integration approach to on-line identification of continuous-time systems," *Automatica*, vol. 26, no. 1, pp. 63–74, 1990.

[14] J. Schoukens, "Modeling of continuous-time systems using a discrete-time representation," *Automatica*, vol. 26, no. 3, pp. 579–583, 1990.

[15] I. Stewart and D. Tall, *Complex Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1987.

[16] H. Unbenhauen and G. P. Rao, "Continuous-time approach to system identification—A survey," *Automatica*, vol. 26, no. 1, pp. 23–35, 1990.

[17] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor, "System identification with generalized orthonormal basis functions," *Automatica*, vol. 31, no. 12, pp. 1821–1834, 1995.

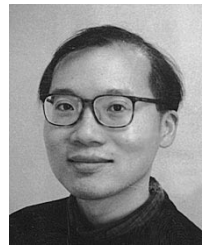
[18] S. Van Huffel and H. Zha, "The restricted total least squares problem: Formulation, algorithm and properties," *SIAM J. Matrix Anal. Appl.*, vol. 12, pp. 292–309, 1991.

[19] M. Verhaegen, D. Westwick, and K. Kearney, "The use of bilinear transformation of the shift operator in subspace model identification," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1422–1428, Aug. 1995.

[20] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans Automat. Contr.*, vol. 36, pp. 551–562, May 1990.

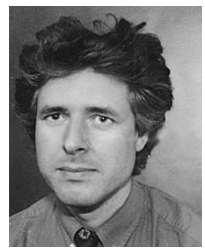
[21] ——, "System identification using Kautz models," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1276–1282, June 1994.

[22] P. C. Young, "Parameter estimation for continuous-time models—A survey," *Automatica*, vol. 17, no. 1, pp. 23–39, 1981.



Chun Tung Chou was born in Hong Kong in 1967. He received the B.A. degree in engineering science from the University of Oxford, Oxford, U.K., in 1988 and the Ph.D. degree in control engineering from University of Cambridge, Cambridge, U.K., in 1994.

He is currently a Post-Doctoral Research Fellow with the Department of Electrical Engineering, Delft University of Technology, Delft, The Netherlands. His current research interests are system identification of linear and nonlinear systems.



Michel Verhaegen was born in Belgium on September 2, 1959. He received the Bachelor's degree in engineering/aeronautics from the Delft University of Technology, Delft, The Netherlands, in August 1982 and the Doctoral degree in applied sciences from the Catholic University Leuven, Leuven, Belgium, in November 1985. During his graduate study, he held an IWONL Research Assistantship in the Department of Electrical Engineering.

From 1985 to 1987, he was a Postdoctoral Associate of the U.S. National Research Council, and during that time, he was affiliated with the NASA Ames Research Center, Mountain View, CA. From 1988 to 1989, he was an Assistant Professor with the Department of Electrical Engineering, Delft University of Technology. Since February 1990, he was a Senior Research Fellow of the Dutch Academy of Arts and Sciences, affiliated with the Network Theory Group of Delft University of Technology. Since 1990, he has held short sabbatical leaves at the University of Uppsala, Uppsala, Sweden, McGill University, Montreal, P.Q., Canada, Lund University of Technology, Lund, Sweden, and the German Aerospace Research Center (DLR), Munich. Currently, he is an Associate Professor with the Delft University of Technology and is Head of the Systems and Control Engineering Group of the Faculty of Information Technology and Systems. He is a project leader of a number of research projects sponsored by the European Community, the Dutch National Science Foundation (STW), and the Delft University of Technology. His main research interest is the interdisciplinary domain of numerical algebra and system theory. In this field, he has published more than 100 papers. Current activities focus on the development and practical application of new identification and simulation tools in model-based industrial controller design.

Dr. Verhaegen received a Best Presentation Award at the 1986 American Control Conference, Seattle, WA, and a Recognition Award from NASA in 1989.

Rolf Johansson received the M.S. degree in technical physics in 1977, the B.M. degree in 1980, and the Doctorate degree in 1983 in control theory. He was appointed Docent in 1985 and received the M.D. degree from Lund University of Technology, Lund, Sweden, in 1986.

Since 1986, he has been an Associate Professor of control theory at the Lund Institute of Technology, with side appointments at the Faculty of Medicine. Currently, he is Coordinating Director of Robotics Research with participants from several departments of Lund University. He has had visiting appointments at Laboratoire d'Automatique de Grenoble, Grenoble, France, California Institute of Technology, Pasadena, and Rice University, Houston, TX. In his scientific work, he has been involved in research in adaptive system theory, mathematical modeling, system identification, robotics, signal processing theory, as well as biomedical research. He has published the book *System Modeling and Identification* (Englewood Cliffs, NJ: Prentice-Hall, 1993).

Dr. Johansson was awarded the biomedical engineering prize (the Ebeling Prize) of the Swedish Society of Medicine in 1995.