Pergamon

0005-1098(95)00074-7

Automatica, Vol. 31, No. 12, pp. 1821–1834, 1995

0005-1098/95 $9.50 + 0.00

System Identification with Generalized Orthonormal Basis Functions*

PAUL M. J. VAN DEN HOF,† PETER S. C. HEUBERGER‡ and JÓZSEF BOKOR§

A set of flexible basis functions that generalizes the classical Laguerre and Kautz bases is shown to possess attractive properties when used in linear least-squares identification.

Key Words—System identification; orthogonal basis functions; FIR models; linear regression; modelling errors; system approximation.

Abstract—A least-squares identification method is studied that estimates a finite number of expansion coefficients in the series expansion of a transfer function, where the expansion is in terms of recently introduced generalized basis functions. The basis functions are orthogonal in $\mathcal{H}_2$, and generalize the pulse, Laguerre and Kautz bases. One of their important properties is that, when chosen properly, they can substantially increase the speed of convergence of the series expansion. This leads to accurate approximate models with only a few coefficients to be estimated. Explicit bounds are derived for the bias and variance errors that occur in parameter estimates as well as in the resulting transfer function estimates.

NOTATION

$(\cdot)^{\mathrm{T}}$	transpose of a matrix;
$\mathbb{R}^{p\times m}$	set of real-valued matrices with dimension $p \times m$;
$\mathbb{C}$	set of complex numbers;
$\mathbb{Z}_+$	set of nonnegative integers;
$\ell_2[0, \infty)$	space of squared summable sequences on the time interval $\mathbb{Z}_+$;
$\ell_2^{p\times m}[0, \infty)$	space of matrix sequences $\{F_k \in \mathbb{R}^{p\times m}\}_{k=0,1,2,\ldots}$ such that $\sum_{k=0}^{\infty} \operatorname{tr}(F_k^{\mathrm{T}}F_k)$ is finite;
$\mathcal{H}_2^{p\times m}$	set of real $p \times m$ matrix functions that are square-integrable on the unit circle;
$\lVert\cdot\rVert_2$	ℓ_2 norm of a vector; induced ℓ_2 norm or spectral norm of a constant matrix, i.e. its maximum singular value;
$\lVert\cdot\rVert_1$	ℓ_1 norm of a vector; induced ℓ_1 norm of a matrix operator;
$\lVert\cdot\rVert_\infty$	ℓ_∞ norm of a vector; induced ℓ_∞ norm of a matrix operator;
$\lVert\cdot\rVert_{\mathcal{H}_2}$	$\mathcal{H}_2$ norm of a stable transfer function.
$\bar{E}$	$\lim_{N\to\infty} (1/N) \sum_{t=1}^{N} E$;
$\otimes$	Kronecker matrix product;
e_i	ith Euclidean basis vector in $\mathbb{R}^n$;
I_n	$n \times n$ identity matrix;
$\delta(t)$	Kronecker delta function, i.e. $\delta(t) = 1$, $t = 0$; $\delta(t) = 0$, $t \neq 0$.
$:=$	'is defined by'.

The scalar transfer function $G(z)$ has an n_b-dimensional state-space realization (A, B, C, D), with $A \in \mathbb{R}^{n_b\times n_b}$, and B, C and D of appropriate dimensions, if $G(z) = C(zI - A)^{-1}B + D$. A realization is minimal if it has minimal dimension. The controllability Gramian P and observability Gramian Q are defined as the solutions to the Lyapunov equations $APA^{\mathrm{T}} + BB^{\mathrm{T}} = P$ and $A^{\mathrm{T}}QA + C^{\mathrm{T}}C = Q$ respectively, and the realization is called (internally) balanced if $P = Q = \Sigma$, with $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_{n_b})$, $\sigma_1 \geq \ldots \geq \sigma_{n_b}$. A system $G \in \mathcal{H}_2$ is called inner if it is stable and it satisfies $G(z^{-1})G(z) = 1$.

1. INTRODUCTION

The use of orthogonal basis functions for the Hilbert space $\mathcal{H}_2$ of stable systems has a long history in modelling and identification of

* Received 22 June 1994; revised 9 January 1995; received in final form 15 February 1995. The original version of this paper was presented at the 10th IFAC Symposium on System Identification, which was held in Copenhagen, Denmark, during 4–6 July 1994. The Published Proceedings of this IFAC meeting may be ordered from: Elsevier Science Limited, The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, U.K. This paper was recommended for publication in revised form by Associate Editor Bo Wahlberg under the direction of Editor Torsten Söderström.
Corresponding author Dr Paul M. J. Van den Hof. Tel. +31 15 784509; Fax +31 15 784717; E-mail vdhof@tudw03.tudelft.nl.

† Mechanical Engineering Systems and Control Group, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.

‡ National Institute of Public Health and Environmental Protection (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands.

§ Computer and Automation Research Institute, Hungarian Academy of Sciences, Kende u. 13-17, H 1502 Budapest, Hungary.

dynamical systems. The main part of this work dates back to the classical work of Lee (1933) and Wiener (1949), as also summarized in Lee (1960).

In the past few decades, orthogonal basis functions (e.g. the Laguerre functions) have been employed for the purpose of system identification by, for example, King and Paraskevopoulos (1979), Nurges and Yaaksoo (1981) and Nurges (1987). In these works the input and output signals of a dynamical system are transformed to a (Laguerre) transform domain, being induced by the orthogonal basis for the signal space. Consecutively, more or less standard identification techniques are applied to the signals in this transform domain. The main motivation for this approach has been directed towards data reduction, since the representation of the measurement data in the transform domain becomes much more efficient once an appropriate basis is chosen.

Wahlberg (1990, 1991, 1994a) applied orthogonal functions for the identification of a finite sequence of expansion coefficients. Given the fact that every stable system has a unique series expansion in terms of a prechosen basis, a model representation in terms of a finite-length series expansion can serve as an approximate model, where the coefficients of the series expansion can be estimated from input–output data.

Consider a stable system $G(z) \in \mathcal{H}_2$, written as

$$G(z) = \sum_{k=0}^{\infty} G_k z^{-k}, \tag{1}$$

with $\{G_k\}_{k=0,1,2,\ldots}$ the sequence of Markov parameters. Let $\{f_k(z)\}_{k=0,1,2,\ldots}$ be an orthonormal basis for the set of systems $\mathcal{H}_2$. Then there exists a unique series expansion

$$G(z) = \sum_{k=0}^{\infty} L_k f_k(z), \tag{2}$$

with $\{L_k\}_{k=0,1,2,\ldots}$ the (real-valued) expansion coefficients. The orthonormality of the basis is reflected by the property that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f_k(e^{i\omega}) f_l(e^{-i\omega})\, d\omega = \begin{cases} 1 & (k = l), \\ 0 & (k \neq l). \end{cases} \tag{3}$$

Note that $f_k(z) = z^{-k}$ is one of the possibilities for choosing such a basis. If a model of the system G is represented by a finite-length series expansion

$$\hat{G}(z) = \sum_{k=0}^{n-1} \hat{L}_k f_k(z) \tag{4}$$

then it is easily understandable that the accuracy of the model, in terms of the minimal possible deviation between system and model (in any prechosen norm), will be essentially dependent on the choice of basis functions $f_k(z)$. Note that the choice $f_k(z) = z^{-k}$ corresponds to the use of so-called FIR (finite impulse response) models (Ljung, 1987).

Since the accuracy of the models is limited by the basis functions, the development of appropriate basis functions is a topic that has attracted considerable interest. The issue here is that it is profitable to design basis functions that reflect the dominant dynamics of the process to be modelled.

Laguerre functions are determined by

$$f_k(z) = \sqrt{1-a^2}\, z \frac{(1-az)^k}{(z-a)^{k+1}}, \quad |a| < 1. \tag{5}$$

(see e.g. Gottlieb, 1938; Szegö, 1975), and they involve a scalar design variable a that has to be chosen in a range that matches the dominating (first-order) dynamics of the process to be modelled. Considerations for optimal choices of a are discussed, for example, by Clowes (1965) and Fu and Dumont (1993). For moderately damped systems, Kautz functions have been employed; these are actually second-order generalizations of the Laguerre functions (see Kautz 1954; Wahlberg, 1990, 1994a, b).

Recently a generalized set of orthonormal basis functions has been developed that is generated by inner (all-pass) transfer functions of any prechosen order (Heuberger and Bosgra, 1990; Heuberger, 1991; Heuberger et al., 1993, 1995). This type of basis functions generalizes the Laguerre and Kautz-type bases, which appear as special cases when choosing first-order and second-order inner functions. Given any inner transfer function (with any set of eigenvalues), an orthonormal basis for the space of stable systems $\mathcal{H}_2$ (and similarly for the signal space ℓ_2) can be constructed.

Using generalized basis functions that contain dynamics can have important advantages in identification and approximation problems. It has been shown by Heuberger et al. (1993, 1995) that if the dynamics of the basis generating system and the dynamics of the system to be modelled approach each other, the convergence rate of a series expansion of the system becomes very fast. Needless to say, the identification of expansion coefficients in a series expansion benefits very much from a fast convergence rate—the number of coefficients to be determined to accurately model the system becomes smaller. This concerns a reduction of both bias and variance contributions in the estimated models.

In this paper, we shall focus on the properties of the identification scheme that estimates expansion coefficients in such series expansions, by using simple (least-squares) linear regression algorithms. To this end, we shall consider the following problem set-up, compatible with the standard framework for identification as presented in Ljung (1987).

Consider a linear, time-invariant, discrete-time data generating system

$$y(t) = G_0(q)u(t) + H_0(q)e_0(t), \tag{6}$$

where $G_0(z) = \sum_{k=0}^{\infty} g_k^{(0)} z^{-k}$ is a scalar stable transfer function (i.e. bounded for $|z| \geq 1$), $y(t)$ and $u(t)$ are scalar-valued output and input signals, q^{-1} is the delay operator, H_0 is a stable rational and monic transfer function, and e_0 is a unit-variance, zero-mean white noise process. We shall also denote $v(t) = H_0(q)e_0(t)$.

A linear model structure will be employed, determined by

$$\hat{y}(t, \theta) = \sum_{k=0}^{n-1} L_k(\theta) f_k(q) u(t), \tag{7}$$

with θ varying over an appropriate parameter space $\Theta \subset \mathbb{R}^d$.

Given data $\{u(t), y(t)\}_{t=1,\ldots,N}$ taken from experiments on this system, the corresponding one-step-ahead prediction error is given by

$$\varepsilon(t, \theta) = y(t) - \sum_{k=0}^{n-1} L_k(\theta) f_k(q) u(t). \tag{8}$$

The least-squares parameter estimate is determined by

$$\hat{\theta}_N(n) = \arg\min_{\theta \in \Theta} \frac{1}{N} \sum_{t=1}^{N} \varepsilon(t, \theta)^2, \tag{9}$$

and the corresponding estimated transfer function by

$$G(z, \hat{\theta}_N) = \sum_{k=0}^{n-1} L_k(\hat{\theta}_N) f_k(z). \tag{10}$$

This identification method has some favourable properties. Firstly, it is a linear regression scheme, which leads to a simple analytic solution; secondly, it is of the *output-error-method type,* which has the advantage that the input–output system $G_0(z)$ can be estimated consistently whenever the unknown noise disturbance $v(t)$ is uncorrelated with the input signal (Ljung, 1987).

However, it is well known that for moderately damped systems, and/or in situations of high sampling rates, it may take a large value of n, the number of coefficients to be estimated, in order to capture the essential dynamics of the system G into its model. If we are able to improve the basis functions in such a way that an accurate description of the model to be estimated can be achieved by a small number of coefficients in a series expansion then this is beneficial from both aspects of bias and variance of the model estimate.

In this paper we shall analyse bias and variance errors for the asymptotic parameter and transfer function estimates, for the general class of orthogonal basis functions recently introduced in Heuberger and Bosgra (1990), Heuberger (1991) and Heuberger *et al.* (1993, 1995). The results presented generalize corresponding results provided by Wahlberg (1991, 1994a) for the Laguerre and Kautz bases.

In Section 2 we first present the general class of orthogonal basis functions, and we formulate the least-squares identification in Section 3. Then in Sections 4 and 5 we introduce and analyse the Hambo transform of signals and systems, induced by the generalized orthogonal basis. This transform plays an important role in the statistical analysis of the asymptotic parameter estimates. This asymptotic analysis is completed in Section 6. A simulation example in Section 7 illustrates the identification method, and the paper is concluded in Section 8 with some summarizing remarks.

2. GENERALIZED ORTHONORMAL BASIS FUNCTIONS

We shall consider the generalized orthogonal basis functions that were introduced by Heuberger *et al.* (1995), based on the preliminary work of Heuberger and Bosgra (1990) and Heuberger (1991). The main result of concern is reflected in the following theorem.

Theorem 2.1. Let $G_b(z)$ be a scalar inner function with McMillan degree $n_b > 0$, having a minimal balanced realization (A, B, C, D). Denote

$$V_k(z) := z(zI - A)^{-1} B G_b^k(z). \tag{11}$$

Then the sequence of scalar rational functions $\{e_i^{\mathrm{T}} V_k(z)\}_{i=1,\ldots,n_b;k=0,\ldots,\infty}$ forms an orthonormal basis for the Hilbert space $\mathcal{H}_2$.

Note that these basis functions exhibit the property that they can incorporate system dynamics in a very general way. One can construct an inner function G_b from any given set of poles, and thus the resulting basis can incorporate dynamics of any complexity, combining, for example, both fast and slow dynamics in damped and resonant modes. A direct result is that for any specifically chosen $V_k(z)$, any strictly proper transfer function $G(z) \in \mathcal{H}_2$ has a unique series expansion

$$G(z) = z^{-1} \sum_{k=0}^{\infty} L_k V_k(z), \tag{12}$$

with $L_k \in \ell_2^{1\times n_b}[0, \infty)$.

For specific choices of $G_b(z)$, well-known classical basis functions can be generated.

(a) With $G_b(z) = z^{-1}$, having minimal balanced realization $(0, 1, 1, 0)$, the standard pulse basis $V_k(z) = z^{-k}$ results.

(b) Choosing a first order inner function $G_b(z) = (1 - az)/(z - a)$, with some real-valued a, $|a| < 1$, and balanced realization

$$(A, B, C, D) = (a, \sqrt{1-a^2}, \sqrt{1-a^2}, -a), \tag{13}$$

the Laguerre basis results:

$$V_k(z) = \sqrt{1-a^2}\, z \frac{(1-az)^k}{(z-a)^{k+1}}. \tag{14}$$

(c) Similarly, the Kautz functions (Kautz, 1954; Wahlberg, 1990, 1994a) originate from the choice of a second-order inner function

$$G_b(z) = \frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c}, \tag{15}$$

with some real-valued b and c satisfying $|c|, |b| < 1$. A balanced realization of $G_b(z)$ can be found to be given by

$$A = \begin{bmatrix} b & \sqrt{(1-b^2)} \\ c\sqrt{(1-b^2)} & -bc \end{bmatrix}, \tag{16}$$

$$B = \begin{bmatrix} 0 \\ \sqrt{(1-c^2)} \end{bmatrix}, \tag{17}$$

$$C = [\gamma_2 \quad \gamma_1], \quad D = -c, \tag{18}$$

with $\gamma_1 = -b\sqrt{(1-c^2)}$ and $\gamma_2 = \sqrt{(1-c^2)(1-b^2)}$, (see also Heuberger *et al.*, 1993, 1995).

The generalized orthonormal basis for $\mathcal{H}_2$ also induces a similar basis for the signal space $\ell_2[0, \infty)$ of squared summable sequences, through inverse z-transformation to the signal domain. Denoting

$$V_k(z) = \sum_{l=0}^{\infty} \phi_k(l) z^{-l}, \tag{19}$$

it follows that $\{e_i^{\mathrm{T}} \phi_k(l)\}_{i=1,\ldots,n_b; k=0,\ldots,\infty}$ is an orthonormal basis for the signal space $\ell_2[0, \infty)$. These ℓ_2 basis functions can also be constructed directly from G_b and its balanced realization (A, B, C, D) (see Heuberger *et al.*, 1995).

3. IDENTIFICATION OF EXPANSION COEFFICIENTS

In this section we shall consider and denote the least-squares identification method in more detail.

The prediction error that results from applying the appropriate model structure conforming to (12) can be written as

$$\varepsilon(t, \theta) = y(t) - \sum_{k=0}^{n-1} L_k V_k(q) u(t-1), \tag{20}$$

with the unknown parameter θ written as

$$\theta := [L_0 \quad \ldots \quad L_{n-1}]^{\mathrm{T}} \in \mathbb{R}^{n_b n}. \tag{21}$$

We shall assume that the input signal $\{u((t)\}$ is quasi-stationary (Ljung, 1987) with spectral density $\Phi_u(\omega)$ having a stable spectral factor $H_u(e^{i\omega})$, i.e. $\Phi_u(\omega) = H_u(e^{i\omega}) H_u(e^{-i\omega})$.

We shall further denote

$$x_k(t) := V_k(q) u(t-1), \tag{22}$$

$$\psi(t) := \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{n-1}(t) \end{bmatrix}, \tag{23}$$

and consequently

$$\varepsilon(t, \theta) = y(t) - \psi^{\mathrm{T}}(t)\theta. \tag{24}$$

Following Ljung (1987), under weak conditions, the parameter estimate $\hat{\theta}_N(n)$ given by (9) will converge with probability 1 to the asymptotic estimate

$$\theta^*(n) = R(n)^{-1} F(n), \tag{25}$$

with

$$R(n) = \bar{E}\psi(t)\psi^{\mathrm{T}}(t), \quad F(n) = \bar{E}\psi(t)y(t). \tag{26}$$

For the analysis of bias and variance errors of this identification scheme, we shall further use the notation

$$G_0(z) = z^{-1} \sum_{k=0}^{\infty} L_k^{(0)} V_k(z),$$

$$\theta_0 = [L_0^{(0)} \quad \ldots \quad L_{n-1}^{(0)}]^{\mathrm{T}},$$

$$\theta_e = [L_n^{(0)} \quad L_{n+1}^{(0)} \quad \ldots]^{\mathrm{T}},$$

$$\psi_e(t) = [x_n^{\mathrm{T}}(t) \quad x_{n+1}^{\mathrm{T}}(t) \quad \ldots]^{\mathrm{T}},$$

$$\Omega_n(e^{i\omega}) := e^{-i\omega}[V_0^{\mathrm{T}}(e^{i\omega}) \quad V_1^{\mathrm{T}}(e^{i\omega}) \quad \ldots \quad V_{n-1}^{\mathrm{T}}(e^{i\omega})]^{\mathrm{T}},$$

$$\Omega_e(e^{i\omega}) := e^{-i\omega}[V_n^{\mathrm{T}}(e^{i\omega}) \quad V_{n+1}^{\mathrm{T}}(e^{i\omega}) \quad \ldots]^{\mathrm{T}},$$

leading to the following alternative description of the data-generating system:

$$y(t) = \psi^{\mathrm{T}}(t)\theta_0 + \psi_e^{\mathrm{T}}(t)\theta_e + v(t), \qquad (27)$$

$$G_0(e^{i\omega}) = \Omega_n^{\mathrm{T}}(e^{i\omega})\theta_0 + \Omega_e^{\mathrm{T}}(e^{i\omega})\theta_e. \qquad (28)$$

In the analysis of the asymptotic parameter estimate $\theta^*(n)$ the block-Toeplitz structured matrix $R(n)$ will play an important role. Note that this matrix has a block-Toeplitz structure with the (j, ℓ) block-element given by $\bar{E}x_j(t)x_\ell^{\mathrm{T}}(t)$. For the analysis of the properties of this block-Toeplitz matrix, we shall employ a signal and system transformation that is induced by the generalized basis. This transformation is presented and discussed in the next two sections.

Remark 3.1. In this paper we consider the identification of strictly proper systems by strictly proper models. All results will also hold true for the case of proper systems and corresponding proper models, by simply adapting the notation in this section.

4. THE HAMBO TRANSFORM OF SIGNALS AND SYSTEMS

The presented generalized orthonormal basis for $\mathcal{H}_2/\ell_2$ induces a transformation of signals and systems to a transform domain. Besides the intrinsic importance of signal and systems analysis in this transform domain (for some of these results see Heuberger, 1991), we can fruitfully use these transformations in the analysis of statistical properties of the identified models, as well as in the derivation of bias and variance error bounds.

Let $\{V_k(z)\}_{k=0,\dots,\infty}$ be an orthonormal basis, as defined in Section 2, and let $\{\phi_k(t)\}_{k=0,\dots,\infty}$ be as defined in (19). Then, for any signal $x(t) \in \ell_2^m$, there exists a unique transformation

$$\mathcal{X}(k) := \sum_{t=0}^{\infty} \phi_k(t)x^{\mathrm{T}}(t), \qquad (29)$$

and we denote the corresponding λ transform of $\mathcal{X}(k)$ by

$$\tilde{x}(\lambda) := \sum_{k=0}^{\infty} \mathcal{X}(k)\lambda^{-k}. \qquad (30)$$

We shall refer to $\tilde{x}(\lambda)$ as the *Hambo* transform of the signal $x(t)$. Note that $x \in \ell_2^m$ and $\tilde{x}(\lambda) \in \mathcal{H}_2^{n_b \times m}$.

Now consider a scalar system $y(t) = G(q)u(t)$, with $G \in \mathcal{H}_2$, and with u and y signals in ℓ_2. Then there exists a Hambo-transformed system $\tilde{G}(\lambda) \in \mathcal{H}_2^{n_b \times n_b}$ such that

$$\tilde{y}(\lambda) = \tilde{G}(\lambda)\tilde{u}(\lambda). \qquad (31)$$

In terms of the sequence of expansion coefficients, this can also be written as

$$\mathcal{Y}(k) = \tilde{G}(q)\mathcal{U}(k) \quad \text{for all } k, \qquad (32)$$

where the shift operator q operates on the sequence index k. The construction of this transformed system is given in the following proposition.

Proposition 4.1. Consider a scalar system $G \in \mathcal{H}_2$ relating input and output signals according to $y(t) = G(q)u(t)$, with $u, y \in \ell_2$, where $G(z) = \sum_{k=0}^{\infty} g_k z^{-k}$. Consider an orthonormal basis $\{V_k(z)\}_{k=0,\dots,\infty}$ as defined in Section 2, generated by a scalar inner transfer function $G_b(z)$ with balanced realization (A, B, C, D). Denote the rational function

$$N(\lambda) := A + B(\lambda - D)^{-1}C. \qquad (33)$$

Then

$$\tilde{y}(\lambda) = \tilde{G}(\lambda)\tilde{u}(\lambda) \qquad (34)$$

with

$$\tilde{G}(\lambda) := \sum_{k=0}^{\infty} g_k N(\lambda)^k. \qquad (35)$$

Proof. See the Appendix.

The interpretation of this proposition is that the Hambo transform of any system G can be obtained by a simple variable transformation $z^{-1} = N(\lambda)$ on the original transfer function.

Note that this result generalizes the situation of a corresponding Laguerre transformation, where it concerns the variable transformation $z = (\lambda + a)/(1 + a\lambda)$ (see also Wahlberg, 1991). However, because in our case the McMillan degree of the inner function that generates the basis is $n_b \geq 1$, the Hambo-transformed system $\tilde{G}$ increases in input–output dimension to $\tilde{G} \in \mathcal{H}_2^{n_b \times n_b}$. Note that, since G is scalar, $N(\lambda)$ is an $n_b \times n_b$ rational transfer function matrix of McMillan degree 1 (since D is scalar). Note also the appealing symmetric structure of this result. Whereas G_b has a balanced realization (A, B, C, D), the variable-transformation function $N(\lambda)$ has a realization (D, C, B, A), which can also be shown to be balanced.

Proposition 4.1 considers scalar ℓ_2 signals and scalar systems. Later in this paper we shall also have to deal with specific situations of multivariable signals, for which there exist straightforward extensions of this result.

Proposition 4.2. Consider a scalar transfer function $G \in \mathcal{H}_2$ relating m-dimensional input and output signals $u, y \in \ell_2^m$, according to†

$$y(t) = [G(z)I_m]u(t). \quad (36)$$

Then

$$\tilde{y}(\lambda) = \tilde{G}(\lambda)\tilde{u}(\lambda), \quad (37)$$

with $\tilde{G}(\lambda)$ as defined in (35).

Proof. For $m = 1$ the result is shown in Proposition 4.1. If we write the relation between y and u componentwise, i.e. $y_i(t) = G(z)u_i(t)$, it follows from Proposition 4.1 that $\tilde{y}_i(\lambda) = \tilde{G}(\lambda)\tilde{u}_i(\lambda)$, where $\tilde{y}_i, \tilde{u}_i \in \mathcal{H}_2^{n_b \times 1}(\lambda)$. It follows directly that

$$\begin{aligned}\tilde{y}(\lambda) &= [\tilde{y}_1(\lambda) \ \ldots \ \tilde{y}_m(\lambda)] \\ &= \tilde{G}(\lambda)[\tilde{u}_1(\lambda) \ \ldots \ \tilde{u}_m(\lambda)] \\ &= \tilde{G}(\lambda)\tilde{u}(\lambda).\end{aligned}$$

□

One of the results that we shall need in the analysis of least-squares related block Toeplitz matrices is formulated in the following proposition.

Proposition 4.3. Consider a scalar inner transfer function $G_b(z)$ generating an orthogonal basis as discussed before. Then

$$\tilde{G}_b(\lambda) = \lambda^{-1}I_{n_b}. \quad (38)$$

Proof. It can be simply verified that for all k, $G_b(q)\phi_k(t) = \phi_{k+1}(t)$. With Proposition 4.2, it follows that $\tilde{\phi}_{k+1}(\lambda) = \tilde{G}_b(\lambda)\tilde{\phi}_k(\lambda)$. Since for each k, $\tilde{\phi}_k(\lambda) = \sum_{t=0}^{\infty} I_{n_b}\delta(t-k)\lambda^{-t}$, it follows that for all k,

$$I_{n_b}\lambda^{-k-1} = \tilde{G}_b(\lambda)I_{n_b}\lambda^{-k}. \quad (39)$$

Since this holds for all k, it proves the result. □

The basis-generating inner function transforms to a simple shift in the Hambo domain. Next we shall consider a result that reflects some properties of the back-transformation of Hambo-transformed systems to the original domain, i.e. the inverse Hambo transform.

Proposition 4.4. Consider a scalar inner transfer function G_b generating an orthogonal basis as discussed before, and let G_b induce a corresponding Hambo transform. Let $H \in \mathcal{H}_2$ be a scalar transfer function with Hambo transform $\tilde{H}$. Then

$$\tilde{H}(G_b(z))V_0(z) = V_0(z)H(z^{-1}). \quad (40)$$

Proof. See the Appendix.

† Since $G(z)$ is scalar, we allow the notation $G(z)I_m$, which more formally should be written as $G(z) \otimes I_m$.

The following lemma relates quadratic signal properties of the transformed signals.

Lemma 4.1. Let $x_j, x_l \in \ell_2^m$ and consider a Hambo transform induced by an orthonormal basis $V_k(z)$ generated by an inner function $G_b(z)$ with McMillan degree $n_b \geq 1$. Then

$$\begin{aligned}\sum_{t=0}^{\infty} x_j(t)x_l^{\mathrm{T}}(t) &= \sum_{k=0}^{\infty} \mathcal{X}_j^{\mathrm{T}}(k)\mathcal{X}_l(k) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{x}_j^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{x}_l(\mathrm{e}^{\mathrm{i}\omega})\,\mathrm{d}\omega. \quad (41)\end{aligned}$$

Proof. This lemma is a direct consequence of the fact that, because the basis is orthonormal, it induces a transformation that is an isomorphism. □

The transformation that is discussed in this section refers to ℓ_2 signals, and the corresponding transformation of systems actually concerns the transformation of the ℓ_2 behaviour or graph of a dynamical system. However, this same orthogonal basis for ℓ_2 can also be employed to induce a transformation of (quasi-)stationary stochastic processes to the transform domain, as briefly considered in the next section.

5. HAMBO TRANSFORMATION OF STOCHASTIC PROCESSES

Let v be a scalar-valued stochastic process or quasi-stationary signal (Ljung, 1987), having a rational spectral density $\Phi_v(\omega)$. let $H_v(\mathrm{e}^{\mathrm{i}\omega})$ be a stable spectral factor of $\Phi_v(\omega)$, and let $h_v(k)$ be its ℓ_2 impulse response, satisfying $H_v(z) = \sum_{k=0}^{\infty} h_v(k)z^{-k}$. Then

$$h_v(t) = H_v(q)\delta(t), \quad (42)$$

and consequently, with Proposition 4.1,

$$\tilde{h}_v = \tilde{H}_v\tilde{\delta}. \quad (43)$$

The Hambo transform of the spectral density $\Phi_v(\omega)$ will be defined as

$$\tilde{\Phi}_v(\omega) := \tilde{H}_v^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{H}_v(\mathrm{e}^{\mathrm{i}\omega}). \quad (44)$$

Let $w(t) = P_{wv}(q)v(t)$, with P_{wv} a stable scalar transfer function. Then

$$\tilde{h}_w = \tilde{P}_{wv}\tilde{h}_v, \quad (45)$$

with $h_w, h_v \in \ell_2$, the impulse responses of stable spectral factors of $\Phi_w(\omega), \Phi_v(\omega)$, respectively. Similarly to Lemma 4.1 we can now formulate some properties of stochastic processes.

Lemma 5.1. Let w and z be m-dimensional stationary stochastic processes, satisfying $w(t) = \sum_{k=0}^{\infty} h_w(k)e(t-k)$ and $z(t) = \sum_{k=0}^{\infty} h_z(k)e(t-k)$, with $\{e(t)\}$ a scalar-valued unit-variance white noise process. Then

$$\bar{E}[w(t)z^{\mathrm{T}}(t)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{h}_w^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{h}_z(\mathrm{e}^{\mathrm{i}\omega})\,\mathrm{d}\omega. \quad (46)$$

Lemma 5.2. Let w, z and v be m-dimensional stationary stochastic processes, satisfying

$$w(t) = P_{wv}(q)I_m v(t), \quad (47)$$

$$z(t) = P_{zv}(q)I_m v(t), \quad (48)$$

where $P_{wv}, P_{zv} \in \mathcal{H}_2$ and $v(t) = \sum_{k=0}^{\infty} h_v(k)e(t-k)$, with $\{e(t)\}$ a scalar-valued unit-variance white noise process. Then

$$\bar{E}w(t)z^{\mathrm{T}}(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{h}_v^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{P}_{wv}^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{P}_{wv}(\mathrm{e}^{\mathrm{i}\omega})\tilde{h}_v(\mathrm{e}^{\mathrm{i}\omega})\,\mathrm{d}\omega.$$

The previous lemmas can be simply shown to hold also in the case of quasi-stationary signals. To this end, we have already used the operator $\bar{E} := \lim_{N\to\infty}(1/N)\sum_{t=0}^{N-1} E$, where E stands for expectation.

6. ASYMPTOTIC ANALYSIS OF BIAS AND VARIANCE ERRORS

6.1. *Analysis of the least squares problem*

As mentioned in Section 3, the main properties of the asymptotic parameter estimate hinge on the characteristics of the matrix $R(n)$, as defined by (26) and (23). We shall first analyse this matrix in terms of its structural properties and its eigenvalues. The Hambo transform is very efficiently employed in the following result.

Theorem 6.1. The matrix $R(n)$ defined in (26) is a block-Toeplitz matrix, being the covariance matrix related to the spectral density function $\tilde{\Phi}_u(\omega)$.

Proof. The (j, ℓ) block-element of the matrix $R(n)$ is given by $\bar{E}x_j(t)x_l^{\mathrm{T}}(t)$. Since $x_j(t) = G_b^j(q)I_{n_b}x_0(t)$, it follows with Lemma 5.2 that

$$\bar{E}x_j(t)x_l^{\mathrm{T}}(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{h}_{x_0}^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega}) \times [\tilde{G}_b^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})]^j[\tilde{G}_b(\mathrm{e}^{\mathrm{i}\omega})]^l\tilde{h}_{x_0}(\mathrm{e}^{\mathrm{i}\omega})\,\mathrm{d}\omega.$$

With Proposition 4.3, it follows that

$$\bar{E}x_j(t)x_l^{\mathrm{T}}(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\omega(j-l)}\tilde{h}_{x_0}^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{h}_{x_0}(\mathrm{e}^{\mathrm{i}\omega})\,\mathrm{d}\omega. \quad (49)$$

Since $x_0(t) = q^{-1}V_0(q)u(t)$, we can write $h_{x_0}(t) = q^{-1}V_0(q)H_u(q)\delta(t)$. Since H_u is scalar, we can write $h_{x_0}(t) = H_u(q)q^{-1}V_0(q)\delta(t) = H_u(q)h_{v_0}(t)$, with $h_{v_0}(t)$ the impulse response of the transfer function $q^{-1}V_0(q)$.

Applying Propositon 4.1 now shows that $\tilde{h}_{x_0} = \tilde{H}_u\tilde{h}_{v_0} = \tilde{H}_u$. The latter equality follows from $\tilde{h}_{v_0} = I_{n_b}$, since the impulse response of $q^{-1}V_0(q)$ exactly matches the first n_b basis functions in the Hambo domain. Consequently,

$$\bar{E}x_j(t)x_l^{\mathrm{T}}(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\omega(j-l)}\tilde{H}_u^{\mathrm{T}}(\mathrm{e}^{-\mathrm{i}\omega})\tilde{H}_u(\mathrm{e}^{\mathrm{i}\omega})\,\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\omega(j-l)}\tilde{\Phi}_u(\omega)\,\mathrm{d}\omega$$

which proves the result. □

Remark 6.1. In Wahlberg (1991), treating the (first-order) Laguerre case, the corresponding Toeplitz matrix is the covariance matrix related to the spectral density

$$\Phi_u\left(\frac{\mathrm{e}^{\mathrm{i}\omega} + a}{1 + a\mathrm{e}^{\mathrm{i}\omega}}\right). \quad (50)$$

This implies that in this case a variable transformation

$$\mathrm{e}^{\mathrm{i}\omega} \to \frac{\mathrm{e}^{\mathrm{i}\omega} + a}{1 + a\mathrm{e}^{\mathrm{i}\omega}} \quad (51)$$

is involved, or equivalently

$$\mathrm{e}^{-\mathrm{i}\omega} \to \frac{1 + a\mathrm{e}^{\mathrm{i}\omega}}{\mathrm{e}^{\mathrm{i}\omega} + a}. \quad (52)$$

Using the balanced realization of a first-order inner function, (13), this implies that, in the setting of this paper, the variable transformation involved is given by $\mathrm{e}^{-\mathrm{i}\omega} \to N(\mathrm{e}^{\mathrm{i}\omega})$, while N has a minimal balanced realization

$$(-a, \sqrt{1-a^2}, \sqrt{1-a^2}, a).$$

This leads directly to the variable transformation (52).

The following proposition bounds the eigenvalues of the block-Toeplitz matrix $R(n)$.

Proposition 6.1. Let the block-Toeplitz matrix $R(n)$ defined in (26) have eigenvalues $\lambda_j(R(n))$. Then

(a) for all n, the eigenvalues of $R(n)$ are bounded by

$$\operatorname*{ess\,inf}_{\omega} \Phi_u(\omega) \le \lambda_j(R(n)) \le \operatorname*{ess\,sup}_{\omega} \Phi_u(\omega);$$

(b) $\lim_{n\to\infty} \max_j \lambda_j(R(n)) = \operatorname{ess\,sup}_\omega \Phi_u(\omega)$.

Proof. See the Appendix.

In the rest of this paper we shall assume that the input spectrum is bounded away from zero, i.e. $\operatorname{ess\,inf}_\omega \Phi_u(\omega) \geq c > 0$.

6.2. *Asymptotic bias error*

The previous results on $R(n)$ can be employed in the derivation of upper bounds for the asymptotic bias errors both in estimated parameters and in the resulting transfer function estimate. Combining (25)–(27), it follows that

$$\theta^* - \theta_0 = R(n)^{-1}\bar{E}[\psi(t)\psi_e^{\mathrm{T}}(t)\theta_e]. \tag{53}$$

Consequently,

$$\|\theta^* - \theta_0\|_2 \leq \|R(n)^{-1}\|_2 \, \|\bar{E}[\psi(t)\psi_e^{\mathrm{T}}(t)]\|_2 \, \|\theta_e\|_2, \tag{54}$$

where, for a (matrix) operator T, $\|T\|_2$ refers to the induced operator 2-norm. For simplicity of notation, we have left out the dependence of θ^* (and θ_0) on n.

We can now formulate the following upper bound on the bias error.

Proposition 6.2. Consider the identification set-up discussed in Section 3. Then

$$\|\theta^* - \theta_0\|_2 \leq \frac{\operatorname{ess\,sup}_\omega \Phi_u(\omega)}{\operatorname{ess\,inf}_\omega \Phi_u(\omega)} \|\theta_e\|_2, \tag{55}$$

where $\|\theta_e\|_2 = \sqrt{\sum_{k=n}^{\infty} L_k^{(0)} L_k^{(0)^{\mathrm{T}}}}$.

Proof. See the Appendix.

In this result, as in the rest of this paper, we have employed an upper bound for $\|R(n)^{-1}\|_2$ as provided by Proposition 6.1. In many situations the input signal and its statistical properties will be known, and $\|R(n)^{-1}\|_2$ can be calculated exactly. In that case we can replace $(\operatorname{ess\,inf}_\omega \Phi_u(\omega))^{-1}$ in (55) by $\|R(n)^{-1}\|_2$.

For the bias in the transfer function estimate, the result corresponding to Proposition 7.13 is as follows.

Proposition 6.3. Consider the identification set-up as discussed in Section 3. Then

(a) for all $\omega_1 \in [-\pi, \pi]$,

$$\begin{aligned} &|G(e^{i\omega_1}, \theta^*) - G_0(e^{i\omega_1})| \\ &\leq \|V_0(e^{i\omega_1})\|_\infty [\|\theta_0 - \theta^*\|_1 + \|\theta_e\|_1] \\ &\leq \|V_0(e^{i\omega_1})\|_\infty \\ &\quad \times \left[\sqrt{n_b n}\frac{\operatorname{ess\,sup}_\omega \Phi_u(\omega)}{\operatorname{ess\,inf}_\omega \Phi_u(\omega)} \|\theta_e\|_2 + \|\theta_e\|_1\right], \end{aligned}$$

where $\|V_0(e^{i\omega_1})\|_\infty$ is the ℓ_∞-induced operator norm of the matrix $V_0(e^{i\omega_1}) \in \mathbb{C}^{n_b \times 1}$, i.e. the maximum absolute value over the elements in $V_0(e^{i\omega_1})$;

(b) the $\mathcal{H}_2$ norm of the model error is bounded by

$$\|G(z, \theta^*) - G_0(z)\|_{\mathcal{H}_2} \leq \sqrt{\|\theta_0 - \theta^*\|_2^2 + \|\theta_e\|_2^2}, \tag{56}$$

$$\leq \left[1 + \frac{\operatorname{ess\,sup}_\omega \Phi_u(\omega)}{\operatorname{ess\,inf}_\omega \Phi_u(\omega)}\right] \|\theta_e\|_2. \tag{57}$$

Proof. See the Appendix.

Note that this latter bound on the bias in the transfer function estimate as well as the previously derived bound are dependent on the basis functions chosen. The factor $\|\theta_e\|_2$ is determined by the convergence rate of the series expansion of G_0 in the generalized basis. The closer the dynamics of the system G_0 is to the dynamics of the inner transfer function G_b, the faster is the convergence rate. An upper bound for this convergence rate is derived in the following proposition, based on the results of Heuberger *et al.* (1995).

Proposition 6.4. Let $G_0(z)$ have poles $\mu_i, i = 1, \ldots, n_s$, and let $G_b(z)$ have poles $\rho_j, j = 1, \ldots, n_b$. Denote

$$\lambda := \max_i \prod_{j=1}^{n_b} \left|\frac{\mu_i - \rho_j}{1 - \mu_i \rho_j}\right|. \tag{58}$$

Then there exists a constant $c \in \mathbb{R}$ such that, for all $\eta > \lambda$,

$$\|\theta_e\|_2 \leq c\frac{\eta^{n+1}}{\sqrt{1 - \eta^2}}. \tag{59}$$

Proof. A sketch of the proof is given in the Appendix. For a detailed proof, see Heuberger (1991) and Heuberger and Van den Hof (1995).

Note that when the two sets of poles converge to each other, λ will tend to 0, the upper bound on $\|\theta_e\|_2$ will decrease drastically, and the bias error will be reduced accordingly. The above result clearly shows the important contribution that an appropriately chosen set of basis functions can have in achieving a reduction of the bias in estimated transfer functions.

The results (55) and (57) show that we achieve consistency of the parameter and transfer function estimates as $n \to \infty$ provided that the input spectrum is bounded away from 0 and $\|\theta_e\|_2 \to 0$ for $n \to \infty$. The latter condition is guaranteed if $G_0 \in \mathcal{H}_2$.

For the FIR case, corresponding to $G_b(z) = z^{-1}$, we know that under specific experimental conditions the finite number of expansion coefficients can also be estimated consistently, irrespective of the tail. This situation can also be formulated for the generalized case.

Corollary 6.1. Consider the identification setup as discussed in Section 3. If $\tilde{H}_u$ is an inner transfer function then it follows that $\theta^* = \theta_0$, and the transfer function error bounds become

(a) $|G(e^{i\omega_1}, \theta^*) - G_0(e^{i\omega_1})| \leq \|V_0(e^{i\omega_1})\|_\infty \|\theta_e\|_1$ for each ω_1;

(b) $\|G(z, \theta^*) - G_0(z)\|_{\mathcal{H}_2} \leq \|\theta_e\|_2$.

Proof. Under the given condition, it can be simply verified that $\bar{\Phi}_u(\omega) = I_{n_b}$. This implies that the block-Toeplitz matrix $R(n) = I$, and that for all $n \geq 1$, $R_{12}(n) = 0$. Employing this relation in the proofs of Propositions 6.2 and 6.3 gives the results. □

Note that a special case of the situation of an inner $\tilde{H}_u$ is obtained if the input signal u is uncorrelated (white noise). Then $H_u = 1$, and consequently $\tilde{H}_u = I_{n_b}$, being inner.

6.3. *Asymptotic variance*

For an analysis of the asymptotic variance of the estimated transfer function, we can generalize the results obtained for the case of Laguerre functions by Wahlberg (1991). From classical analysis of prediction error identification methods (Ljung, 1987), we know that, under fairly weak conditions,

$$\sqrt{N}(\hat{\theta}_N(n) - \theta^*) \to \mathcal{N}(0, Q_n) \quad \text{as } N \to \infty, \tag{60}$$

where $\mathcal{N}(0, Q_n)$ denotes a Gaussian distribution with zero mean and covariance matrix Q_n. For output error identification schemes, as applied in this paper, the asymptotic covariance matrix satisfies

$$Q_n = [\bar{E}\psi(t)\psi^T(t)]^{-1}[\bar{E}\bar{\psi}(t)\bar{\psi}^T(t)][\bar{E}\psi(t)\psi^T(t)]^{-1}, \tag{61}$$

where $\bar{\psi}(t) = \sum_{i=0}^{\infty} h_0(i)\psi(t+i)$ and $h_0(i)$ is the impulse response of the corresponding transfer function H_0.

We know that, according to Theorem 6.1, the block-Toeplitz matrix $R(n) = \bar{E}\psi(t)\psi^T(t)$ is related to the spectral density function $\bar{\Phi}_u(\omega)$. For the block-Toeplitz matrix $P(n) = \bar{E}\bar{\psi}(t)\bar{\psi}^T(t)$, we can formulate a similar result.

Lemma 6.1. The block-Toeplitz matrix $P(n) = \bar{E}\bar{\psi}(t)\bar{\psi}^T(t)$ is the covariance matrix related to the spectral density function $\bar{\Phi}_u(\omega)\Phi_v(\omega)$.

Proof. The proof follows similar lines to that of Theorem 6.1. □

From the asymptotic covariance of the parameter estimate, we can derive an expression for the transfer function estimate:

$$\frac{N}{n_b n}\operatorname{cov}(\hat{G}(e^{i\omega_1}), \hat{G}(e^{i\omega_2})) \to \frac{1}{n_b n}\Omega_n^T(e^{i\omega_1})Q_n\Omega_n(e^{-i\omega_2}) \tag{62}$$

as $N \to \infty$, where $\operatorname{cov}(\cdot,\cdot)$ is the cross-covariance matrix in the joint asymptotic distribution of

$$[G(e^{i\omega_1}, \hat{\theta}_N) - G(e^{i\omega_1}, \theta^*), \; G(e^{i\omega_2}, \hat{\theta}_N) - G(e^{i\omega_2}, \theta^*)].$$

In (62) the term $n_b n$ is the number of scalar parameters that is estimated.

We now have the ingredients for formulating an expression for the asymptotic covariance $(n \to \infty, N \to \infty)$ of the estimated transfer functions.

Theorem 6.2. Assume the spectral density $\Phi_u(\omega)$ to be bounded away from zero and sufficiently smooth. The for $N, n \to \infty$, $n^2/N \to 0$,

$$\frac{N}{n_b n}\operatorname{cov}(G(e^{i\omega_1}, \hat{\theta}_N), G(e^{i\omega_2}, \hat{\theta}_N)) \to \begin{cases} 0 & \text{for } G_b(e^{i\omega_1}) \neq G_b(e^{i\omega_2}), \\ V_0^T(e^{i\omega_1})V_0(e^{-i\omega_1})\dfrac{\Phi_v(\omega_1)}{\Phi_u(\omega_1)} & \\ & \text{for } \omega_1 = \omega_2. \end{cases}$$

Proof. See the Appendix.

Theorem 6.2 gives a closed-form expression for the asymptotic covariance. Note that it implies that the variance of the transfer function estimate for a specific ω_1 is given by

$$\frac{n}{N}V_0^T(e^{i\omega_1})V_0(e^{-i\omega_1})\frac{\Phi_v(\omega_1)}{\Phi_u(\omega_1)}, \tag{63}$$

which is the noise to input signal ratio weighted with an additional weighting factor that is determined by the basis functions. This additional weighting, which is not present in the case of FIR estimation, again generalizes the weighting that is also present in the case of Laguerre basis functions (see Wahlberg, 1991). Since the frequency function $V_0(e^{i\omega})$ has a low-pass character, it ensures that the variance will have a roll-off at high frequencies. This is unlike the case of FIR estimation, where the absolute variance generally increases with increasing frequency.

The role of V_0 in this variance expression clearly shows that there is a design variable involved that can be chosen also from a point of view of variance reduction. In that case V_0 has to be chosen in such a way that it reduces the effect of the noise ($\Phi_v(\omega)$) in those frequency regions where the noise is dominating.

The result of the theorem also shows that for $n_b = 1$ the transfer function estimates will be asymptotically uncorrelated. In that case it can simply be shown that $G_b(e^{i\omega_1}) \neq G_b(e^{i\omega_2})$ implies $\omega_1 \neq \omega_2$ for $\omega_1, \omega_2 \in [0, \pi]$. In the case $n_b > 1$ this latter situation is not guaranteed.

7. SIMULATION EXAMPLE

In order to illustrate the identification method considered in this paper, we shall give results for an example where an identification is performed on the basis of simulation data. The simulated system is determined by

$$G_0(z) = \frac{b_1 z^{-1} + \ldots + b_5 z^{-5}}{1 + a_1 z^{-1} + \ldots + a_5 z^{-5}}, \tag{64}$$

with $b_1 = 0.2530$, $b_2 = -0.9724$, $b_3 = 1.4283$, $b_4 = -0.9493$, $b_5 = 0.2410$; $a_1 = -4.15$, $a_2 = 6.8831$, $a_3 = -5.6871$, $a_4 = 2.3333$ and $a_5 = -0.3787$. The system has poles at $0.95 \pm 0.2i$, $0.85 \pm 0.09i$ and 0.55. The static gain of the system is 1.

An ouptut noise is added to the simulated output, coloured by a second-order noise filter

$$H_0(z) = \frac{1 - 1.38z^{-1} + 0.4z^{-2}}{1 - 1.9z^{-1} + 0.91z^{-2}}. \tag{65}$$

As input signal a zero-mean unit-variance white noise signal is chosen, while the input to the noise filter is also a white noise signal with variance 0.0025, leading to a signal-to-noise ratio at the output of 11.6 dB, which is equivalent to around 30% noise disturbance in amplitude on the noise-free output signal.

Orthogonal basis functions have been chosen generated by a fourth-order inner function having poles at $0.9 \pm 0.3i$ and $0.7 \pm 0.2i$. As the choice of basis functions can highly influence the accuracy of the identified model, we have chosen basis functions with poles that are only slightly in the direction of the system dynamics. They are not really accurate representations of the system poles, thus avoiding the use of knowledge that is generally not available in an identification situation.

We have used a data set of input and output signals with length $N = 1200$, and have estimated five coefficients of the series expansion.

The output signal and its noise contribution are depicted in Fig. 1. The Bode amplitude plot of G_0 is sketched in Fig. 2, together with the amplitude plots of each of the four components of V_0, i.e. the first four basis functions. Note that all other basis functions will show the same Bode amplitude plot, since they differ only in multiplication by a scalar inner function, which does not change its amplitude. We have used five different realizations of 1200 data points to estimate five different models. Their Bode amplitude plots are given in Fig. 3 and the corresponding step responses in Fig. 4.

Figure 5 shows the relevant expressions in the asymptotic variance expression (63). This refers to the plots of $V_0^T V_0$ and of the noise spectrum $\Phi_v(\omega)$ as well as their product. Since $\Phi_u(\omega) = 1$, this latter product determines the asymptotic variance of the estimated transfer function.

To illustrate the power of the identification

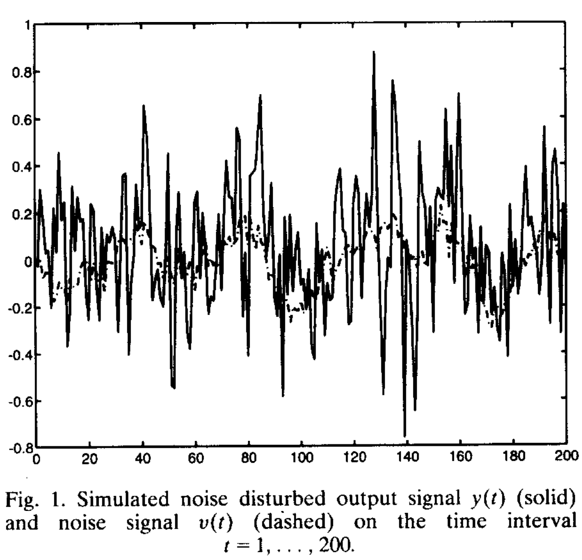

Fig. 1. Simulated noise disturbed output signal $y(t)$ (solid) and noise signal $v(t)$ (dashed) on the time interval $t = 1, \ldots, 200$.

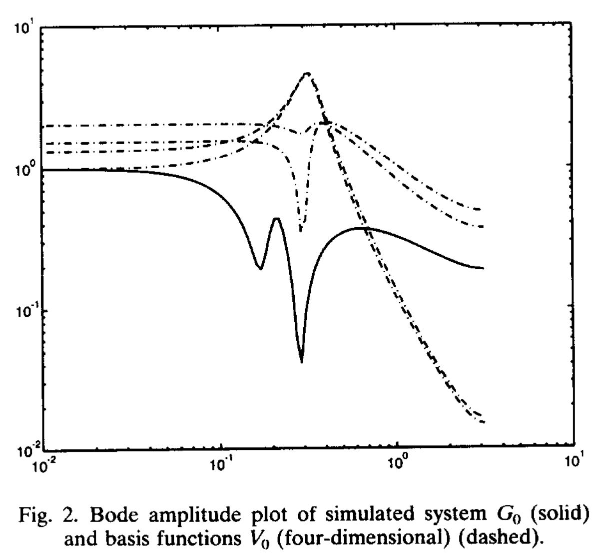

Fig. 2. Bode amplitude plot of simulated system G_0 (solid) and basis functions V_0 (four-dimensional) (dashed).

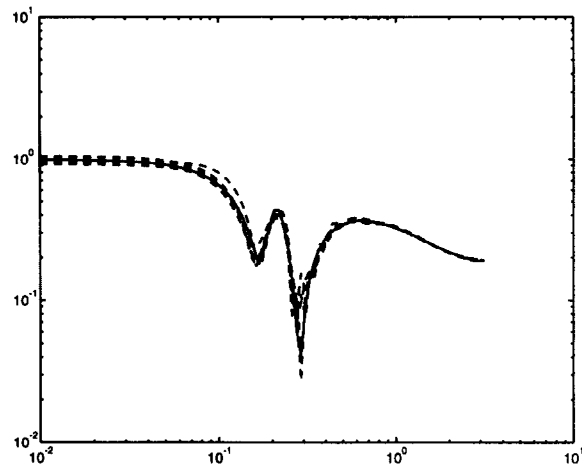

Fig. 3. Bode amplitude plot of simulated system G_0 (solid) and five estimated models with $n = 5$ and $N = 1200$, using five different realizations of the input–output data (dashed).

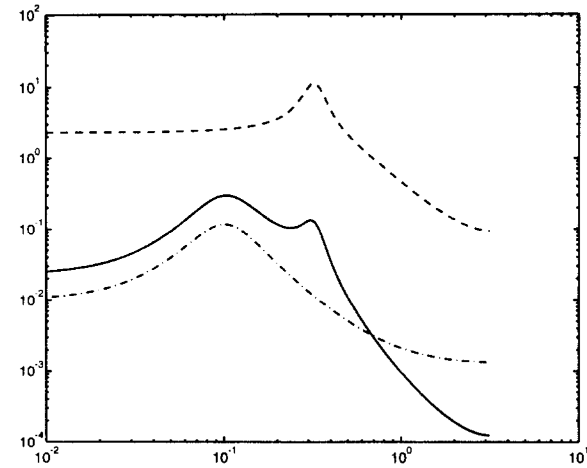

Fig. 5. Bode amplitude plot of $n_b^{-1} V_0^T V_0$ (dashed), spectrum $\Phi_v(\omega)$ (dashed-dotted) and their product (solid).

method, we have made a comparison with the identification of fifth-order (least-squares) output error models, dealing with a parametrized prediction error

$$\varepsilon(t, \theta) = y(t) - \frac{b_1 q^{-1} + b_2 q^{-2} + \ldots + b_5 q^{-5}}{1 + a_1 q^{-1} + \ldots + a_5 q^{-5}} u(t). \tag{66}$$

Note that G_0 is also a fifth-order system.

The results of the estimated fifth-order output error models are sketched in Figs 6 and 7. Here also, five different realizations of the input–output data are used. It can be observed that the models based on the generalized orthogonal basis functions have a good ability to identify the resonant behaviour of the system in the frequency range from 0.15 to 0.5 rad s^{-1}, while the output error models clearly have poorer performance here. The variance of both types of identification methods seem to be comparable. Note that the OE algorithm requires nonlinear optimization, whereas the identification of expansion coefficients is a convex optimization problem. Providing the output error algorithm with an initial parameter estimate based on the poles of the basis functions did not influence the obtained results essentially.

8. DISCUSSION

In this paper we have analysed some asymptotic properties of linear estimation schemes that identify a finite number of expansion coefficients in a series expansion of a linear stable transfer function, employing recently developed generalized orthogonal basis functions. The basis functions generalize the well-known pulse, Laguerre and Kautz basis functions, and are shown to provide flexible design variables that, when properly chosen, provide fast convergence of the series expansion. In an identification context, this implies that only a few coefficients have to be estimated to obtain

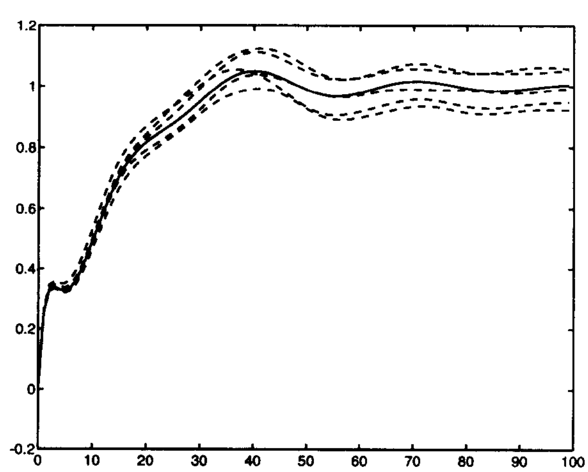

Fig. 4. Step response of simulated system G_0 (solid) and five estimated models with $n = 5$ and $N = 1200$, using five different realizations of the input–output data (dashed).

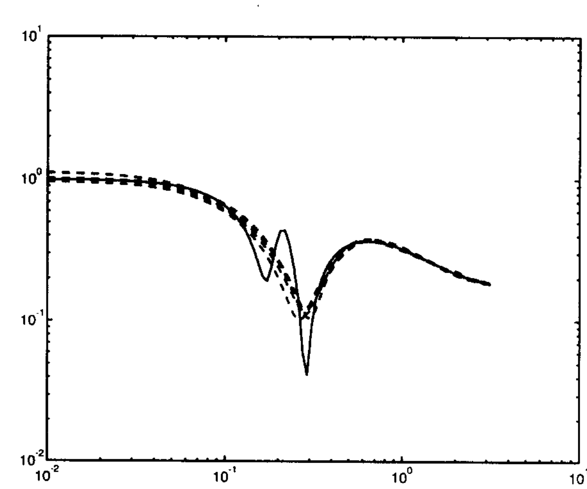

Fig. 6. Bode amplitude plot of G_0 (solid) and five fifth-order OE models (dashed).

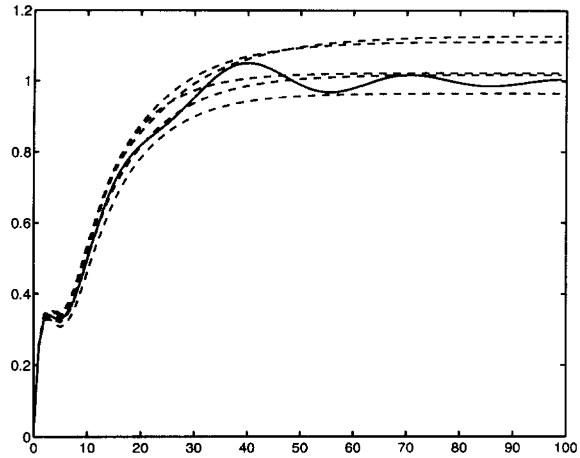

Fig. 7. Step response of G_0 (solid) and five fifth-order OE models (dashed).

accurate estimates, while simple linear regression schemes can be used. Both bias and variance errors are analysed and error bounds established.

As the accuracy of the chosen basis functions can substantially improve the identification results in both bias and variance, the introduced method points to the use of iterative procedures, where the basis functions are updated iteratively. Previously estimated models can then be used to dictate the poles of the basis. Such iterative methods have already been applied successfully in practical experiments (see e.g. De Callafon *et al.*, 1993). The flexibility of the introduced basis functions provides a means to introduce uncertain *a priori* knowledge into the identification procedure. In contrast with other identification techniques, where *a priori* knowledge definitely has to be certain, this *a priori* knowledge (i.e. the system poles) is allowed to be approximative. The only consequence is that the accuracy of the identified models will be higher, the 'better' the *a priori* knowledge is. Apart from the identification of nominal models, the basis functions introduced here have also been applied in the identification of model error bounds (see De Vries, 1994; Hakvoort and Van den Hof, 1994).

Acknowledgement—The authors would like to thank Douwe de Vries, Bo Wahlberg and Brett Ninness for interesting discussions and for their contributions to the results presented in this paper. The authors also acknowledge the Dutch Systems and Control Theory Network for its financial support.

REFERENCES

Clowes, G. J. (1965). Choice of the time scaling factor for linear system approximations using orthonormal Laguerre functions. *IEEE Trans. Autom. Control*, **AC-10**, 487–489.

De Callafon, R. A., P. M. J. Van den Hof and M. Steinbuch (1993). Control relevant identification of a compact disc pick-up mechanism. In *Proc. 32nd IEEE Conf. on Decision and Control*, San Antonio, TX, pp. 2050–2055.

De Vries, D. K. (1994). Identification of model uncertainty for control design. Doctoral dissertation, Mechanical Engineering Systems and Control Group, Delft University of Technology.

Fu, Y. and G. A. Dumont (1993). An optimum time scale for discrete Laguerre network. *IEEE Trans. Autom. Control*, **AC-38**, 934–938.

Gottlieb, M. J. (1938). Concerning some polynomials orthogonal on finite or enumerable set of points. *Am. J. Math.*, **60**, 453–458.

Grenander, U. and G. Szegö (1958). *Toeplitz Forms and Their Applications*. University of California Press, Berkeley.

Hakvoort, R. G. and P. M. J. Van den Hof (1994). An instrumental variable procedure for identification of probabilistic frequency response uncertainty regions. In *Proc. 33rd IEEE Conf. on Decision and Control*, Lake Buena Vista, FL, pp. 3596–3601.

Hannan, E. J. and B. Wahlberg (1989). Convergence rates for inverse Toeplitz matrix forms. *J. Multiv. Anal.*, **31**, 127–135.

Heuberger, P. S. C. (1991). On approximate system identification with system based orthonormal functions. Doctoral dissertation, Delft University of Technology, The Netherlands.

Heuberger, P. S. C. and O. H. Bosgra (1990). Approximate system identification using system based orthonormal functions. In *Proc. 29th IEEE Conf. on Decision and Control*, Honolulu, HI, pp. 1086–1092.

Heuberger, P. S. C. and P. M. J. Van den Hof (1995). A new signals and systems transform induced by generalized orthonormal basis functions. Report N-487, Mechanical Engineering Systems and Control Group, Delft University of Technology, June 1995.

Heuberger, P. S. C., P. M. J. Van den Hof and O. H. Bosgra (1995). A generalized orthonormal basis for linear dynamical systems. *IEEE Trans. Autom. Control*, **AC-40**, 451–465.

Heuberger, P. S. C., P. M. J. Van den Hof and O. H. Bosgra (1993). A generalized orthonormal basis for linear dynamical systems. In *Proc. 32nd IEEE Conf. on Decision and Control*, San Antonio, TX, pp. 2850–2855.

Kautz, W. H. (1954). Transient syythesis in the time domain. *IRE Trans. Circ. Theory*, **CT-1**, 29–39.

King, R. E. and P. N. Paraskevopoulos (1979). Parametric identification of discrete time SISO systems. *Int. J. Control*, **30**, 1023–1029.

Lee, Y. W. (1933). Synthesis of electrical networks by means of the Fourier transforms of Laguerre functions. *J. Maths and Phys.*, **11**, 83–113.

Lee, Y. W. (1960). *Statistical Theory of Communication*. Wiley, New York.

Ljung, L. (1987). *System Identification—Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.

Ljung, L. and Z. D. Yuan (1985). Asymptotic properties of black-box identification of transfer functions. *IEEE Trans. Autom. Control*, **AC-30**, 514–530.

Nurges, Y. (1987). Laguerre models in problems of approximation and identification of discrete systems. *Autom. Rem. Control*, **48**, 346–352.

Nurges, Y. and Y. Yaaksoo (1981). Laguerre state equations for a multivariable discrete system. *Autom. Rem. Control*, **42**, 1601–1603.

Szegö, G. (1975). *Orthogonal Polynomials*, 4th ed. American Mathematical Society, Providence, RI.

Van den Hof, P. M. J., P. S. C. Heuberger and J. Bokor (1994). System identification with generalized orthonormal basis functions. In *Proc. 33rd IEEE Conf. on Decision and Control*, Lake Buena Vista, FL, pp. 3382–3387.

Wahlberg, B. (1990). On the use of orthogonalized exponentials in system identification. Report LiTH-ISY-1099, Department of Electrical Engineering, Linköping University, Sweden.

Wahlberg, B. (1991). System identification using Laguerre models. *IEEE Trans. Autom. Control*, **AC-36**, 551–562.

Wahlberg, B. (1994a). System identification using Kautz models. *IEEE Trans. Autom. Control*, **AC-39**, 1276–1282.

Wahlberg, B. (1994b). Laguerre and Kautz models. In *Preprints 10th IFAC Symp. on System Identification*, Copenhagen, Vol. 3, pp. 1-12.
Wiener, N. (1949). *Extrapolation, Interpolation and Smoothing of Stationary Time Series.* MIT Press, Cambridge, MA.

APPENDIX/PROOFS

Proof of Proposition 4.1

Let us consider the situation for $G(z) = z^{-1}$. Here $y(t) = u(t-1)$. Consequently, $\mathscr{Y}(k) := \sum_{t=0}^{\infty} \phi_k(t)y(t) = \sum_{t=0}^{\infty} \phi_k(t)u(t-1)$. From Heuberger *et al.* (1995), we know that for each $k \in \mathbb{Z}$,

$$\begin{bmatrix} \phi_0(t) \\ \phi_1(t) \\ \vdots \\ \phi_k(t) \end{bmatrix} = A_{k+1} \begin{bmatrix} \phi_0(t-1) \\ \phi_1(t-1) \\ \vdots \\ \phi_k(t-1) \end{bmatrix}, \tag{A.1}$$

with the matrix $A_k \in \mathbb{R}^{kn_b \times kn_b}$ given by

$$A_k = \begin{bmatrix} A & 0 & 0 & \dots & 0 \\ BC & A & 0 & \dots & 0 \\ BDC & BC & A & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ BD^{k-2}C & BD^{k-3}C & \dots & BC & A \end{bmatrix}. \tag{A.2}$$

Using this in the expression for $\mathscr{Y}(k)$ shows that

$$\begin{aligned} \mathscr{Y}(k) &= \sum_{t=0}^{\infty} \phi_k(t)u(t-1) \\ &= [0 \quad 0 \quad \dots \quad I_{n_b}]A_{k+1} \sum_{t=0}^{\infty} \begin{bmatrix} \phi_0(t-1) \\ \phi_1(t-1) \\ \vdots \\ \phi_k(t-1) \end{bmatrix} u(t-1) \\ &= [0 \quad 0 \quad \dots \quad I_{n_b}]A_{k+1} \begin{bmatrix} \mathscr{U}(0) \\ \mathscr{U}(1) \\ \vdots \\ \mathscr{U}(k) \end{bmatrix}. \end{aligned} \tag{A.3}$$

As a result, for each k,

$$\mathscr{Y}(k) = A\mathscr{U}(k) + BC\mathscr{U}(k-1) + BDC\mathscr{U}(k-2) + \dots + BD^{k-1}C\mathscr{U}(0).$$

This immediately shows that

$$\tilde{y}(\lambda) = [A + B(\lambda - D)^{-1}C]\tilde{u}(\lambda), \tag{A.4}$$

leading to $\tilde{G}(\lambda) = [A + B(\lambda - D)^{-1}C]$. Putting several time delay transfer functions in cascade, it follows straightforwardly that $G(z) = z^{-k}$ leads to $\tilde{G}(\lambda) = N(\lambda)^k$, which proves the result. □

Lemma A.1. Consider a scalar inner transfer function G_b generating an orthogonal basis as discussed in Section 2, with V_0 and N as defined by (11) and (33) respectively. Then

$$V_0(z)z = N(G_b(z))V_0(z). \tag{A.5}$$

Proof.

$$\begin{aligned} N(G_b(z))V_0(z)z^{-1} &= \{A + B[G_b(z) - D]^{-1}C\}(zI - A)^{-1}B \\ &= A(zI - A)^{-1}B \\ &\quad + B[G_b(z) - D]^{-1}C(zI - A)^{-1}B \\ &= A(zI - A)^{-1}B \\ &\quad + B[C(zI - A)^{-1}B]^{-1}C(zI - A)^{-1}B \\ &= A(zI - A)^{-1}B + B \\ &= zI(zI - A)^{-1}B = V_0(z). \end{aligned}$$

Proof of Proposition 4.4.

Write H in its Laurent expansion: $H(z) = \sum_{i=0}^{\infty} h_i z^{-i}$. Then $V_0(z)H(z^{-1}) = \sum_{i=0}^{\infty} h_i V_0(z)z^i$. Using Lemma A.1, it follows that this equals $\sum_{i=0}^{\infty} h_i N^i(G_b(z))V_0(z)$. Using (35) now shows that this is equivalent to $H(G_b(z))V_0(z)$. □

Proof of Proposition 6.1

(a) Denote the vector $\mu(n) := [\mu_0^{\mathrm{T}} \ \mu_1^{\mathrm{T}} \ \dots \ \mu_{n-1}^{\mathrm{T}}]^{\mathrm{T}}$, $\mu_i \in \mathbb{R}^{n_b}$. Then

$$\mu(n)^{\mathrm{T}}R(n)\mu(n) = \sum_{k=0}^{n-1}\sum_{l=0}^{n-1} \mu_k^{\mathrm{T}} c_{kl}\mu_l, \tag{A.6}$$

with

$$c_{kl} = \frac{1}{2\pi}\int_{-\pi}^{\pi} V_k(e^{i\omega})V_l^{\mathrm{T}}(e^{-i\omega})\Phi_u(\omega)\,d\omega. \tag{A.7}$$

Denoting $\eta(e^{i\omega}) := \sum_{k=0}^{n-1} \mu_k^{\mathrm{T}} V_k(e^{i\omega})$, it follows that

$$\mu(n)^{\mathrm{T}}R(n)\mu(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \eta(e^{i\omega})\Phi_u(\omega)\eta(e^{-i\omega})\,d\omega, \tag{A.8}$$

while the orthonormality of the basis functions implies that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |\eta(e^{i\omega})|^2\,d\omega = \mu(n)^{\mathrm{T}}\mu(n). \tag{A.9}$$

Since $\operatorname{ess\,inf}_\omega \Phi_u(\omega)\mu(n)^{\mathrm{T}}\mu(n) \le \mu(n)^{\mathrm{T}}R(n)\mu(n) \le \operatorname{ess\,sup}_\omega \Phi_u(\omega)\mu(n)^{\mathrm{T}}\mu(n)$, it follows that

$$\operatorname*{ess\,inf}_{\omega} \Phi_u(\omega) \le \|R(n)\|_2 \le \operatorname*{ess\,sup}_{\omega} \Phi_u(\omega). \tag{A.10}$$

The latter equation can be verified by noting that, since $R(n)$ is symmetric, there exists $Q(n)$ satisfying $R(n) = Q(n)Q(n)^{\mathrm{T}}$, leading to

$$\operatorname*{ess\,inf}_{\omega} \Phi_u(\omega) \le \|Q(n)\|_2^2 \le \operatorname*{ess\,sup}_{\omega} \Phi_u(\omega).$$

(b) The Hermitian form $T_n := \mu^{\mathrm{T}}(n)R(n)\mu(n)$ can be written as

$$T_n = v^{\mathrm{T}}(n)\,\mathrm{diag}\,\{\lambda_1^{(n)}, \dots, \lambda_{n_b n}^{(n)}\}\,v(n)$$

through unitary transformation preserving the norm, i.e. $v^{\mathrm{T}}(n)v(n) = \mu^{\mathrm{T}}(n)\mu(n)$. Consequently,

$$\frac{\mu(n)^{\mathrm{T}}R(n)\mu(n)}{\mu^{\mathrm{T}}(n)\mu(n)} \le \max_i \lambda_i^{(n)}, \tag{A.11}$$

which is known to be bounded by $\operatorname{ess\,sup}_\omega \Phi_u(\omega)$. Since the Hermitian form T_n is related to the Toeplitz form (A.8), Theorem 5.2.1 of Grenander and Szegö (1958) leads directly to the result that

$$\lim_{n\to\infty} \max_i \lambda_i^{(n)} = \operatorname*{ess\,sup}_{\omega} \Phi_u(\omega).$$

Proof of Proposition 6.2

With Proposition 6.1(a), it follows that $\|R(n)^{-1}\|_2 \le [\operatorname{ess\,inf}_\omega \Phi_u(\omega)]^{-1}$. For constructing a bound on the second term on the right-hand side of (54), we first establish the following notation. Denote $R_{12}(n) := \bar{E}[\psi(t)\psi_e^{\mathrm{T}}(t)]$; then we can write

$$\begin{aligned} R := \bar{E}\left\{\begin{bmatrix} \psi(t) \\ \psi_e(t) \end{bmatrix}[\psi^{\mathrm{T}}(t) \quad \psi_e^{\mathrm{T}}(t)]\right\} \\ = \begin{bmatrix} R(n) & R_{12}(n) \\ R_{21}(n) & R_{22}(n) \end{bmatrix}, \end{aligned}$$

which is an infinite block-Toeplitz matrix, of which $R(n)$ is a finite part. Since

$$R_{12}(n) = [I_{n_b n} \quad 0]R\begin{bmatrix} 0_{n_b n \times \infty} \\ I \end{bmatrix}$$

it follows that

$$\begin{aligned} \|R_{12}(n)\|_2 &\le \|[I_{n_b n} \quad 0]\|_2\, \|R\|_2 \left\|\begin{bmatrix} 0_{n_b n \times \infty} \\ I \end{bmatrix}\right\|_2 \\ &\le \|R\|_2. \end{aligned}$$

As a result,

$$\|R_{12}(n)\|_2 \le \lim_{n\to\infty} \max_j \lambda_j(R(n)),$$

which, by Proposition 6.1(b), is equal to $\operatorname{ess\,sup}_\omega \Phi_u(\omega)$. This proves the result. □

Proof of Proposition 6.3

Writing

$$G(e^{i\omega}, \theta^*) - G_0(e^{i\omega}) = [\Omega^{\mathrm{T}}(e^{i\omega}) \quad \Omega_e^{\mathrm{T}}(e^{i\omega})]\begin{bmatrix} \theta^* - \theta_0 \\ \theta_e \end{bmatrix}, \tag{A.12}$$

it follows that, for each ω_1,

$$|G(e^{i\omega_1}, \theta^*) - G_0(e^{i\omega_1})| \leq \|[\Omega^T(e^{i\omega_1}) \quad \Omega_e^T(e^{i\omega_1})]\|_1 \left\| \begin{bmatrix} \theta^* - \theta_0 \\ \theta_e \end{bmatrix} \right\|_1,$$

where $\|\cdot\|_1$ denotes the induced ℓ_1 matrix norm and the ℓ_1 norm respectively. It follows from the fact that G_b is inner that

$$\begin{aligned} &|G(e^{i\omega_1}, \theta^*) - G_0(e^{i\omega_1})| \\ &\leq \|V_0^T(e^{i\omega_1})\|_1 \left\| \begin{bmatrix} \theta^* - \theta_0 \\ \theta_e \end{bmatrix} \right\|_1 \\ &= \|V_0^T(e^{i\omega_1})\|_1 (\|\theta^* - \theta_0\|_1 + \|\theta_e\|_1) \\ &= \|V_0(e^{i\omega_1})\|_\infty (\|\theta^* - \theta_0\|_1 + \|\theta_e\|_1). \end{aligned}$$

Part (a) of the proposition now follows by substituting the error bound obtained in Proposition 6.2 and using the inequality $\|\theta^* - \theta_0\|_1 \leq \sqrt{n_b n}\, \|\theta^* - \theta_0\|_2$.

Because of the orthonormality of the basis functions on the unit circle, it follows that

$$\|G(z, \theta^*) - G_0(z)\|_{\mathcal{H}_2} = \left\| \begin{bmatrix} \theta^* - \theta_0 \\ \theta_e \end{bmatrix} \right\|_2,$$

which, together with Proposition 6.2, proves the result (b). □

Proof of Proposition 6.4

We shall only give a sketch of the proof of this result. For a full proof see Heuberger (1991) and Heuberger and Van den Hof (1995). The proof is constructed along the following line of reasoning.

(a) Let $x_0(t)$ be the impulse response of $zG_0(z) = \sum_{k=0}^{\infty} L_k V_k(z)$. Then $\tilde{x}_0^T(\lambda) = \sum_{k=0}^{\infty} L_k \lambda^{-k}$. This follows straightforwardly by applying the definitions of the Hambo transformations. Consequently, the poles of $\tilde{x}_0(\lambda)$ will determine (a lower bound on) the speed of convergence of the series expansion of G_0.

(b) $\tilde{x}_0(\lambda) = \tilde{G}_0(\lambda) C^T \dfrac{\lambda}{\lambda - D}$. This follows from applying the Hambo transform to the relation $x_0(t) = G_0(z)\delta(t+1)$, with $\delta(t)$ the unit pulse.

(c) Let $\{\mu_i\}_{i=1,\ldots,n_s}$ denote the poles of G_0. Then the poles of $\tilde{G}_0$ are given by $\lambda_i = \{G_b^{-1}(\mu_i)\}_{i=1,\ldots,n_s}$, leading to

$$|\lambda_i| = \prod_{j=1}^{n_b} \left| \frac{\mu_i - \rho_j}{1 - \mu_i \rho_j} \right|.$$

The first statement was proved by Van den Hof *et al.* (1994), and the latter implication follows directly by substituting the appropriate expressions.

(d) The poles of $\tilde{x}_0(\lambda)$ are the poles of $\tilde{G}_0(\lambda)$, possibly minus poles in $\lambda = 0$. This follows from careful analysis of the expression in (b).

(e) Since λ in (58) is the maximum amplitude of poles of $\sum_{k=0}^{\infty} L_k \lambda^{-k}$, it follows that there exist $\alpha_i \in \mathbb{R}$ such that for all $\eta > \lambda$, $|L_{k,i}| \leq \alpha_i \eta^{k+1}$, with $L_{k,i}$ the ith element of L_k. As $\|\theta_e\|_2^2 = \sum_{k=n}^{\infty} \sum_{i=1}^{n_b} L_{k,i}^2$, the result follows by substitution.

Lemma A.2. Let v be a scalar-valued stationary stochastic process with rational spectral density $\Phi_v(\omega)$. Let $e^{-i\tilde{\omega}} = G_b(e^{i\omega})$. Then

$$V_0^T(e^{i\omega}) \tilde{\Phi}_v(\tilde{\omega}) V_0(e^{-i\omega}) = V_0^T(e^{i\omega}) \Phi_v(\omega) V_0(e^{-i\omega}). \quad \text{(A.13)}$$

Proof. Let H_v be a stable spectral factor of Φ_v, satisfying $\Phi_v(\omega) = H_v(e^{-i\omega}) H_v(e^{i\omega})$. Then $\tilde{\Phi}_v(\omega) = \tilde{H}_v^T(e^{-i\omega}) \tilde{H}_v(e^{i\omega})$, and it follows that

$$\begin{aligned} &V_0^T(e^{i\omega}) \tilde{\Phi}_v(\tilde{\omega}) V_0(e^{-i\omega}) \\ &= V_0^T(e^{i\omega}) \tilde{H}_v^T(G_b(e^{i\omega})) \tilde{H}_v(G_b(e^{-i\omega})) V_0(e^{-i\omega}). \end{aligned}$$

Using Proposition 4.4, this latter expression is equal to $H_v(e^{-i\omega}) V_0^T(e^{i\omega}) V_0(e^{-i\omega}) H_v(e^{i\omega})$. Since H_v is scalar, this proves the lemma. □

Proof of Theorem 6.2

Using (62) and substituting $\Omega_n(e^{i\omega})$ shows that

$$\begin{aligned} &\frac{1}{n_b n} \Omega_n^T(e^{i\omega_1}) Q_n \Omega_n(e^{-i\omega_2}) \\ &= \frac{1}{n_b n} [V_0^T(e^{i\omega_1}) \quad V_1^T(e^{i\omega_1}) \quad \ldots \quad V_{n-1}^T(e^{i\omega_1})] Q_n \\ &\quad \times [V_0^T(e^{-i\omega_2}) \quad V_1^T(e^{-i\omega_2}) \quad \ldots \quad V_{n-1}^T(e^{-i\omega_2})]^T. \end{aligned}$$

Note that this latter expression can be written as

$$\begin{aligned} &\frac{1}{n_b n} V_0^T(e^{i\omega_1}) [I \quad G_b(e^{i\omega_1}) I \quad \ldots \quad G_b^{n-1}(e^{i\omega_1}) I] Q_n \\ &\quad \times [I \quad G_b(e^{-i\omega_2}) I \quad \ldots \quad G_b^{n-1}(e^{-i\omega_2}) I]^T V_0(e^{-i\omega_2}). \quad \text{(A.14)} \end{aligned}$$

We now evaluate the expression

$$\begin{aligned} &\frac{1}{n} [I \quad G_b(e^{i\omega_1}) I \quad \ldots \quad G_b^{n-1}(e^{i\omega_1}) I] Q_n \\ &\quad \times [I \quad G_b(e^{-i\omega_2}) I \quad \ldots \quad G_b^{n-1}(e^{-i\omega_2}) I]^T. \quad \text{(A.15)} \end{aligned}$$

Since G_b is an inner function, we can consider the variable transformation

$$e^{-i\tilde{\omega}} := G_b(e^{i\omega}). \quad \text{(A.16)}$$

Employing this transformation in the expression (A.15), the latter is equivalent to

$$\frac{1}{n} [I \quad e^{-i\tilde{\omega}_1} I \quad \ldots \quad e^{-i\tilde{\omega}_1 (n-1)} I] Q_n [I \quad e^{i\tilde{\omega}_2} I \quad \ldots \quad e^{i\tilde{\omega}_2 (n-1)} I]^T.$$

The convergence result of Hannan and Wahlberg (1989) and Ljung and Yuan (1985) now show that, as $n \to \infty$, this expression converges to

$$\begin{cases} 0 & \text{if } \tilde{\omega}_1 \neq \tilde{\omega}_2, \\ Q(\tilde{\omega}_1) & \text{if } \tilde{\omega}_1 = \tilde{\omega}_2, \end{cases} \quad \text{(A.17)}$$

where $Q(\omega)$ is the spectral density related to the Toeplitz matrix Q_n in the limit as $n \to \infty$.

Employing Lemma A.2, together with (61), it follows that, as $n \to \infty$, the spectrum related to the Toeplitz matrix Q_n is given by $\tilde{\Phi}_v(\omega) \tilde{\Phi}_u(\omega)^{-1}$. This is due to the fact that the symbol (spectrum) of a Toeplitz matrix that is the product of several Toeplitz matrices asymptotically (as $n \to \infty$) equals the product of the symbols (spectra) of the separate Toeplitz matrices (see Grenander and Szegö, 1958). Combining this with (A.14) now shows that, as $n \to \infty$,

$$\begin{aligned} &\frac{1}{n_b n} \Omega_n^T(e^{i\omega_1}) Q_n \Omega_n(e^{-i\omega_2}) \\ &= \frac{1}{n_b} V_0^T(e^{i\omega_1}) Q(\tilde{\omega}) V_0(e^{-i\omega_2}) \\ &= \begin{cases} 0 & \text{if } \tilde{\omega}_1 \neq \tilde{\omega}_2, \\ \dfrac{1}{n_b} V_0^T(e^{i\omega_1}) \tilde{\Phi}_v(\tilde{\omega}_1) \tilde{\Phi}_u(\tilde{\omega}_1)^{-1} V_0(e^{-i\omega_1}) & \text{if } \tilde{\omega}_1 = \tilde{\omega}_2. \end{cases} \end{aligned} \quad \text{(A.18)}$$

Employing Lemma A.2 now shows that the resulting expression in (A.18) becomes

$$\frac{1}{n_b} V_0^T(e^{i\omega_1}) \Phi_v(\omega_1) \Phi_u(\omega_1)^{-1} V_0(e^{-i\omega_1}).$$

In using Lemma A.2 it must be realized that there will always exist a scalar-valued stationary stochastic process z with rational spectrum Φ_z such that $\tilde{\Phi}_z = \tilde{\Phi}_v \tilde{\Phi}_u^{-1}$. □