Frequency Domain Identification with Generalized Orthonormal Basis Functions

Douwe K. de Vries and Paul M. J. Van den Hof, *Senior Member, IEEE*

Abstract— A method is considered for the identification of linear parametric models based on a least squares identification criterion that is formulated in the frequency domain. To this end, use is made of the empirical transfer function estimate (ETFE), identified from time-domain data. As a parametric model structure use is made of a finite expansion sequence in terms of recently introduced generalized basis functions, being generalizations of the classical pulse and Laguerre and Kautz types of bases. An asymptotic analysis of the estimated models is provided and conditions for consistency are formulated. Explicit and transparent bias and variance expressions are established, the latter ones also valid in a situation of undermodeling.

Index Terms—Asymptotic analysis, frequency domain method, system approximation, system identification, transfer function estimation.

I. INTRODUCTION

THE IDENTIFICATION of models on the basis of frequency domain data is a subject that attracts a growing number of researchers and engineers. Especially in application areas where experimental data of a process with (partly) unknown dynamics can be taken relatively cheaply, the excitation of the process with periodic signals (e.g., sinusoids) is an attractive way of extracting accurate information of the process dynamics from experiments. Due to the commercial availability of frequency analyzers that can handle huge amounts of data by special purpose hardware, the experimental determination of frequency responses of dynamical systems has gained increasing interest in application areas as the modeling of, e.g., mechanical servo systems and flexible space structures.

Identification on the basis of frequency domain data can have a number of advantages when compared to the "classical" time-domain approach. For a very nice overview of these arguments the reader is referred to [28]. Here we would like to limit attention to the fact that dealing with frequency domain data allows us to achieve a substantial data reduction, thus enabling us to handle huge amounts of time-domain data with a corresponding reduced variance of the model estimates. Additionally, the formulation of an identification criterion in the frequency domain can be beneficial, especially in those situations where the application of the model dictates a performance evaluation in terms of frequency domain properties. This last situation occurs often when the identified models are used as a basis for model-based control design.

On the other hand, it should be stressed that any frequency domain data is obtained by some data handling/processing mechanism that starts off with time-domain data. This is reason not to overestimate the difference between time-domain and frequency-domain identification; see, e.g., [18].

The common way of formulating an identification problem in the frequency domain is by assuming the availability of the exact frequency response of the (unknown) linear system, disturbed by some additive (frequency-domain) noise with specific properties, e.g., independence among the several frequencies. For this situation a large number of identification methods exist, mostly dealing with least squares criteria [14], [29], [38], [31]. Maximum amplitude (ℓ_∞) criteria are considered in [32] and [9], while special purpose multivariable algorithms are discussed, e.g., in [15] and [1]. Recently subspace algorithms also have been analyzed for frequency domain identification [21]. Many more references and techniques can be found in [30] and [28]. A related approach to the problem based on the discrete Fourier transforms of input and output data in [18] shows a close resemblance of results with the standard time-domain approach.

In this paper we will formulate and analyze an identification problem with an identification criterion formulated in the frequency domain, for a situation in which we consider time-domain data of input and output samples to be available from measurements. These time-domain signals generate a (frequency-domain) empirical transfer function estimate (ETFE) which represents the data in a least squares identification criterion. As a parametric model structure we will use a linear regression form, using a finite series expansion of the model transfer function in terms of very flexible orthonormal basis functions as recently introduced in [12] and [34]. This model structure is very powerful in accurately describing system's transfer functions with only few parameters. It generalizes the situation of Laguerre functions, for which a frequency domain identification analysis has been provided in [3]. The resulting identification scheme will be analyzed and expressions for bias and variance will be formulated, relating the identification results to the properties of the original

excitation and disturbance signals in the time domain. This method will build on the standard identification framework as developed in [17].

Concerning notation we will denote R as the set of reals and Z as the set of integer numbers; $\mathcal{R}\mathcal{H}_2$ is the Hilbert space of transfer functions that are analytic in $|z| \geq 1$ and square integrable on the unit circle; $(\cdot)^*$ will denote complex conjugate transpose of a complex-valued matrix. The remainder of this paper is organized as follows. First, we will formalize our problem setting. Subsequently, we will analyze the estimation results for an identification criterion over a finite number of frequencies. Next the situation of an infinite number of frequencies will be considered. Asymptotic bias and asymptotic variance expressions will be provided, for infinite and for finite model order (undermodeling). The presented identification method will be illustrated by way of a simulation example and by application to data taken from a high-precision mechanical servosystem. The proofs of all results are collected in the Appendix. This paper is an extended version of [6]; some considerations concerning the multivariable situation are given in [26].

II. PROBLEM SETTING

It is assumed that a data generating system, and the measurement data that is obtained from this system, allow a description

$$y(t) = G_o(q)u(t) + v(t) \tag{1}$$

where $G_o(z)$ is a rational transfer function, analytic in $|z| \geq 1$, $v(t)$ is a realization of a stationary stochastic process with rational spectral density $\Phi_v(\omega)$, satisfying $v(t) = H_o(q)e(t)$, where $e(t)$ is a sequence of independent identically distributed random variables, having zero mean, variance σ_e^2, and bounded moments of all orders. H_o is a stable proper minimum-phase transfer function; $u(t)$ will be a bounded quasistationary deterministic signal; q is the shift operator $qu(t) = u(t+1)$.

As further notational convention, let T^N denote the integer interval $T^N := Z \cap [0, N-1]$. For a signal $x(t)$ being defined on T^N, we will denote

$$X_N(e^{i\omega_k}) := \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} x(t)e^{-i\omega_k t} \quad \omega_k \in \Omega_N. \tag{2}$$

Specific sets of frequency points that arise are denoted as

$$\Omega_N := \left\{ \frac{2\pi k}{N}, \ k \in T^N \right\}$$
$$\Omega_N^u := \left\{ \omega_k \in \Omega_N \mid |U_N(e^{i\omega_k})| > 0 \right\}. \tag{3}$$

The identification problem that we consider now can be stated as follows.

Frequency Domain Identification Problem: Given measurements of input and output data $\{u(t), y(t)\}$, $t \in T^N$, taken from the system (1), and given a frequency grid $\mathcal{W}_{N_p}^u$

$$\mathcal{W}_{N_p}^u = \{\omega_1, \cdots, \omega_{N_p}\} \subset \Omega_N^u \tag{4}$$

such that all frequencies $e^{i\omega_k}$ with $\omega_k \in \mathcal{W}_{N_p}^u$ come in complex conjugate pairs, then we consider the identification

problem

$$\hat{\theta} := \arg\min_\theta W_{N,N_p}(\theta) \tag{5}$$

with

$$W_{N,N_p}(\theta) = \frac{1}{N_p} \sum_{k=1}^{N_p} \left| \frac{Y_N(e^{i\omega_k})}{U_N(e^{i\omega_k})} - G(e^{i\omega_k}, \theta) \right|^2 \cdot |K(e^{i\omega_k})|^2$$

where $G(e^{i\omega_k}, \theta)$ is the parameterized frequency response of the model, and θ ranges over an appropriate parameter space $\Theta \in R^{n_p}$ while $n_p \leq N_p$. The specific (linear) parameterization of the model is considered in the next section. $K(e^{i\omega_k})$ is a user-chosen weighting function that allows us to emphasize specific frequency regions more than others; for ease of notation (and without loss of generality) the weighting is assumed to satisfy $|K(e^{i\omega_k})| > 0$ for all $\omega_k \in \mathcal{W}_{N_p}^u$.

Note that even with a constant weighting over frequency, in the least squares criterion (5) more attention is paid to those frequency regions which have a relatively denser frequency grid within $\mathcal{W}_{N_p}^u$.

The chosen identification criterion was suggested before in [33]. It closely matches the criterion treated in [18] and [19] which can be written as

$$\hat{\theta} = \arg\min \frac{1}{N_p} \sum_{k=1}^{N_p} |Y_N(e^{i\omega_k}) - G(e^{i\omega_k}, \theta)U_N(e^{i\omega_k})|^2|$$
$$\cdot \frac{1}{|H(e^{i\omega_k}, \theta)|^2}$$

and that is also employed in [30] in the more general setting of errors-in-variables models. In our choice of $W_{N,N_p}(\theta)$ we focus attention on the role of estimating the input–output dynamics of the plant, thus restricting to the situation of a fixed noise model $H(e^{i\omega_k}) = U_N(e^{i\omega_k})/K(e^{i\omega_k})$. In this particular case, we will investigate the effect of a specific choice of parameterization for G, as presented in the next section; this identification approach can be referred to as an output-error method, according to the notions used for time-domain methods [17]. The (time-domain) method is known for its property that the plant G_0 can be estimated consistently, while its variance will be minimal when the (fixed) noise model matches the true noise process dynamics. This latter situation is not treated here.

III. MODEL PARAMETERIZATION WITH ORTHONORMAL BASIS FUNCTIONS

In the identification problem (5) we will employ a linear model parameterization

$$G(e^{i\omega_k}, \theta) = \phi^T(e^{i\omega_k})\theta \tag{6}$$

that is very flexible, based on a recently introduced generalized orthonormal basis. The related theory, as presented in [12], shows that for any scalar stable all-pass transfer function G_b with balanced realization (A, B, C, D) the sequence of functions

$$\mathcal{V}_k(z) = (zI - A)^{-1}B \cdot G_b(z)^k \tag{7}$$

generates an orthonormal basis for the space of stable systems $\mathcal{RH}_2$, with the property that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{V}_j(e^{i\omega}) \mathcal{V}_k^T(e^{-i\omega}) \, d\omega = \begin{cases} I, & j = k \\ 0, & j \neq k. \end{cases} \quad (8)$$

As a result there exist unique D_o and $\{L_k\}_{k=0,\cdots,\infty}$ such that

$$G_o(z) = D_o + \sum_{k=0}^{\infty} L_k^T \mathcal{V}_k(z). \quad (9)$$

Note that $\mathcal{V}_k(z) \in \mathcal{RH}_2^{n_b \times 1}$ with n_b the McMillan degree of G_b (dimension of A), and $L_k \in R^{n_b \times 1}$.

The advantage of this generalized basis is that if an appropriate choice of dynamics (set of poles) is incorporated into G_b, and thus into the basis functions $\mathcal{V}_k$, then the series expansion (9) shows an increasing rate of convergence. Consequently, the accuracy of a finite expansion model will substantially increase; for more details on the use of these basis functions see [12], [34], and [25]. Here we limit the discussion by remarking that for specific choices of all-pass functions, well-known "classical" basis functions result, as the standard shift $\mathcal{V}_k(z) = z^{-k}$ is the result of $G_b(z) = z^{-1}$ and the Laguerre functions induced by a first-order all-pass function $G_b(z) = (1 - az)/(z - a)$ for some $|a| < 1$.

With respect to our formulated frequency domain identification problem, it follows that—for a specifically chosen basis $\mathcal{V}_k(z)$—the regression vector $\phi^T(e^{i\omega_k})$ in (6) is given by

$$\phi(e^{i\omega_k}) = \begin{bmatrix} 1 & \mathcal{V}_0^T(e^{i\omega_k}) & \cdots & \mathcal{V}_{n-1}^T(e^{i\omega_k}) \end{bmatrix}^T \quad (10)$$

which is a column vector with length $n_p = n_b n + 1$, equal to the number of parameters. The estimated parameter vector is denoted by

$$\hat{\theta} = \begin{bmatrix} \hat{D} & \hat{L}_0^T & \cdots & \hat{L}_{n-1}^T \end{bmatrix}^T. \quad (11)$$

IV. PARAMETER ESTIMATE

Building upon (9) for the data generating system, we will write

$$G_o(e^{i\omega_k}) = \phi^T(e^{i\omega_k})\theta_o + Z(e^{i\omega_k}) \quad (12)$$

with

$$Z(e^{i\omega_k}) = \sum_{k=n}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega_k}) $$

and

$$\theta_o = \begin{bmatrix} D & L_0^T & \cdots & L_{n-1}^T \end{bmatrix}^T. \quad (13)$$

Using the system's equations, similar as in [17], we now can express the ETFE by

$$\frac{Y_N(e^{i\omega_k})}{U_N(e^{i\omega_k})} = G_o(e^{i\omega_k}) + S(e^{i\omega_k}) + F(e^{i\omega_k}) \quad (14)$$

where

$$F(e^{i\omega_k}) = \frac{V_N(e^{i\omega_k})}{U_N(e^{i\omega_k})} \quad (15)$$

and where $S(e^{i\omega_k})$ is a term due to the past of the input signal, $\{u(t)\}_{t<0}$. Written in vector notation we can rewrite this into

$$\begin{aligned} \mathsf{G} &= \mathsf{G}_o + \mathsf{S} + \mathsf{F} \\ &= \Phi\theta_o + \mathsf{Z} + \mathsf{S} + \mathsf{F} \end{aligned} \quad (16)$$

with

$$\begin{aligned}
\mathsf{G} &= \begin{bmatrix} \dfrac{Y_N(e^{i\omega_1})}{U_N(e^{i\omega_1})} & \cdots & \dfrac{Y_N(e^{i\omega_{N_p}})}{U_N(e^{i\omega_{N_p}})} \end{bmatrix}^T \\
\mathsf{G}_o &= \begin{bmatrix} G_o(e^{i\omega_1}) & \cdots & G_o(e^{i\omega_{N_p}}) \end{bmatrix}^T \\
\mathsf{S} &= \begin{bmatrix} S(e^{i\omega_1}) & \cdots & S(e^{i\omega_{N_p}}) \end{bmatrix}^T \\
\mathsf{F} &= \begin{bmatrix} F(e^{i\omega_1}) & \cdots & F(e^{i\omega_{N_p}}) \end{bmatrix}^T \\
\mathsf{Z} &= \begin{bmatrix} Z(e^{i\omega_1}) & \cdots & Z(e^{i\omega_{N_p}}) \end{bmatrix}^T \\
\Phi &= \begin{bmatrix} 1 & \mathcal{V}_0^T(e^{i\omega_1}) & \cdots & \mathcal{V}_{n-1}^T(e^{i\omega_1}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mathcal{V}_0^T(e^{i\omega_{N_p}}) & \cdots & \mathcal{V}_{n-1}^T(e^{i\omega_{N_p}}) \end{bmatrix}
\end{aligned} \quad (17)$$

where Z reflects a term due to undermodeling, S represents the effect of unknown past inputs, and F reflects a term due to the noise.

The parameter estimate can be constructed from matrix operations according to

$$\hat{\theta} = \Psi\mathsf{G} \quad (18)$$

where

$$\Psi = (\Phi^*\Lambda\Phi)^{-1}\Phi^*\Lambda \quad (19)$$

with

$$\Lambda = \text{diag}\big(|K(e^{i\omega_1})|^2 \cdots |K(e^{i\omega_{N_p}})|^2\big) $$

and provided $\Phi^*\Lambda\Phi$ is nonsingular. Given (16), (18), and (19) it follows that the least squares parameter estimate (5) can be expressed as

$$\hat{\theta} = \theta_o + \Psi\mathsf{Z} + \Psi\mathsf{S} + \Psi\mathsf{F}. $$

For the bias and covariance of the estimated parameters (5) it can simply be verified that

$$\mathbb{E}[\hat{\theta}] = \theta_o + \Psi\mathsf{Z} + \Psi\mathsf{S} \quad (20)$$

$$\text{cov}[\hat{\theta}] = \Psi\, \mathbb{E}[\mathsf{F}\mathsf{F}^*]\Psi^*. \quad (21)$$

The transfer function estimate of the identified model will be given by

$$\hat{G}(e^{i\omega}) := G(e^{i\omega}, \hat{\theta}) = \phi^T(e^{i\omega})\hat{\theta}. \quad (22)$$

Invertability of the matrix $\Phi^*\Lambda\Phi$ in (19) is guaranteed if a sufficient number of different frequencies is selected in the frequency grid, formalized in the following lemma.[1]

---

[1] The authors are indebted to P. Heuberger for his contribution to the construction of the proof of this lemma.

Lemma 4.1: If $N_p \geq n_p$, then Φ has full column rank, and so Ψ (19) is well defined.

As a result of this lemma, the parameter estimate (5), (18) will be unique, provided that a sufficient number of sinusoids is present in the input signal. This situation is similar to the corresponding time-domain least squares problem. Note that in the case considered here, the regression vector ϕ is composed of basis functions only.

In order to relate the parameter estimate to the dynamic properties of the underlying system, we have to impose additional conditions on the input signal. To this end the following assumption is formulated.

Assumption 4.2: The input signal $u(t)$ is periodic, having r periods of length N_o, with $N = r \cdot N_o$, $r \geq 1$, such that $|U_{N_o}(e^{i\omega_k})|^2$ is bounded and bounded away from zero for all $\omega_k \in \mathcal{W}_{N_p}^u$, while for the situation $N_o \to \infty$ it is assumed that there exists a bounded function $\Phi_u(\omega)$ such that

$$\lim_{N_o \to \infty} |U_{N_o}(e^{i\omega_k})|^2 = \Phi_u(\omega_k) \quad \forall \omega_k \in \mathcal{W}_{N_p}^u \qquad (23)$$

where $\Phi_u^{-1}(\omega)$, $\omega \in [0, 2\pi)$ is continuous, except for at most a finite number of points, and bounded.

It has to be noted that for $N_o \to \infty$, the notation $\Phi_u(\omega_k)$ is used for the spectrum of one period of the signal. This is different from the spectrum of the periodic signal, since it can be verified that

$$U_N(e^{i\omega_k}) = \sqrt{\frac{N}{N_o}} \cdot U_{N_o}(e^{i\omega_k}). \qquad (24)$$

Further, the number of frequencies N_p will necessarily satisfy the inequality $N_p \leq N_0$.

For analyzing the statistical properties of the estimate defined in this section, use will be made of the following classical result for ETFE estimates.

Lemma 4.3: Let the input signal satisfy Assumption 4.2. Then

1) (Brillinger [2]) for finite N_0 and $N \to \infty$ the stochastic variable $\sqrt{r}\, \mathsf{F}$ tends to a normal distribution

$$\sqrt{r}\, \mathsf{F} \in \mathcal{N}(0, Q)$$

with

$$Q = \begin{bmatrix} \frac{\Phi_v(\omega_1)}{|U_{N_o}(e^{i\omega 1})|^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\Phi_v(\omega_{N_p})}{|U_{N_o}(e^{i\omega N_p})|^2} \end{bmatrix}$$

2)

$$\lim_{N \to \infty} r\, \mathbb{E}[\mathsf{F}\mathsf{F}^*] = Q \qquad (25)$$

where the rate of convergence is $1/N$ elementwise, and there exists a constant $c_1 \in R$ such that

$$\|r\, \mathbb{E}[\mathsf{F}\mathsf{F}^*] - Q\|_2 \leq \frac{c_1}{r}. \qquad (26)$$

The result of Brillinger shows that in order for F to obtain a covariance that tends to zero for increasing N, the input signal needs to be periodic with $r \to \infty$. This follows by observing that $|U_{N_o}(e^{i\omega_k})|^2$ is a bounded function.

V. ASYMPTOTIC ANALYSIS OF PARAMETER ESTIMATE

A. Analysis for Finite Number N_p of Frequencies

We will first restrict the analysis of the identified model to the situation where the number of frequencies N_p in the criterion is finite and the input signal is periodic with finite period $N_o \geq N_p$.

Using Lemma 4.3 we can state the following result for the estimated parameters in the asymptotic case when $N \to \infty$.

Proposition 5.1: Consider Assumption 4.2 and the expressions (20) and (21) of the estimated parameters $\hat{\theta}$ (5), while $n_p \leq N_p \leq N_o < \infty$. Then

1) $\text{plim}_{N \to \infty} \hat{\theta} = \theta^*$, where $\theta^* = \theta_o + \Psi\mathsf{Z}$;
2) for $N \to \infty$, the stochastic variable $\sqrt{N} \cdot (\hat{\theta} - \theta^*)$ tends to a normal distribution: $\sqrt{N} \cdot (\hat{\theta} - \theta^*) \in \mathcal{N}(0, P)$, where the asymptotic covariance matrix P is determined by its scalar elements

$$P_{ik} = N_o \sum_{\ell=1}^{N_p} \Psi_{i\ell} \Psi_{k\ell}^* \frac{\Phi_v(\omega_\ell)}{\Phi_u(\omega_\ell)},$$
$$i, k = 1, \cdots, n_p. \qquad (27)$$

Additionally

$$\lim_{N \to \infty} N \cdot \text{cov}(\hat{\theta}) = P.$$

Part 1) of the proposition shows that in case there is no undermodeling, i.e., $\mathsf{Z} \equiv 0$, the parameter estimate is (weakly) consistent. A similar consistency result will hold for the estimated frequency response of the model.

Proposition 5.1 provides an asymptotic expression for the covariance of the estimated parameters. However, (27) does not provide much insight in the underlying mechanisms. Structurally more simple expressions can be obtained when we let the number N_p of frequencies tend to infinity also. This will be discussed in the following sections.

B. Analysis for Infinite Number N_p of Frequencies

We will derive expressions for the bias and variance in the parameter estimate (5), for the asymptotic situation that the identification criterion is calculated over an infinite number N_p of frequencies. To this end we will consider the situation of a periodic input signal with r periods of length N_p. In these asymptotic results we restrict the choice of frequency grid in such a way that the frequency points are equidistantly spaced.

Assumption 5.2: The frequency grid $\mathcal{W}_{N_p}^u \subset \Omega_N^u$ (4), (3) satisfies

$$\omega_k = \frac{2\pi(k-1)}{N_p} \quad \forall N_p > 0$$

with $N_p = N_o$, and the frequency domain weighting $|K(e^{i\omega_k})|$ is fixed to one for all $\omega_k \in \mathcal{W}_{N_p}^u$.

This assumption is made to facilitate the analysis of the asymptotic case and to fully employ the orthonormality property of the basis functions used. The possibility of loosening these conditions will be discussed later.

We start the asymptotic bias analysis by evaluating the properties of the matrix $\Phi^*\Phi$.

Lemma 5.3: Consider Assumption 5.2. Then

$$\frac{1}{N_p}\Phi^*\Phi \to I \quad \text{as} \quad \frac{n}{N_p} \to 0$$

where the rate of convergence is n/N_p elementwise.

Additionally, there exists a constant $c_4 \in R$ such that

$$\left\|\frac{1}{N_p}\Phi^*\Phi - I\right\|_2 \le c_4 \frac{n_p n}{N_p}.$$

This lemma shows that the parameter estimate is numerically very well conditioned in the asymptotic case.

The next lemma is also instrumental in the analysis of the estimated models.

Lemma 5.4: Consider Assumption 5.2. Then there exists a $c_5 \in R$ such that

$$\left\|\Psi - \frac{1}{N_p}\Phi^*\right\|_2 \le c_5 \frac{n_p n}{N_p^{3/2}} \quad \text{as} \quad \frac{n_p n}{N_p} \to 0. \quad (28)$$

This lemma shows that in the considered limiting situation the pseudoinverse of the matrix Φ can simply be obtained by taking $\frac{1}{N_p}\Phi^*$. This of course greatly facilitates both calculation and analysis of the asymptotic least squares estimates.

The following theorem addresses the asymptotic properties of the estimated parameters.

Theorem 5.5: Consider Assumptions 4.2 and 5.2 and the estimated parameters (5). Then:

1) $\text{plim}_{N,N_p\to\infty}\hat{\theta} = \theta^*$, with $\theta^* = \theta_o$;
2) $\lim_{N,N_p,r\to\infty} N\,\text{cov}(\hat{\theta}) = R$, where the bounded matrix R is given by

$$R = \lim_{N_p\to\infty} \frac{1}{N_p}\Phi^*Q\Phi \quad (29)$$

with Q as defined in (25).

This theorem shows that asymptotically the parameterized part of the model can be estimated unbiasedly, despite the possible presence of undermodeling. If additionally the input signal is periodic with the number of periods tending to infinity also, the estimate becomes (weakly) consistent, as in that case the variance of the estimated parameters tends to zero.

Note that the consistency of the parameter estimate is quite remarkable, as we do accept the presence of unmodeled dynamics. For time-domain identification using an FIR model structure, a similar result of consistent parameter estimates holds also in the case of unmodeled dynamics, but in that case the input signal is restricted to be white noise; see [24]. In the considered situation here, it is the orthonormality property of the basis functions that actually provides this property.

For the bias result 1) of the theorem it is not required that the input signal be periodic, i.e., $r = 1$ also is allowed. The condition on the input signal is that it contains N_p different frequency components unequal to zero, while $N_p \to \infty$, and thus also the situation $N_p = N$ is allowed. The condition $r \to \infty$ in part 2) is a technical condition that is used in the proof. It is conjectured that in future work this condition can be shown to be superfluous.

In order to get further insight into the specific properties of the asymptotic covariance of the parameters, (29) is further analyzed in the next corollary.

Corollary 5.6: Consider the situation of Theorem 5.5-2). Then

$$\lim_{N,N_p,r\to\infty} N\,\text{cov}[\hat{L}_j, \hat{L}_k]$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{V}_0(e^{i\omega})\mathcal{V}_0^*(e^{i\omega})G_b^{j-k}(e^{i\omega})\frac{\Phi_v(\omega)}{\Phi_u(\omega)}\,d\omega \quad (30)$$

$$\lim_{N,N_p,r\to\infty} N\,\text{cov}[\hat{L}_j, \hat{D}] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{V}_0(e^{i\omega})G_b^j(e^{i\omega})\frac{\Phi_v(\omega)}{\Phi_u(\omega)}\,d\omega \quad (31)$$

and

$$\lim_{N,N_p\to\infty} N\,\text{var}[\hat{D}] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\Phi_v(\omega)}{\Phi_u(\omega)}\,d\omega. \quad (32)$$

It can be observed from this corollary that the estimated parameters become uncorrelated when $\Phi_v(\omega)/\Phi_u(\omega)$ is constant over frequency. In order to verify this, use has to be made of the property that the inner product between two distinct (orthonormal) basis functions will be zero. The situation that $\Phi_v(\omega)/\Phi(\omega) = 1$ refers to the case where the input signal over one period of u has the same spectral density as the noise and so does not imply that the noise signal is assumed to be periodic.

Note that the asymptotic covariance of the estimated parameters does not depend on the number of periods r in the input signal. Also, the asymptotic covariance does not depend on the number of estimated parameters $n_p = n_b n + 1$.

Remark 5.7: While in this section the analysis of the parameter estimate is performed for the system configuration (1) with output noise taken from a stationary stochastic process, it has to be noted that the results as formulated in Lemmas 5.3 and 5.4 can also be used to analyze the parameter estimate for different noise paradigms, e.g., the situation of bounded additive noise on the frequency response as considered in [11], [8], [27], and bounded additive errors on the output data as handled in [23], [37], and [10].

Remark 5.8: In recent work [19], [22] a slightly stronger result for the asymptotic parameter estimate is formulated as $\hat{\theta} \to \theta^*$ with probability one, with

$$\theta^* = \arg\min_\theta \frac{1}{2\pi}\int_{-\Omega}^{\Omega}|G_0(e^{i\omega}) - G(e^{i\omega},\theta)|^2 \cdot |W(e^{i\omega})|^2\,d\omega$$

with W a weighting function accounting for both frequency weighting $K(e^{i\omega})$ and a nonequidistant frequency grid. This result does not require the conditions of Assumption 5.2. We consider here the more restrictive case in order to fully exploit the orthonormality property of the basis functions when analyzing the covariance of the estimate. In this respect the more general situation would require basis functions that are orthogonal with respect to an inner product that accounts for the weighting W.

VI. ASYMPTOTIC ANALYSIS OF TRANSFER FUNCTION ESTIMATE

A. Finite-Order Case

Building on the results for the parameter estimates in the previous section, we will now take a closer look at the consequences for the estimated transfer function, i.e., its frequency response.

For the asymptotic bias we have the following result.

Corollary 6.1: Consider the situation of Theorem 5.5-1). Then for all $\omega \in [0, 2\pi)$

$$G_o(e^{i\omega}) - G(e^{i\omega}, \theta^*) = \sum_{k=n}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega}).$$

Proof: This result is direct from Theorem 5.5-1). □

Additionally, it can be shown that the amplitude of this asymptotic bias term can be bounded by an expression

$$\left| \sum_{k=n}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega}) \right| \leq \mathcal{K}^T |\mathcal{V}_0(e^{i\omega})| \frac{\eta^n}{1 - \eta} \qquad (33)$$

with $\mathcal{K} \in R^{n_b \times 1}$ and $1 > \eta \in R$, while η becomes smaller when the poles that are present in the all-pass function G_b approach the poles of the actual system G_o. More specifically, as analyzed in [34], η can be chosen any value $\eta > \lambda$, with

$$\lambda := \max_i \prod_{j=1}^{n_b} \left| \frac{\alpha_i - \rho_j}{1 - \alpha_i \rho_j} \right|$$

where α_i are the poles of G_0 and ρ_j are the poles of G_b. If the two sets of poles approach each other, then λ will tend to zero, and the series expansion (9) will have a high rate of convergence. In that situation the asymptotic bias in Corollary 6.1 will be accordingly small. For more details on this bound see [34].

It appears that the asymptotic model $G(e^{i\omega}, \theta^*)$ can be constructed from G_0 in a very special and explicit way by applying a frequency windowing operation to the system's frequency response. This mechanism which is introduced for the Laguerre case in [3], is formulated in the following proposition.

Proposition 6.2: Consider the situation of Theorem 5.5-1). Then the asymptotic transfer function model $G(e^{i\omega}, \theta^*)$ satisfies

$$G(e^{i\omega}, \theta^*) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{P}_n(\omega, \zeta) G_0(e^{i\zeta}) \, d\zeta \qquad (34)$$

with

$$\mathcal{P}_n(\omega, \zeta) := \phi^T(e^{i\omega}) \phi(e^{-i\zeta}) = 1 + \sum_{k=0}^{n-1} \mathcal{V}_k^T(e^{i\omega}) \mathcal{V}_k(e^{-i\zeta}). \qquad (35)$$

Interpreting this proposition, one can observe that the frequency domain identification approach as handled in this paper comes down to smoothing the frequency response of the plant by applying a frequency window that is explicitly given and that is dependent on the choice of basis functions and the choice of model order n.
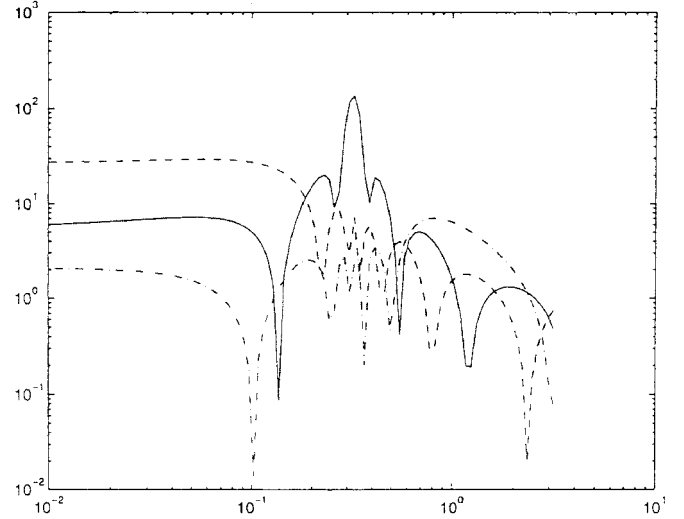

Fig. 1. Bode amplitude plot of $\mathcal{P}_n(\omega_1, \omega)$ as a function of ω for a basis induced by four poles: $0.9 \pm 0.3i$, $0.7 \pm 0.2i$, with $n = 3$ and evaluated at three values of ω_1, $\omega_1 = 0.32175$ rad/s (solid), $\omega_1 = 1$ rad/s (dash-dot), and $\omega_1 = 0.05$ rad/s (dashed).

A related, and actually similar interpretation, is given by the observation that in the asymptotic situation considered in the previous section, i.e., $N_p \to \infty$, the estimated frequency response $G(e^{i\omega}, \hat{\theta})$ converges in probability to a smoothed ETFE of the plant, where the smoothing operation is determined by the window given above

$$G(e^{i\omega}, \hat{\theta}) \to \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{P}_n(\zeta, \omega) \frac{Y_N(e^{i\zeta})}{U_N(e^{i\zeta})} \, d\zeta. \qquad (36)$$

Note that in contrast with the (real-valued) windows that are normally used in spectral analysis, the window $\mathcal{P}_n(\omega, \zeta)$ used here is complex valued, and moreover it is localized in frequency, i.e., it does not reflect a simple convolution in the frequency domain; the window is essentially dependent on ω and not only on $\omega - \zeta$.

By using the specific properties of the basis functions, $\mathcal{V}_k(e^{i\omega}) = \mathcal{V}_0(e^{i\omega}) G_b^k(e^{i\omega})$ with G_b being a stable all-pass function, and using the summation rule for finite power series, it follows that for $\omega \neq \zeta$

$$\mathcal{P}_n(\omega, \zeta) = 1 + \mathcal{V}_0^T(e^{i\omega}) \mathcal{V}_0(e^{-i\zeta}) \frac{\left(\frac{G_b(e^{i\omega})}{G_b(e^{i\zeta})} \right)^n - 1}{\frac{G_b(e^{i\omega})}{G_b(e^{i\zeta})} - 1} \qquad (37)$$

while

$$\mathcal{P}_n(\omega, \omega) = 1 + n \|\mathcal{V}_0(e^{i\omega})\|_2^2. \qquad (38)$$

This frequency window is illustrated in two figures, where an example has been chosen of an orthonormal basis induced by a scalar fourth-order all-pass function G_b having poles $0.9 \pm 0.3i$ and $0.7 \pm 0.2i$. In Fig. 1 the amplitude of $\mathcal{P}_n(\omega_1, \omega)$ is given as a function of ω for $n = 3$ and for three different values of ω_1. The first value $\omega_1 = 0.32175$ is the frequency of the resonant mode determined by the first-mentioned pole-pair. Fig. 2 shows the amplitude of $\mathcal{P}_n(\omega_1, \omega)$ for ω_1 fixed to $\omega_1 = 0.32175$ and for $n = 3, 20$.
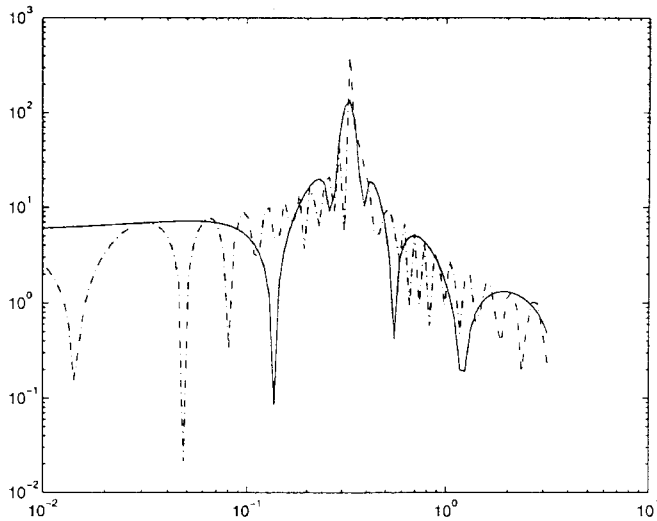
Fig. 2. Bode amplitude plot of $\mathcal{P}_n(\omega_1, \omega)$ as a function of ω for the same basis as in Fig. 1, $\omega_1 = 0.321\,75$ rad/s, and evaluated for $n = 3$ (solid) and $n = 20$ (dash-dot).

From Fig. 1 one can clearly observe the ω_1-dependence of the frequency window leading to a smoothing operation that is nonfixed over the frequency axis. Fig. 2 illustrates the effect that for increasing n the window will become more concentrated around $\omega = \omega_1$, finally tending toward a dirac-function as $n \to \infty$.

The smoothing operation in the frequency domain that is actually performed by the considered identification setup can now also be used in order to tune the dynamics to be chosen in the basis functions, i.e., the poles in the all-pass function G_b. These dynamics will determine the character of the smoothing operation for a given number of parameters n_p to be estimated. However, explicit design rules in view of this phenomenon have not been formulated yet.

For the covariance of the frequency response estimate we have the following result.

Proposition 6.3: Consider Assumptions 4.2 and 5.2 and the estimated transfer function (22). Let $\omega_1, \omega_2 \in [0, 2\pi)$. Then

$$\lim_{N, N_p, r \to \infty} \frac{N}{n_p} \text{cov}[\hat{G}(e^{i\omega_1}), \hat{G}(e^{i\omega_2})]$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{n_p} \mathcal{P}_n(\omega_1, \zeta) \mathcal{P}_n^*(\omega_2, \zeta) \frac{\Phi_v(\zeta)}{\Phi_u(\zeta)} \, d\zeta \quad (39)$$

with $\mathcal{P}_n(\omega, \zeta)$ as defined in (35).

Before discussing this result, we will first present the following corollary.

Corollary 6.4: Consider the situation of Theorem 6.3 and in addition let $\Phi_v(\zeta)/\Phi_u(\zeta) = 1$ for all $\zeta \in [0, 2\pi)$. Then

$$\lim_{N, N_p, r \to \infty} \frac{N}{n_p} \text{cov}[\hat{G}(e^{i\omega_1}), \hat{G}(e^{i\omega_2})] = \frac{1}{n_p} \mathcal{P}_n(\omega_1, \omega_2). \quad (40)$$

Theorem 6.3 provides an expression for the covariance of the estimated transfer function for the situation of a finite model order. At first sight (39) may seem a bit complicated; however, one has to realize that when given a basis generating system G_b and a model order n, the window (35) can be calculated directly. The fact that an accessible expression is

provided for the variance, which also holds for finite model orders and in a situation of undermodeling, is opposed to the situation for time-domain prediction error methods; see, e.g., [16], [20], [17], and [36]. However, in order to be able to establish these expressions, we need a periodic input signal instead of a quasistationary one, which is required in the time domain identification case. Note that (39) shows that the asymptotic covariance of the transfer function estimate does not depend on the number of periods r in the input signal and that the asymptotic covariance between transfer function estimates at different frequencies is nonzero for finite model orders.

Corollary 6.4 shows that the variance of the transfer function estimate will be relatively large where $\|\mathcal{V}_0(e^{i\omega})\|_2^2$ is large (provided that $\Phi_v(\zeta)/\Phi_u(\zeta)$ is not relatively small for those frequencies). Thus, in general the variance will be relatively large near the poles of the basis generating system and hence near the poles of the system itself for a properly chosen basis generating system (see [12] and [34]). Additionally, noise at frequencies not situated near the poles of the basis generating system (e.g., high-frequency noise) usually will be relatively harmless.

The frequency region (with regard to ζ and a fixed ω) where $\mathcal{P}_n(\omega, \zeta)$ is relatively large becomes more concentrated around $\zeta = \omega$ when the model order n increases. If $\Phi_v(\omega)/\Phi_u(\omega)$ is relatively small, then an increase of the model order can very well lead to a reduction of the variance in this frequency region. It follows that, up to a certain model order, variance considerations do not necessarily conflict with bias considerations for those frequency regions where $\Phi_v(\zeta)/\Phi_u(\zeta)$ is relatively small. This is a somewhat surprising and (to the authors' best knowledge) new result. The existing literature only addresses the situation where the model order n goes to infinity also. In that case the variance always increases with the model order; see, e.g., [17] and Theorem 6.8 below.

Remark 6.5: The results obtained in this and the previous section are in accordance with and actually are further generalizations of the results obtained in [3], where a similar identification problem is analyzed for an integral identification criterion $(N_p \to \infty)$ and Laguerre basis functions, where no undermodeling is considered. In [3] the frequency smoothing by orthogonal basis functions modeling is introduced.

B. Infinite-Order Case

Similar to the situation for the time-domain identification of finite expansion models, we can arrive at simple expressions for the asymptotic variance of the estimated parameters and transfer function in the situation that the model order n goes to infinity also.

In order to formulate this result we will build upon (30), and we will show that an asymptotic-in-order result can be formulated in a simple way, utilizing an orthonormal transformation that is induced by the basis functions that are applied. This transformation is presented in [34] and further elaborated in [13].

Definition 6.6 (Hambo Transform [34]): Let $\Phi(w)$ be a rational spectral density function with stable spectral factor
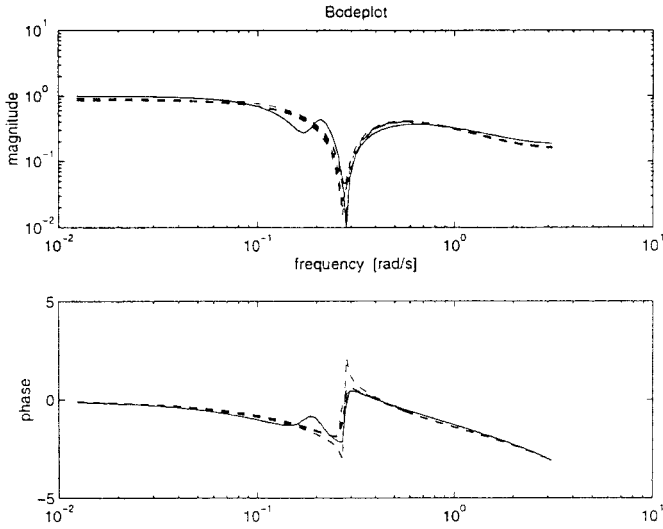
Fig. 3. Bodeplot of G_0 (solid) and estimates using orthogonal basis functions for different realizations of the noise (dashed); number of parameters (n_p) is five.

$H(e^{i\omega})$, and let G_b be a scalar stable all-pass function with balanced realization (A, B, C, D) that induces an orthonormal basis $\mathcal{V}_k(z)$ for $\mathcal{H}_2$. Then the Hambo transformed spectrum $\tilde{\Phi}(\omega)$ is defined by

$$\tilde{\Phi}(\omega) := \tilde{H}^T(e^{-i\omega})\tilde{H}(e^{i\omega})$$

with

$$\tilde{H}(\lambda) = H(z)|_{z^{-1}=N(\lambda)}$$

and

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

When applying this orthonormal transformation to the result (30), an analysis similar to the one applied in the time-domain case [34] leads to a formulation for the spectral density function that generates the asymptotic covariance matrix R (29), composed of block-elements given by (30).

Lemma 6.7: Consider the covariance matrix $\bar{R}$ being composed of block elements $\bar{R}_{j+1,k+1} = N \operatorname{cov}[\hat{L}_j, \hat{L}_k]$ with $j, k = 0, \cdots n - 1$, as given in (30).

Then $\bar{R}_{j+1,k+1} = F(j - k)$, with $F(\tau), \tau \in Z$, the (matrix) covariance function related to the $n_b \times n_b$ spectral density function $\tilde{\Phi}_v(\omega)\tilde{\Phi}_u(\omega)^{-1}$.

Note that the matrix $\bar{R}$ is obtained from R by discarding the first row and column. In this way a block-symmetric matrix is obtained, limiting attention to the vector-valued parameters $\hat{L}_j$ and not addressing the scalar direct feedthrough $\hat{D}$.

The representation of the asymptotic covariance matrix in this spectral form enables the formulation of the asymptotic-in-order result for the asymptotic variance of the estimated transfer function, as formulated next. This result is similar to the related time-domain identification result in [34].

Theorem 6.8: Consider Assumptions 4.2 and 5.2 and the estimated transfer function (22). Let $\omega_1, \omega_2 \in [0, 2\pi)$. Then we have

$$\lim_{\substack{N,N_p,r,n\to\infty \\ \frac{n_p n}{N_p} \to 0}} \frac{N}{n_p} \operatorname{cov}[\hat{G}(e^{i\omega_1}) - \hat{D}, \hat{G}(e^{i\omega_2}) - \hat{D}]$$

$$= \begin{cases} 0, & \text{for } G_b(e^{i\omega_1}) \neq G_b(e^{i\omega_2}) \\ \frac{1}{n_b} V_0^T(e^{i\omega_1})V_0(e^{-i\omega_1})\frac{\Phi_v(\omega_1)}{\Phi_u(\omega_1)}, & \text{for } \omega_1 = \omega_2. \end{cases}$$

The asymptotic-in-order result shows a very simple expression for the asymptotic variance of the estimated transfer function, where the asymptotic variance at a specific frequency is equal to a frequency-weighted noise-to-signal ratio at that frequency. The frequency weighting is completely determined by the basis functions that are used in our model structure. This result exactly matches the equivalent form that has been derived for time-domain identification [34]. However, in the current situation a periodic input signal is required instead of a general quasistationary signal.

VII. APPLICATION

In order to illustrate the presented approach, the resulting identification algorithm will be applied to data from a simulation example and to data that is taken from the mechanical servosystem in a Compact Disc drive.

Simulation Example: The fifth-order system is simulated in (41), as shown at the bottom of the page, having pole locations: $0.9509 \pm 0.1982i$, $0.8497 \pm 0.0964i$, and 0.5490. The input signal is periodic with a period length of 256 points, where the input over one period is white noise. There is 50% (in standard deviation) colored noise $v(t) = H_o(q)e(t)$ on the output, where the noise filter is given by

$$H_o(z) = \frac{1 - 1.38z^{-1} + 0.40z^{-2}}{1 - 1.90z^{-1} + 0.91z^{-2}} \tag{42}$$

and $e(t)$ is normally distributed white noise.

The identification procedure presented in this paper is applied to data from this system. Basis functions are chosen to be generated by the pole locations 0.9, $0.8 \pm 0.2i$, 0.7. The estimates are made over five datasets with different noise realizations, while each dataset consists of 2048 datapoints. No particular frequency weighting is considered.

In Fig. 3 the results are shown for estimated models with five parameters $(n_p = 5)$ and in Fig. 4 for $n_p = 13$.

As a reference, Fig. 5 shows the results that are obtained when applying a fifth-order output error (OE) time-domain identification procedure, having a correct model structure to model the plant G_0 exactly (ten parameters).

It is observed that for a choice of basis poles that can hardly be considered to be very accurate, the identification results for the orthogonal basis functions method are good; they improve from the five-parameter situation to the 13-parameter situation

$$G_o(z) = \frac{0.2530z^{-1} - 0.9724z^{-2} + 1.4283z^{-3} - 0.9493z^{-4} + 0.2410z^{-5}}{1 - 4.1500z^{-1} + 6.8831z^{-2} - 5.6871z^{-3} + 2.3333z^{-4} - 0.3787z^{-5}} \tag{41}$$
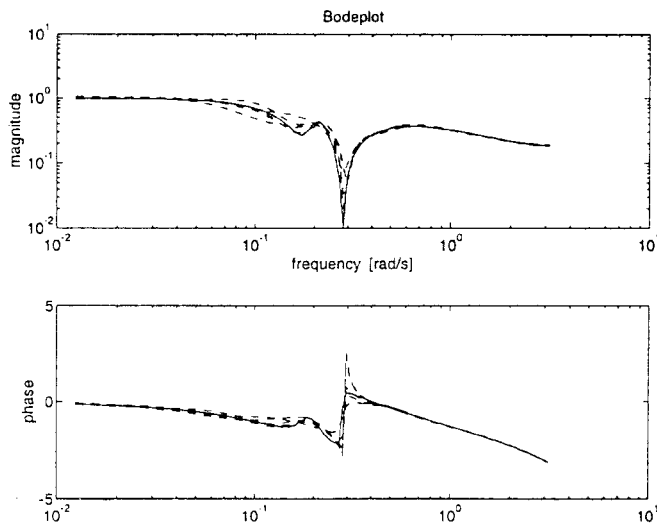
Fig. 4. Bodeplot of G_0 (solid) and estimates using orthogonal basis functions for different realizations of the noise (dashed); number of parameters (n_p) is 13.
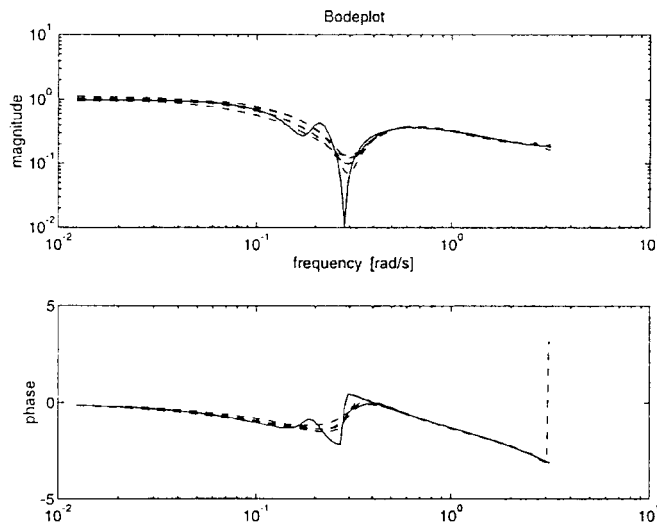

Fig. 5. Bodeplot of G_0 (solid) and estimates using a fifth-order OE model for different realizations of the noise (dashed); number of parameters is ten.

and are at least as good as the results of a standard OE method. Moreover, the orthogonal basis functions method relies only on a linear regression type of algorithm, while the OE estimate requires a nonlinear optimization procedure.

Identification of the Radial Servosystem in a Compact-Disc Drive: The introduced procedure has also been applied to measured data from the radial servo system in the optical device of a Compact-Disc player. This system is operating in closed loop with a stabilizing controller, and for this example the identification of the sensitivity function of this loop will be considered. To this end a reference signal consisting of a Schroeder-phased multisine ($N_p = 397$) has been applied with eight periods of 1024 samples each, with a frequency grid from 100 Hz until 10 kHz. The sampling frequency is 25 kHz.

The poles of the basis functions have been chosen on the basis of a curve-fit of the frequency response measured by a frequency-analyzer [35]. Both eighth- and tenth-order basis
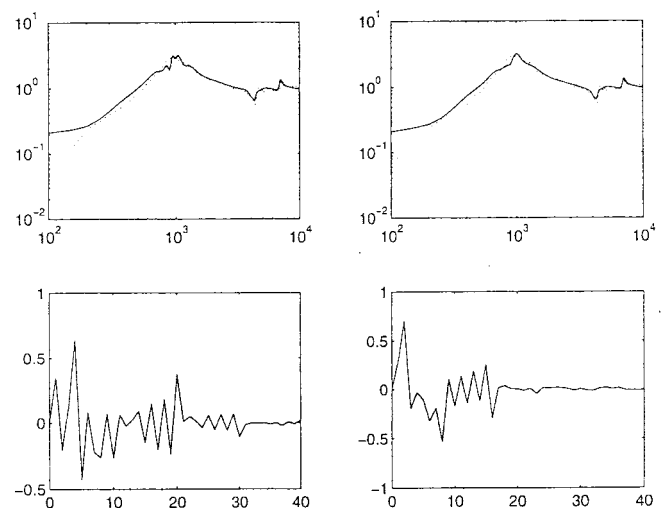

Fig. 6. Upper figures: Amplitude Bode plot of estimated model (solid) and result of frequency analyzer (dotted) for $n_p = 40$, tenth-order basis (left) and eighth-order basis (right). Lower figures: estimated coefficients $\hat{\theta}_i$, $i = 1, \cdots 40$, for tenth-order basis (left) and eighth-order basis (right).

functions have been applied.[2] Fig. 6 shows the results of identified models with $n_p = 40$.

The estimated coefficients show that the convergence of the sequence is dependent on the basis chosen. In the given example the eighth-order basis leads to a convergence that is faster than the tenth-order basis. Both estimated models provide good results; however, the model with the eighth order basis is preferable as it uses less parameters—not more than 20 parameters contribute essentially to the model dynamics. The less accurate fit in the low-frequency region (100–250 Hz) is caused by a considerable noise power that occurs in this frequency region.

VIII. CONCLUSION

We have obtained asymptotic bias and variance expressions for the estimated parameters and transfer function in an identification problem that is based on time-domain data and an identification criterion that is formulated in the frequency domain. First the situation is considered of infinite time-domain data and a criterion over a finite number of frequencies. Additionally, the number of frequencies is taken to tend to infinity also. As a model structure, a linear parameterization in terms of recently introduced generalized basis functions is employed, being a generalization of the classical pulse and Laguerre- and Kautz-type bases.

For the infinite-frequency criterion, it is shown that the parameterized part of the model can be estimated consistently, despite of the presence of neglected dynamics (undermodeling). The approximate model is shown to be equivalent to a model obtained by frequency-smoothing the system's frequency response by a frequency-localized window. Variance expressions are derived for situations of finite and infinite order models.

---

[2] The authors are indebted to E. van Donkelaar for performing the appropriate experiments.

The identification procedure presented here matches the identification setup that is used in [4] and [5] for identification of model uncertainty. The results presented here can therefore be used to guide the crucial design choices to be made in order to achieve minimal uncertainty bounds.

APPENDIX

Lemma 8.1: Let X be a complex function that is differentiable on the unit circle. Then for $N_p \to \infty$

$$\left| \frac{1}{N_p} \sum_{\ell=1}^{N_p} X(e^{i\frac{2\pi\ell}{N_p}}) - \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{i\omega}) \, d\omega \right|$$
$$\leq \frac{\sqrt{2}\pi}{N_p} \cdot \max_{\omega} |X'(e^{i\omega})| \qquad (A.1)$$

where $X'(e^{i\omega}) := dX(e^{i\omega})/d\omega$.

Proof: First we decompose $X(e^{i\omega})$ into its real and imaginary part, $X(e^{i\omega}) = X_r(e^{i\omega}) + iX_{\mathrm{im}}(e^{i\omega})$, and we will bound the expression on the left-hand side of (A.1) for X_r and X_{im} separately

$$\left| \frac{1}{N_p} \sum_{\ell=1}^{N_p} X_r(e^{i\frac{2\pi\ell}{N_p}}) - \frac{1}{2\pi} \int_{-\pi}^{\pi} X_r(e^{i\omega}) \, d\omega \right|$$
$$= \left| \sum_{\ell=1}^{N_p} \left(\frac{1}{N_p} X_r(e^{i\frac{2\pi\ell}{N_p}}) - \frac{1}{2\pi} \int_{\frac{\pi(2\ell-1)}{N_p}}^{\frac{\pi(2\ell+1)}{N_p}} X_r(e^{i\omega}) d\omega \right) \right|$$
$$\leq \frac{1}{N_p} \sum_{\ell=1}^{N_p} \max_{\pi(2\ell-1)/N_p \leq \zeta_\ell \leq \pi(2\ell+1)/N_p} \left| X_r(e^{i\frac{2\pi\ell}{N_p}}) - X_r(e^{i\zeta_\ell}) \right|$$
$$\leq \frac{1}{N_p} \sum_{\ell=1}^{N_p} \frac{\pi}{N_p} \max_{\pi(2\ell-1)/N_p \leq \zeta_\ell \leq \pi(2\ell+1)/N_p} |X_r'(e^{i\zeta_\ell})|$$
$$\leq \frac{\pi}{N_p} \max_{\omega} |X_r'(e^{i\omega})| \leq \frac{\pi}{N_p} \max_{\omega} |X'(e^{i\omega})|.$$

As a similar result holds for $X_{\mathrm{im}}(e^{i\omega})$, the combination of bounding the real and imaginary parts separately leads to the result of the Lemma. □

Proof of Lemma 4.1: First we will consider the case $n = 1$. In this situation we evaluate under which conditions the matrix

$$\Phi^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathcal{V}_0(e^{i\omega_1}) & \mathcal{V}_0(e^{i\omega_2}) & \cdots & \mathcal{V}_0(e^{i\omega_{N_p}}) \end{bmatrix} \qquad (A.2)$$

has full row rank. Now consider a real-valued vector $[\delta \; \gamma^T]$ with $\gamma \in R^{n_b}$ and $\delta \in R$, and consider the dynamical system $H(z)$ being defined by the state-space realization (A, B, γ^T, δ). Then it can simply be verified using the structure of $\mathcal{V}_0$ that the equation

$$[\delta \quad \gamma^T]\Phi^T = 0 \qquad (A.3)$$

is equivalent to $H(e^{i\omega_k}) = 0$ for $k = 1, \cdots, N_p$. Since H has McMillan degree smaller than or equal to n_b, it can maximally have n_b finite zeros. If $N_p > n_b$, it follows that (A.3) implies $[\delta \; \gamma^T] = 0$ which proves the full row rank property of Φ^T.

The situation $n > 1$ can simply be handled by employing [12, Proposition 5.4], which shows that for any $n > 1$ there exists an input-balanced pair of matrices A_n, B_n with $\tilde{\mathcal{V}}_0(z) := (zI - A_n)^{-1} B_n$ such that

$$\tilde{\mathcal{V}}_0(z) = \begin{bmatrix} \mathcal{V}_0(z) \\ \mathcal{V}_1(z) \\ \vdots \\ \mathcal{V}_{n-1}(z) \end{bmatrix}. \qquad (A.4)$$

This implies that we can apply the same analysis as for the case $n = 1$, now using zeros of the system $H_n(z)$ with realization $(A_n, B_n, \gamma^T, \delta)$ with $\gamma \in R^{n_b n}$, which has McMillan degree smaller than or equal to $n_b n$.

Since Λ is a diagonal matrix with positive elements, the full row rank property of Φ^T is sufficient to guarantee that $\Phi^T \Lambda \Phi$ is nonsingular. □

Proof of Lemma 4.3: Part 1) is a direct result of [2, Th. 4.4.1].

For part 2) results in [17, eqs. (6.30), (6.31)] can be utilized to show that

$$\left| \mathbb{E}\left[|V_N(e^{i\omega_k})|^2\right] - \Phi_v(\omega_k) \right| \leq \frac{c_1}{N}$$

and

$$\left| \mathbb{E}[V_N(e^{i\omega_k}) V_N(e^{i\omega_\ell})^*] \right| \leq \frac{c_1}{N} \quad \text{for} \quad \omega_k \neq \omega_\ell.$$

With (24) the result follows directly. The norm bound follows by noting that the induced two-norm of a matrix is overbounded by its Frobenius norm. □

Proof of Proposition 5.1: From (20) it can be verified that for $N \to \infty$, $\mathbb{E}[\hat{\theta}] = \theta_o + \Psi Z$. This is caused by the fact that for $N \to \infty$ $\Psi S = 0$, as the effect of initial conditions can be written as $S(e^{i\omega_k}) = R_N(e^{i\omega_k})/U_N(e^{i\omega_k})$, while $R_N(e^{i\omega_k})$ decays as $1/\sqrt{N}$; see, e.g., [17].

Asymptotic normality of the parameter estimate follows from the fact that for $N \to \infty$ the Discrete Fourier Transform $V_N(e^{i\omega_k})$ becomes normally distributed (see [2] and [4, Th. 5.3.2.]) and so does $F(e^{i\omega_k})$.

The asymptotic covariance matrix follows as a direct result of Lemma 4.3 and (21). Since the asymptotic parameter covariance $N \operatorname{cov}(\hat{\theta})$ is bounded, $\hat{\theta}$ converges to θ^* in mean square sense and thus also in probability. □

Proof of Lemma 5.3: We can write $\frac{1}{N_p} \Phi^* \Phi = \frac{1}{N_p} \sum_{\ell=1}^{N_p} \phi(e^{-i\omega_\ell}) \phi^T(e^{i\omega_\ell})$. It follows straightforwardly from the orthonormality of the basis functions that $\frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(e^{-i\omega_\ell}) \phi^T(e^{i\omega_\ell}) = I$. The result can now be proven by applying Lemma 8.1 to all the scalar elements of the matrix $\phi(e^{-i\omega_\ell}) \phi^T(e^{i\omega_\ell})$. To this end, we write

$$\phi(e^{-i\omega_\ell}) \phi^T(e^{i\omega_\ell}) = \begin{bmatrix} 1 & f^* \\ f & P \end{bmatrix}$$

with $f = [\mathcal{V}_0^T(e^{-i\omega_\ell}) \; \cdots \; \mathcal{V}_{n-1}^T(e^{-i\omega_\ell})]^T$, and the $n_b \times n_b$ block element P_{qs} of P defined by

$$P_{\mathrm{qs}} = \mathcal{V}_q(e^{-i\omega_\ell}) \mathcal{V}_s^T(e^{i\omega_\ell}).$$

Since for all $q \leq n - 1$ the scalar elements $\left| \frac{d\mathcal{V}_q(e^{i\omega})_i}{d\omega} \right|$, $i = 1, \cdots, n_b$ are uniformly bounded, this shows that Lemma 8.1 can be successfully applied to the first row and column of $\Phi^* \Phi$.

For the scalar (i, j)-elements in P_{qs} it follows that

$$
\left| \frac{d}{d\omega} [\mathcal{V}_q(e^{-i\omega}) \mathcal{V}_s^T(e^{i\omega})]_{ij} \right|
$$

$$
= \left| \frac{d}{d\omega} [\mathcal{V}_0(e^{-i\omega}) \mathcal{V}_0^T(e^{i\omega})]_{ij} G_b^{s-q}(e^{i\omega}) \right|
$$

$$
\leq \left| \frac{d}{d\omega} [\mathcal{V}_0(e^{-i\omega}) \mathcal{V}_0^T(e^{i\omega})]_{ij} \right| |G_b(e^{i\omega})|^{s-q}
$$

$$
+ |[\mathcal{V}_0(e^{-i\omega}) \mathcal{V}_0^T(e^{i\omega})]_{ij}| \cdot |s - q| \cdot |G_b(e^{i\omega})|^{s-q-1}
$$

$$
\cdot \left| \frac{d}{d\omega} G_b(e^{i\omega}) \right|.
$$

Since the elements of $\mathcal{V}_0$ as well as G_b are in $\mathcal{RH}_2$, and $|G_b(e^{i\omega})| = 1$, all terms in this inequality are uniformly bounded, except for $|s - q|$. However, as $|s - q| \leq n - 1$ for $n \geq 1$ it follows that for all elements of the matrix $\phi(e^{-i\omega_\ell}) \phi^T(e^{i\omega_\ell})$ there exists a constant $c \in R$ such that $\max_\omega |X'(e^{i\omega})| \leq c \cdot n$. Application of Lemma 8.1 then proves the result.

The second part of the lemma follows from the element-wise convergence, together with the fact that the matrix has finite dimension n_p and that the induced two-norm of a matrix satisfies

$$
\|A\|_2 \leq \sqrt{\sum_i \sum_k |A_{ik}|^2}. \qquad \square
$$

Proof of Lemma 5.4: From simple manipulations it follows that

$$
\left\| \Psi - \frac{1}{N_p} \Phi^* \right\|_2 = \left\| \left[I - \left(\frac{1}{N_p} \Phi^* \Phi \right)^{-1} \right] \frac{1}{N_p} \Phi^* \right\|_2
$$

$$
\leq \left\| I - \left(\frac{1}{N_p} \Phi^* \Phi \right)^{-1} \right\|_2 \cdot \left\| \frac{1}{N_p} \Phi^* \right\|_2. \quad \text{(A.5)}
$$

Next we will analyze the two separate components in this expression.

For a matrix X satisfying $\|I - X\|_2 < 1$ it holds (see, e.g., [7, p. 59]) that

$$
\|I - X^{-1}\|_2 \leq \frac{\|I - X\|_2}{1 - \|I - X\|_2}.
$$

Using Lemma 5.3 it follows that for $c_4 \frac{n_p n}{N_p} < 1$

$$
\left\| I - \left(\frac{1}{N_p} \Phi^* \Phi \right)^{-1} \right\|_2 \leq \frac{c_4 \frac{n_p n}{N_p}}{1 - c_4 \frac{n_p n}{N_p}}. \quad \text{(A.6)}
$$

For $n_p n / N_p \to 0$ this reduces to

$$
\left\| I - \left(\frac{1}{N_p} \Phi^* \Phi \right)^{-1} \right\|_2 \leq c_4 \frac{n_p n}{N_p}. \quad \text{(A.7)}
$$

For the second term in (A.5) we write

$$
\left\| \frac{1}{N_p} \Phi \right\|_2 = \sqrt{\frac{1}{N_p} \left\| \frac{1}{N_p} \Phi^* \Phi \right\|_2}
$$

$$
= \sqrt{\frac{1}{N_p} \left\| I - \left(I - \frac{1}{N_p} \Phi^* \Phi \right) \right\|_2}
$$

$$
\leq \sqrt{\frac{1}{N_p} \left(\|I\|_2 + \left\| I - \frac{1}{N_p} \Phi^* \Phi \right\|_2 \right)}
$$

$$
\leq \sqrt{\frac{1}{N_p} \left(1 + c_4 \frac{n_p n}{N_p} \right)}. \quad \text{(A.8)}
$$

Combining (A.8) with (A.7) now proves the result. $\square$

Proof of Theorem 5.5—Part 1): Referring to the Proof of Proposition 5.1 it follows that $\Psi \mathsf{S} = 0$ as $n_p n / N \to 0$. Thus we have to show that $\Psi \mathsf{Z} \to 0$ for $n_p n / N_p \to 0$. This proof will be made in two steps. First we will show that $\Psi \mathsf{Z} \to \frac{1}{N_p} \Phi^* \mathsf{Z}$, and in the second step we will prove that this latter expression tends to zero in the matrix two-norm.

For analyzing Z we find from (33) and resulting from [12] and [34] that

$$
\sum_{k=n}^{\infty} |L_k| \leq \mathcal{K} \frac{\eta^n}{1 - \eta}. \quad \text{(A.9)}
$$

The two-norm of the vector Z now can be bounded as

$$
\|\mathsf{Z}\|_2 = \sqrt{\sum_{\ell=1}^{N_p} \left| \sum_{k=n}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega_\ell}) \right|^2}
$$

$$
\leq \sqrt{\sum_{\ell=1}^{N_p} \left(\mathcal{K}^T |\mathcal{V}_0(e^{i\omega_\ell})| \frac{\eta^n}{1 - \eta} \right)^2}
$$

$$
\leq \sqrt{N_p} \frac{\eta^n}{1 - \eta} |\mathcal{K}^T V_\infty| \quad \text{(A.10)}
$$

with $V_\infty \in R^{n_b \times 1}$ the vector of $\mathcal{H}_\infty$-norms of the scalar elements in $\mathcal{V}_0$, and where it should be noted that $\mathcal{K} \in R^{n_b \times 1}$ and that n_b is assumed to be fixed and finite. This result, together with Lemma 5.4, shows that

$$
\left\| \left[\Psi - \frac{1}{N_p} \Phi^* \right] \mathsf{Z} \right\|_2 \leq \left\| \Psi - \frac{1}{N_p} \Phi^* \right\|_2 \cdot \|\mathsf{Z}\|_2 \leq c_6 \frac{n_p n}{N_p}. \quad \text{(A.11)}
$$

Next we have to analyze

$$
\frac{1}{N_p} \Phi^* \mathsf{Z}
$$

$$
= \frac{1}{N_p} \begin{bmatrix} \sum_{\ell=1}^{N_p} \sum_{k=n}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega_\ell}) \\ \sum_{\ell=1}^{N_p} \mathcal{V}_0(e^{-i\omega_\ell}) \sum_{k=n}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega_\ell}) \\ \vdots \\ \sum_{\ell=1}^{N_p} \mathcal{V}_{n-1}(e^{-i\omega_\ell}) \sum_{k=n}^{\infty} L_k \mathcal{V}_k(e^{i\omega_\ell}) \end{bmatrix}
$$

$$
= \frac{1}{N_p} \begin{bmatrix} \sum_{k=n}^{\infty} L_k^T \sum_{\ell=1}^{N_p} \mathcal{V}_k(e^{i\omega_\ell}) \\ \sum_{k=n}^{\infty} \left(\sum_{\ell=1}^{N_p} \mathcal{V}_0(e^{-i\omega_\ell}) \mathcal{V}_k^T(e^{i\omega_\ell}) \right) L_k \\ \vdots \\ \sum_{k=n}^{\infty} \left(\sum_{\ell=1}^{N_p} \mathcal{V}_{n-1}(e^{-i\omega_\ell}) \mathcal{V}_k^T(e^{i\omega_\ell}) \right) L_k \end{bmatrix}.
$$

As in the above summations, k is not bounded away from N_p, and the conditions for applying Lemma 8.1 are not

satisfied. To this end we split the above expression by way of introducing $Z = Z_1 + Z_2$ with

$$Z_1(e^{i\omega_\ell}) = \sum_{k=n}^{\sqrt{N_p}} L_k^T \mathcal{V}_k(e^{i\omega_\ell})$$

$$Z_2(e^{i\omega_\ell}) = \sum_{k=\sqrt{N_p}+1}^{\infty} L_k^T \mathcal{V}_k(e^{i\omega_\ell})$$

where, for ease of notation, we have assumed that $\sqrt{N_p}$ is an integer. If this is not the case $\sqrt{N_p}$ should be rounded to the nearest integer.

From the proof of Lemma 5.3 it follows that there exist constants $C_1 \in R^{n_b \times 1}$ and $C_2 \in R^{n_b \times n_b}$ such that

$$\left| \frac{1}{N_p} \sum_{\ell=1}^{N_p} \mathcal{V}_s(e^{i\omega_\ell}) \right| \le \frac{C_1}{\sqrt{N_p}} \quad \text{for } s \in N, \ s \le \sqrt{N_p}$$

$$\left| \frac{1}{N_p} \sum_{\ell=1}^{N_p} \mathcal{V}_r(e^{-j\omega_\ell}) \mathcal{V}_s^T(e^{i\omega_\ell}) \right|$$
$$\le \frac{C_2}{\sqrt{N_p}} \quad \text{for } r, s \in N, \ r < s \le \sqrt{N_p}.$$

Since there exists a real-valued η, $0 < \eta < 1$, and $C_3 \in R^{n_b \times 1}$ such that for $k \ge n$, $|L_k| \le C_3 \eta^n$, it can be shown that

$$\left\| \frac{1}{N_p} \Phi^* Z_1 \right\|_2 \le \sqrt{n_p} \frac{\eta^n}{\sqrt{N_p}} c_7. \tag{A.12}$$

For analyzing $\frac{1}{N_p} \Phi^* Z_2$, we write

$$\frac{1}{N_p} \Phi^* Z_2$$
$$= \frac{1}{N_p} \begin{bmatrix} \sum_{k=\sqrt{N_p}+1}^{\infty} \sum_{\ell=1}^{N_p} L_k^T \mathcal{V}_k(e^{i\omega_\ell}) \\ \sum_{k=\sqrt{N_p}+1}^{\infty} \sum_{\ell=1}^{N_p} \mathcal{V}_0(e^{-i\omega_\ell}) \mathcal{V}_k^T(e^{i\omega_\ell}) L_k \\ \vdots \\ \sum_{k=\sqrt{N_p}+1}^{\infty} \sum_{\ell=1}^{N_p} \mathcal{V}_{n-1}(e^{-i\omega_\ell}) \mathcal{V}_k^T(e^{i\omega_\ell}) L_k \end{bmatrix}.$$

Using the fact that $|\mathcal{V}_k(e^{i\omega})| = |\mathcal{V}_0(e^{i\omega})||G_b^k(e^{i\omega})| = |\mathcal{V}_0(e^{i\omega})|$, and denoting $V_\infty \in R^{n_b \times 1}$ to be the vector of $\mathcal{H}_\infty$-norms of the scalar entries in $\mathcal{V}_0$, and employing (A.9) shows that

$$\left| \frac{1}{N_p} \Phi^* Z_2 \right| \le \frac{\eta^{\sqrt{N_p}+1}}{1-\eta} \begin{bmatrix} \mathcal{K}^T V_\infty \\ V_\infty V_\infty^T \mathcal{K} \\ \vdots \\ V_\infty V_\infty^T \mathcal{K} \end{bmatrix}. \tag{A.13}$$

For the two-norm of this vector it follows that there exists a $c_{12} \in R$ such that

$$\left\| \frac{1}{N_p} \Phi^* Z_2 \right\|_2 \le \sqrt{n_p} \eta^{\sqrt{N_p}} c_{12}. \tag{A.14}$$

As $\|\Phi^* Z\|_2 \le \|\Phi^* Z_1\|_2 + \|\Phi^* Z_2\|_2$, the result of Part 1) of the theorem now follows by combining (A.12) and (A.14).

Part 2): We will prove the correctness of the covariance matrix by showing that in the limit situation, as formulated in this part of the theorem, it follows that

$$\left\| N \operatorname{cov}[\hat{\theta}] - \frac{1}{N_p} \Phi^* Q \Phi \right\|_2 \to 0. \tag{A.15}$$

Using (21) we find

$$\left\| N \operatorname{cov}[\hat{\theta}] - \frac{1}{N_p} \Phi^* Q \Phi \right\|_2$$
$$= \left\| N (\Phi^* \Phi)^{-1} \Phi^* \, \mathbb{E}[FF^*]((\Phi^* \Phi)^{-1} \Phi^*)^* - \frac{1}{N_p} \Phi^* Q \Phi \right\|_2$$
$$= \left\| \frac{1}{N_p} \left(I - \left(I - \left(\frac{1}{N_p} \Phi^* \Phi \right)^{-1} \right) \right) \Phi^* \right.$$
$$\times \left(Q - \left(Q - \frac{N}{N_p} \mathbb{E}[FF^*] \right) \right)$$
$$\left. \cdot \Phi \left(I - \left(I - \left(\frac{1}{N_p} \Phi^* \Phi \right)^{-1} \right) \right)^* - \frac{1}{N_p} \Phi^* Q \Phi \right\|_2. \tag{A.16}$$

Straightforward elaboration of (A.16), using (26), (A.7), and (A.8) now gives (A.15).

Additionally

$$\|Q\|_2 = \left\| \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \right\|_\infty$$

which together with (A.8) shows that $\|\frac{1}{N_p} \Phi^* Q \Phi\|_2$ converges to a finite value. $\square$

Proof of Corollary 5.6: The result follows directly from Theorem 5.5 and Lemma 8.1. $\square$

Proof of Proposition 6.2: Because of the orthonormality of the basis functions, it follows that

$$L_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{V}_k(e^{-i\omega}) G_0(e^{i\omega}) \, d\omega$$

$$D = \frac{1}{2\pi} \int_{-\pi}^{\pi} G_0(e^{i\omega}) \, d\omega.$$

As

$$G(e^{i\omega}, \theta_o) = D + \sum_{k=0}^{n-1} \mathcal{V}_k^T(e^{i\omega}) L_k$$

substitution of the previous equations delivers

$$G(e^{i\omega}, \theta_o)$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[G_0(e^{i\zeta}) + \sum_{k=0}^{n-1} \mathcal{V}_k^T(e^{i\omega}) \mathcal{V}_k(e^{-i\zeta}) G_0(e^{i\zeta}) \right] d\zeta$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} [\phi^T(e^{i\omega}) \phi(e^{-i\zeta})] G_0(e^{i\zeta}) \, d\zeta$$

which proves the result. $\square$

Proof of Proposition 6.3: Using the results of Theorem 5.5-2) it follows that in the considered limit

$$\frac{N}{n_p} \text{cov}[\hat{G}(e^{i\omega_1}), \hat{G}(e^{i\omega_2})]$$
$$= \frac{1}{n_p N_p} \phi^T(e^{i\omega_1}) \Phi^* Q \Phi \phi(e^{-i\omega_2}). \qquad (A.17)$$

This can be verified by using the fact that there exists a $c \in R$ such that $\|\phi(e^{i\omega})\|_2 \leq c\sqrt{n_p}$. Expression (A.17) equals

$$\frac{1}{n_p N_p} \phi^T(e^{i\omega_1}) \sum_{\ell=1}^{N_p} \phi(e^{-i\omega_\ell}) \frac{\Phi_v(\omega_\ell)}{\Phi_u(\omega_\ell)} \phi^T(e^{i\omega_\ell}) \phi(e^{-i\omega_2}).$$

By applying Lemma 8.1, in the limit $N_p \to \infty$ this expression will be equal to

$$\frac{1}{n_p} \frac{1}{2\pi} \int_{-\pi}^{\pi} [\phi^T(e^{i\omega_1}) \phi(e^{-i\omega})] \frac{\Phi_v(\omega)}{\Phi_u(\omega)} [\phi^T(e^{i\omega}) \phi(e^{-i\omega_2})] \, d\omega \tag{A.18}$$

which proves the result. □

Proof of Corollary 6.4: Substituting $\Phi_v(\omega)/\Phi_u(\omega) = 1$ in (A.18), and using the orthonormality of the basis functions, i.e.,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(e^{-i\omega}) \phi^T(e^{i\omega}) \, d\omega = I_{n_p}$$

it follows that (A.18) simplifies to $\frac{1}{n_p} \phi^T(e^{i\omega_1}) \phi(e^{-i\omega_2})$, which proves the result. □

Proof of Lemma 6.7: Let the spectral densities Φ_v and Φ_u have stable minimum phase spectral factors, respectively, H_v and H_u, and denote $H = H_v(H_u)^{-1}$. Then it follows from Parsseval's relation that we can write the right-hand side expression of (30) as

$$\sum_{t=0}^{\infty} x_j(t) x_k^T(t)$$

where $x_j(t) = G_b^j(q) V_0(q) H(q) \delta(t)$, with δ the unit pulse signal, or equivalently $x_j(t) = G_b^j(q) H(q) v_0(t)$, with $\sum_{t=0}^{\infty} v_0(t) z^{-t} = \mathcal{V}_0(z)$. Note that x_j, x_k are n_b-dimensional ℓ_2-signals. It is proven in [34] that by applying the appropriate Hambo transformation, it follows that

$$\sum_{t=0}^{\infty} x_j(t) x_k^T(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega(j-k)} \tilde{H}^T(e^{-i\omega}) \tilde{H}(e^{i\omega}) \, d\omega$$

and since $\tilde{H}^T(e^{-i\omega}) \tilde{H}(e^{i\omega}) = \tilde{\Phi}_v(\omega) \tilde{\Phi}_u(\omega)^{-1}$, this proves the result. □

Proof of Theorem 6.8: For the frequency response of the estimated model we can write $\hat{G}(e^{i\omega}) - \hat{D} = [\mathcal{V}_0^T(e^{i\omega}) \cdots \mathcal{V}_{n-1}^T(e^{i\omega})] \hat{\theta}_r$, while the asymptotic covariance matrix of $\hat{\theta}_r$ is the block-Toeplitz matrix induced by the spectral density function $\tilde{\Phi}_v(\omega) \tilde{\Phi}_u(\omega)^{-1}$. This situation is identical to the situation for the time domain identification method described and analyzed in [34] leading to the corresponding result of the theorem. □

REFERENCES

[1] D. S. Bayard, "High-order multivariable transfer function curve fitting: Algorithms, sparse matrix methods and experimental results," *Automatica*, vol. 30, pp. 1439–1444, 1994.
[2] D. R. Brillinger, *Time Series—Data Analysis and Theory*. San Francisco, CA: Holden-Day, 1981.
[3] W. R. Cluett and L. Wang, "Frequency smoothing using Laguerre model," *Proc. Inst. Elec. Eng.*, vol. 139, pt. D, pp. 88–96, 1992.
[4] D. K. De Vries, "Identification of model uncertainty for control design," Ph.D. dissertation, Delft Univ. Technol., The Netherlands, 1994.
[5] D. K. De Vries and P. M. J. Van den Hof, "Quantification of uncertainty in transfer function estimation: A mixed probabilistic—Worst-case approach," *Automatica*, vol. 31, pp. 543–557, 1995.
[6] ———, "Frequency domain identification with generalized orthonormal basis functions," in *Proc. 34th IEEE Conf. Decision Contr.*, New Orleans, LA, 1995, pp. 1240–1245.
[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: John Hopkins Univ. Press, 1989.
[8] G. Gu and P. P. Khargonekar, "Linear and nonlinear algorithms for identification in $\mathcal{H}_\infty$ with error bounds," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 953–963, 1992.
[9] R. G. Hakvoort and P. M. J. Van den Hof, "Frequency domain curve fitting with maximum amplitude criterion and guaranteed stability," *Int. J. Contr.*, vol. 60, pp. 809–825, 1994.
[10] ———, "Consistent parameter bounding identification for linearly parametrized model sets," *Automatica*, vol. 31, pp. 957–969, 1995.
[11] A. J. Helmicki, C. A. Jacobson, and C. N. Nett, "Control oriented system identification: A worst-case/deterministic approach," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1163–1176, 1991.
[12] P. S. C. Heuberger, P. M. J. Van den Hof, and O. H. Bosgra, "A generalized orthonormal basis for linear dynamical systems," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 451–465, 1995.
[13] P. S. C. Heuberger and P. M. J. Van den Hof, "The Hambo transform: A signal and system transform induced by generalized orthonormal basis functions," in *Proc. 13th IFAC World Congr.*, J. J. Gertler *et al.*, Eds., San Francisco, CA, June 30–July 5, 1996, vol. I, pp. 357–362.
[14] E. C. Levy, "Complex curve fitting," *IRE Trans. Automat. Contr.*, vol. 4, pp. 37–43, 1959.
[15] P. L. Lin and Y. C. Wu, "Identification of multi-input multi-output linear systems from frequency response data," *J. Dyn. Syst., Meas. Contr.*, vol. 4, pp. 37–43, 1982.
[16] L. Ljung, "Asymptotic variance expressions for identified black-box transfer function models," *IEEE Trans. Automat. Contr.*, vol. 30, pp. 834–844, 1985.
[17] ———, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[18] ———, "Some results on identifying linear systems using frequency domain data," in *Proc. 32nd IEEE Conf. Decision Contr.*, San Antonio, TX, 1993, pp. 3534–3538.
[19] ———, "Building models from frequency domain data," in *Adaptive Control, Filtering, and Signal Processing*, IMA Vol. in Math. and Its Appl., K. J. Aström *et al.*, Eds. New York: Springer Verlag, 1995, vol. 74, pp. 229–239.
[20] L. Ljung and Z. D. Yuan, "Asymptotic properties of black-box identification of transfer functions," *IEEE Trans. Automat. Contr.*, vol. 30, pp. 514–530, 1985.
[21] T. McKelvey and H. Akcay, "An efficient frequency domain state-space identification algorithm: Robustness and stochastic analysis," in *Proc. 33rd IEEE Conf. Decision Contr.*, Lake Buena Vista, FL, 1994, pp. 3348–3353.
[22] T. McKelvey, *Identification of State-Space Models from Time and Frequency Domain Data*, Linköping Studies in Sci. and Technol., Diss. no. 380, 1995.
[23] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, pp. 997–1009, 1991.
[24] C. T. Mullis and R. A. Roberts, "The use of second order information in the approximation of discrete time linear systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 24, pp. 226–238, 1976.
[25] B. M. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," in *Proc. 33rd IEEE Conf. Decision Contr.*, Lake Buena Vista, FL, 1994, pp. 3388–3393; also *IEEE Trans. Automat. Contr.*, vol. 42, pp. 515–521, 1997.
[26] B. M. Ninness, J. C. Gomez, and S. R. Weller, "MIMO system identification using orthonormal basis functions," in *Proc. 34th IEEE Conf. Decision Contr.*, New Orleans, LA, 1995, pp. 703–708.

[27] J. R. Partington, "Robust identification and interpolation in $\mathcal{H}_\infty$," *Int. J. Contr.*, vol. 54, pp. 1281–1290, 1991.

[28] R. Pintelon, P. Guillaume, Y. Rolain, J. Schoukens, and H. Vanhamme, "Parametric identification of transfer functions in the frequency domain—A survey," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 2245–2260, 1994.

[29] C. K. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE Trans. Automat. Contr.*, vol. 8, pp. 56–58, 1963.

[30] J. Schoukens and R. Pintelon, *Identification of Linear Systems. A Practical Guideline to Accurate Modeling.* Oxford, UK: Pergamon, 1991.

[31] M. D. Sidman, F. E. DeAngelis, and G. C. Verghese, "Parametric system identification on logarithmic frequency response data," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1065–1070, 1991.

[32] J. T. Spanos, "Algorithms for ℓ_2 and ℓ_∞ transfer function curve fitting," in *Proc. AIAA Guid., Navig. and Contr. Conf.*, New Orleans, LA, 1991, pp. 1739–1747.

[33] A. W. M. van de Enden and G. A. L. Leenknegt, "Design of optimal IIR filters with arbitrary amplitude and phase requirements," in *Signal Processing III: Theories and Applications*, I. T. Young *et al.*, Eds. North-Holland: Elsevier Sci., 1986.

[34] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor, "System identification with generalized orthonormal basis functions," *Automatica*, vol. 31, pp. 1821–1834, 1995.

[35] E. T. van Donkelaar, "Identification of model uncertainty for a compact disc pick-up mechanism," Mech. Eng. Syst. and Contr. Group, Delft Univ. Technology, Rep. A-747, Aug. 1995.

[36] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 551–562, 1991.

[37] E. Walter and H. Piet-Lahanier, "Estimation of parameter bounds from bounded error data: A survey," *Math. Comput. Simul.*, vol. 32, pp. 449–468, 1990.

[38] A. H. Whitfield, "Asymptotic behavior of transfer function synthesis methods," *Int. J. Contr.*, vol. 45, pp. 1083–1092, 1987.

Douwe K. de Vries was born in The Netherlands, in 1964. In 1989 he graduated from the Department of Mechanical Engineering and Marine Technology, Delft University of Technology, The Netherlands. In 1994 he received the Ph.D. degree from the same university.

After a period with the Computer Management Group, he joined the Department of Systems Engineering, Policy Analysis and Management, Delft University of Technology. His research interests include identification, focusing on uncertainty quantification, the interplay between identification and control design, and the interplay between identification and physical modeling. Currently he is working on logistic planning and control problems in applications.

Paul M. J. Van den Hof (S'85–M'88–SM'95) was born in Maastricht, The Netherlands, in 1957. He obtained the M.Sc. and Ph.D. degrees, both from the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands, in 1982 and 1989, respectively.

Since 1986 he has been working in the Mechanical Engineering Systems and Control Group, Delft University of Technology, The Netherlands, where since 1991 he holds a position as an Associate Professor. In 1992 he held a short-term visiting position at the Centre for Industrial Control Science, The University of Newcastle, NSW, Australia. His research interests include issues of system identification, parameterization, and the interplay between identification and robust control design, with applications in mechanical servo systems and industrial process control systems. He is an Associate Editor of *Automatica*.