JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **166**, 428–441 (1992)

Robust Identification in H_∞

JONATHAN R. PARTINGTON

School of Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom

Submitted by C. Foias

Received November 8, 1990

We consider system identification in the Banach space H_∞ in the framework proposed by Helmicki, Jacobson, and Nett. It is shown that there is no robustly convergent linear algorithm for identifying exponentially stable systems in the presence of noise which is not tuned to prior information about the unknown system or noise. Various nonlinear algorithms, some closely related to one of Gu and Khargonekar, are analysed, and results on trigonometric interpolation used to provide new error bounds. An application of these techniques to approximation is given, and finally some numerical results are provided for illustration. © 1992 Academic Press, Inc.

1. INTRODUCTION

The H_∞ identification of a linear system (in either discrete or continuous time) with transfer function G involves the determination of a set of n frequency response measurements, which are then used to construct an approximation to the original transfer function. In addition it is desirable that an identification algorithm be stable under small perturbations of the data (noise): for example, Lagrange interpolation (fitting a polynomial of minimal degree through the data) is known not to have this property and to be generally badly behaved. We therefore seek algorithms with better stability properties.

Section 2 of this paper sets up the mathematical background, and in Section 3 we discuss the existence and non-existence of linear identification algorithms with the desired properties. Section 4 introduces various non-linear algorithms and gives error bounds establishing their approximation power. In Section 5 we consider applications of these techniques to model reduction. Finally Section 6 illustrates the theory with some examples.

The mathematical framework in which we shall work is principally due to Helmicki, Jacobson, and Nett, and I am very grateful to Professor Nett for introducing me to the problems discussed in this paper.

0022-247X/92 $5.00

2. Mathematical Formulation of the Problem

We shall be considering stable, infinite-dimensional, linear time-invariant systems, and our results will be applicable to both discrete-time and continuous time systems. In continuous time the transfer function $G(s)$ is a bounded analytic function in the right half complex plane, i.e., an element of $H_\infty(\mathbb{C}_+)$ which acts on $H_2(\mathbb{C}_+)$ by multiplication; in discrete time, the transfer function $g(z)$ (obtained by taking z-transforms) is analytic and bounded on the set of complex numbers of modulus greater than one. In either case we can, and will, transform the situation in a norm-preserving manner so as to consider functions analytic and bounded on the unit disc: in the first case by defining $f(z) = G(\mathcal{M}z)$ so that $G(s) = f(\mathcal{M}s)$, where $\mathcal{M}$ is the Möbius map $\mathcal{M}z = (1-z)/(1+z)$; and in the second case by defining $f(z) = g(1/z)$.

Let $f(z)$ be a function in the disc algebra $A(D) = H_\infty \cap C(\mathbb{T})$, i.e., a function analytic and bounded on the unit disc with continuous boundary values. An H_∞ identification algorithm can be regarded as a sequence (T_n) of (not necessarily linear) mappings from subsets $S_n \subseteq \mathbb{C}^n$ into H_∞: given a set of possibly corrupted values of the function f on the unit circle, say $(f_1, ..., f_n) \in S_n$, obtained from evaluating f at points $(z_1, ..., z_n)$ on the circle, and adding in noise function η, the function $T_n(f_1, ..., f_n)$ defines an approximation to f. The noise η can be regarded as lying in $l_\infty(\mathbb{T})$, the space of bounded functions on the circle $\mathbb{T}$. Thus $f_1 = f(z_1) + \eta(z_1), ..., f_n = f(z_n) + \eta(z_n)$. It is usual to evaluate the function at the roots of unity, i.e., $z_j = \exp(2\pi ij/n)$, for $j = 1, ..., n$, but we shall not initially assume this.

Let A be the set of functions that we wish the algorithm to identify and $E \subseteq l_\infty(\mathbb{T})$ a set in which the noise η is assumed to lie; then T_n defines a mapping $\tilde{T}_n$ from $A \times E$ into H_∞, by the formula

$$\tilde{T}_n(f, \eta) = T_n(f(z_1) + \eta(z_1), ..., f(z_n) + \eta(z_n)). \tag{2.1}$$

Following Helmicki, Jacobson, and Nett [6] we say that an algorithm is *convergent* over $H_\infty(D_\rho, M)$, the set of functions analytic in $D_\rho = \{z: |z| < \rho\}$ $(\rho > 1)$ with $\sup\{|f(z)|: |z| < \rho\} \leqslant M$, if

$$\lim_{n \to \infty, \varepsilon \to 0} \sup_{\|\eta\|_\infty < \varepsilon} \sup_{f \in H_\infty(D_\rho, M)} \|\tilde{T}_n(f + \eta) - f\|_\infty \to 0. \tag{2.2}$$

Moreover an algorithm is said to be *robustly convergent* when the above is true for all $\rho > 1$ and $M > 0$, and *untuned* if it does not depend on prior information about ρ, M, and ε.

It is shown that no untuned algorithm can exist which is robustly convergent and such that the maps T_n are linear in the data: this was conjected by Helmicki, Jacobson, and Nett [6] who provided a robustly convergent untuned algorithm using nonlinear mappings, as later did Gu and Khargonekar [5]. More such nonlinear algorithms are discussed and error bounds provided guaranteeing asymptotically correct approximation of all f in the disc algebra; and hence, by transforming to the half-plane, of all $G(s) \in H_\infty(\mathbb{C}_+)$ such that G is continuous on the imaginary axis, including at ∞.

3. Linear Algorithms

Helmicki, Jacobson, and Nett [7] have constructed a convergent linear algorithm, tuned to prior data ρ, M, ε such that the identified function f lies in $H_\infty(D_\rho, M)$ and the noise satisfies $\|\eta\|_\infty < \varepsilon$. Moreover they conjectured that no linear robustly convergent algorithm exists. In this section we establish their conjecture, and in fact prove a slightly stronger result.

To establish notation, let us suppose that an identification map T_n uses values of $f+\eta$ at a set of points $\{z_{n,1}, ..., z_{n,n}\}$ on the circle, and write E_n for the evaluation map taking $g \in l_\infty(\mathbb{T})$ to $(g(z_{n,1}), ..., g(z_{n,n}))$.

Theorem 3.1. *There is no sequence of linear maps* $(U_n)\colon \mathbb{C}^n \to H_\infty$ *such that, writing* E_n *for the evaluation map above,*

(i) $\|U_n E_n(\eta)\|_\infty \to 0$ *as* $\|(\eta(z_{n,1}), ..., \eta(z_{n,n}))\|_\infty \to 0$ *and* $n \to \infty$ *jointly; and also*

(ii) $\|U_n E_n(g) - g\|_\infty \to 0$ *in norm for every* g *which is a polynomial.*

The first condition implies that the zero function is identified correctly in the limit as the noise level tends to zero and the number of evaluation points tends to infinity, jointly. The second implies that in the zero-noise case, we obtain correct identification of polynomials in the limit.

Proof. From condition (i) there exists a $\delta > 0$ and a number N such that if $n > N$ and $\|\eta(z_{n,1}), ..., \eta(z_{n,n})\|_\infty < \delta$ then $\|U_n E_n(\eta)\| < 1$. Now given any function $f \in C(\mathbb{T})$, there is a constant $A > 0$ such that $\|f/A\|_\infty < \delta$. This implies that $\|U_n E_n f\|_\infty < A$ for all $n > N$ and hence the complete sequence $(U_n E_n f)$ is bounded in norm.

This in turn implies, by the Banach–Steinhaus principle of uniform boundedness (see, e.g., [14]) that the linear operators $V_n = U_n E_n$ are uniformly bounded in norm, considered as mappings from $C(\mathbb{T})$ into H_∞. We also have that $V_n f \to f$ for any polynomial f. Because the maps are uniformly bounded in norm this is also true for any f in the closure of the polynomials, namely the disc algebra $A(D) = H_\infty \cap C(\mathbb{T})$.

We therefore need to show that there is no uniformly bounded sequence of linear maps, (V_n) from $C(\mathbb{T})$ into H_∞ such that $V_n f \to f$ for all $f \in A(D)$.

It is known that there is no continuous projection from $C(\mathbb{T})$ onto $A(D)$, but we are claiming something stronger here because we do not require convergence for f not in $A(D)$. We adapt the proof of this weaker result as can be found in Hoffman [9, p. 155].

Given $f \in C(\mathbb{T})$ and $0 \leqslant \theta \leqslant 2\pi$ define $f_\theta \in C(\mathbb{T})$ by $f_\theta(z) = f(e^{i\theta}z)$. So, e.g., for $f(z) = z^n$, $f_\theta(z) = e^{in\theta}z^n$.

Now define S_n as bounded maps from $C(\mathbb{T})$ into H_∞ by

$$(S_n f, g) = \int_0^{2\pi} (V_n f_\theta, g_\theta)\, d\theta/2\pi,$$

where g is any function in $L_1(\mathbb{T})$. Initially all we can say is that $S_n f$ is in L_∞ but in fact

$$(S_n z^k, z^m) = \int_0^{2\pi} (V_n z^k, z^m)\, e^{i(k-m)\theta}\, d\theta/2\pi,$$

which gives 0 for $k \neq m$ and $(V_n z^k, z^k)$ for $k = m$ and this last term is zero for $k < 0$ since V_n maps into H_∞.

Therefore we have uniformly continuous maps S_n from $C(\mathbb{T})$ into H_∞ such that

$$S_n(z^k) = \begin{cases} a_{k,n} z^k, & \text{if } k \geqslant 0, \\ 0, & \text{if } k < 0, \end{cases}$$

where the $(a_{k,n})$ are constants. Also the numbers $a_{k,n}$ are uniformly bounded and for fixed k, $a_{k,n} \to 1$ as $n \to \infty$.

We therefore get a continuous projection S from $C(\mathbb{T})$ into H_∞ by defining S on trigonometric polynomials as the limit of S_n and extending by continuity. Clearly S has to be bounded, but we see from above that

$$S(z^k) = \begin{cases} z^k, & \text{if } k \geqslant 0, \\ 0, & \text{if } k < 0, \end{cases}$$

but this map is well known not to be bounded (see, e.g., [9, p. 150]).

Hence such an algorithm cannot exist. ∎

We have actually proven a stronger assertion, namely that even if the noise is known to be in $C(\mathbb{T})$, i.e., continuous on the circle, robust identification is still impossible with a linear, untuned algorithm. It is therefore necessary to concentrate on nonlinear algorithms, which we do in the next section.

4. Nonlinear Algorithms and Error Bounds

A general class of robustly convergent nonlinear algorithms for identification in H_∞ operates as follows:

1. Given data $(f_1, ..., f_n) = (f(z_j) + \eta(z_j))_{j=1}^n$ form an approximation $L_n(f+\eta) \in L_\infty$ to f, in such a way that

   (i) in the zero-noise case, $L_n(f) \to f$ as $n \to \infty$ for f in a suitably large class of functions;

   (ii) the approximation is robust with respect to the data, that is, for any $\varepsilon > 0$ there is a $\delta(\varepsilon) > 0$ such that, provided that $\|(\eta(z_1), ..., \eta(z_n))\|_\infty < \delta$, then $\|L_n(f+\eta) - L_n(f)\|_\infty < \varepsilon$ for all f.

2. If $L_n(f+n)$ is not already rational, perform a model reduction scheme to obtain a rational approximation $R_n(f+\eta)$ to $L_n(f+\eta)$ in the L_∞ norm. Generally steps 1 and 2 are linear in the data.

3. Define $A_n(f+\eta)$ to be the function in H_∞ which is closest to $R_n(f+\eta)$. This can be calculated by solving a Nehari extension problem (see, e.g., [15, 12]) and it is at this stage that we lose linearity in the data.

In the common case that R_n is a trigonometric polynomial $\sum_{k=-m}^{m} c_k z^k$, the solution is given as follows: let Γ be the Hankel matrix

$$\Gamma = \begin{pmatrix} c_{-1} & c_{-2} & \cdots & c_{-m} \\ c_{-2} & c_{-3} & \cdot^{\cdot^{\cdot}} & 0 \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ c_{-m} & 0 & \cdots & 0 \end{pmatrix}$$

and let $v = (v_1, ..., v_m)$ be a nonzero vector such that $\|\Gamma v\|_2 = \|\Gamma\| \, \|v\|$ and $\Gamma v = (w_1, ..., w_m)$, say. Then the closest element of H_∞ to R_n is

$$\sum_{k=-m}^{m} c_k z^k - \frac{\sum_{i=1}^{m} w_i z^{-i}}{\sum_{j=1}^{m} v_j z^{j-1}}, \tag{4.1}$$

a rational function of degree with numerator degree at most $2m-1$ and denominator degree at most $m-1$.

Then, as regards error bounds, we see that

$$\|A_n(f+\eta) - R_n(f+\eta)\|_\infty \leqslant \|f - R_n(f+\eta)\|_\infty$$

and so

$$\begin{aligned} \|A_n(f+\eta) - f\|_\infty &\leqslant 2\,\|f - R_n(f+\eta)\|_\infty \\ &\leqslant 2\,\|f - R_n(f)\|_\infty + 2\,\|R_n(f+\eta) - R_n(f)\|_\infty, \end{aligned} \tag{4.2}$$

and both terms will be small if n is sufficiently large and η is sufficiently small. Further, if R_n is linear in the data, then

$$\|A_n(f+\eta)-f\|_\infty \leqslant 2\,\|f-R_n(f)\|_\infty + 2\,\|R_n(\eta)\|_\infty. \tag{4.3}$$

Helmicki, Jacobson, and Nett [6] performed stage 1 by means of a linear spline through the given data, and stage 2 by a truncated Fourier series approximation (for which the formulae are particularly convenient in this case). Gu and Khargonekar [5] combined stages 1 and 2 by taking discrete Fourier transforms of the data, and using Cesàro averages. We shall start by giving a generalization of their method, together with some new error bounds guaranteeing convergence. After this we shall present some further algorithms with rates of convergence.

We recall, from Zygmund [16, Vol. 2, Chap. X], the following results on Jackson (trigonometric) polynomials.

Let $z=e^{i\theta}$, with $0\leqslant\theta\leqslant 2\pi$, be a point of the unit circle $\mathbb{T}$. The Dirichlet kernel is defined by

$$D_n(\theta)=\frac{\sin(n+1/2)\,\theta}{2\sin\theta/2}$$
$$=(1/2)\sum_{k=-n}^{n} z^k.$$

Moreover the Fejér Kernel is given by

$$K_n(\theta)=(1/(n+1))\sum_{k=0}^{n} D_k(\theta).$$

We now wish to interpolate data $f(t_0), ..., f(t_{2n})$, where $t_r=2\pi r/(2n+1)$, $r=0, ..., 2n$, with a trigonometric polynomial $\sum_{k=-n}^{n} c_k e^{ik\theta}$ as follows.

Define the generalized Jackson (trigonometric) polynomials $(J_{n,m})_{m>n\geqslant 0}$ by

$$J_{n,m}(f,\theta)=(2/m)\sum_{r=0}^{m-1} f(2\pi r/m)\,K_n(\theta-(2\pi r/m). \tag{4.4}$$

This is a trigonometric polynomial of degree n, i.e., a power series from z^{-n} to z^n. Also let

$$I_{n,k}(f,\theta)=(2/(2n+1))\sum_{r=0}^{2n} f(t_r)\,D_k(t_r-\theta),$$

for $k=0, 1, ..., n$.

Two especially interesting cases present themselves:

(i) If $m = n+1$ we obtain the Jackson polynomials $J_{n,n+1} = J_n$; these interpolate the data exactly at the $(n+1)$st roots of unity and have zero derivative at those points.

(ii) If $m = 2n+1$ we obtain the Marcinkiewicz polynomials

$$B_{n,n} = J_{n,2n+1} = (1/(n+1)) \sum_{k=0}^{n} I_{n,k}. \tag{4.5}$$

These polynomials have the following properties:

(a) $J_{n,m}$ remains within the same bounds as f, i.e., interpolation is a norm 1 map from the data (infinity norm) to the continuous functions on the circle.

(b) $J_{n,m}(f, \theta) \to f$ uniformly for any continuous f as $n \to \infty$, keeping $m > n$.

The nonlinear algorithm of Gu and Khargonekar [5] is a close relation of the algorithm using (4.4) above, using an even number of interpolated values rather than an odd number and essentially producing the trigonometric polynomial $J_{n,2n}$. They established robust convergence of their algorithm for all f in $H_\infty(D_\rho, M)$ for some $\rho > 1$ and $M > 0$. Results from [16] show that we may expect convergence over a much wider class of functions.

Since for each k,

$$|f(\theta) - I_{n,k}(f, \theta)| \leqslant 2 \sum_{|j| > k} |a_j|, \tag{4.6}$$

where a_j are the (Fourier) coefficients of f (see [16, Chap. X, Theorem 5.16]), clearly the same error bound holds for the Marcinkiewicz polynomials $B_{n,n}$. This gives us an effective error bound as follows.

THEOREM 4.1. *Suppose f is an analytic function in the Wiener algebra on the circle, that is,*

$$f(z) = \sum_{m=0}^{\infty} a_m z^m \qquad \textit{with} \qquad \sum_{m=0}^{\infty} |a_m| < \infty.$$

Let $B_{n,n} = J_{n,2n+1}$ be the Marcinkiewicz trigonometric polynomial interpolating the function $f(e^{i\theta}) + \eta(e^{i\theta})$, with $\eta \in l_\infty(\mathbb{T})$, and let F_n be the identified *model in H_∞ formed by taking the closest point in H_∞ to $B_{n,n}$, as in* (4.1) *above. Then*

$$\begin{aligned}\|F_n - f\|_\infty &\leqslant \frac{4}{n}\sum_{k=1}^{n}\sum_{|j|>k}|a_j| + 2\|\eta\|_\infty\\ &= \frac{4}{n}\sum_{k=1}^{n}(k-1)|a_k| + 4\sum_{k>n}|a_k| + 2\|\eta\|_\infty. \end{aligned} \tag{4.7}$$

Proof. This follows immediately from (4.5), (4.6), and (4.3). ∎

In the particular case that $f \in H_\infty(D_\rho, M)$ we observe that $|a_k| \leqslant M\rho^{-k}$ from which one obtains very similar error bounds to those of Gu and Khargonekar.

The following result is particularly important for the identification of continuous time delay systems $G(s)$, where the transfer function transformed to the disc, $G((1-z)/(1+z))$, does not lie in any $H_\infty(D_\rho, M)$. We shall give more explicit bounds later. We write

$$D_n(f) = \|f - B_{n,n}(f)\|_\infty. \tag{4.8}$$

THEOREM 4.2. *The algorithm above using $B_{n,n}$ approximates all functions f in the disc algebra, with an error bound of the form $2D_n(f) + 2\varepsilon$, and $D_n(f) \to 0$ as $n \to \infty$.*

Proof. This follows from (4.8) and (4.3). ∎

The algorithm using $J_{n,n+1}$ seems to perform less satisfactorily, giving error bounds $O(n^{-1/3}) + O(\varepsilon)$ in the case that f has a continuous derivative. However, it resembles the spline approximants of Helmicki, Jacobson, and Nett in that it starts by interpolating the data (unlike the other algorithm).

A further algorithm of some interest may be constructed using some results from Natanson [11, Vol. III, Chap. 4]. There no less than four procedures are given for obtaining trigonometric interpolants to data on the circle. Of these the first two seem to be most useful when combined with a Nehari extension. The first one, the "First Bernstein procedure" (which we denote by (B1)) leads to the Jackson polynomials $J_{n,2n+1}$ that we have already discussed.

The second, the "Second Bernstein procedure" (which we denote by (B2)) operates as follows: let T_n be the trigonometric polynomial of degree n interpolating f at the $(2n+1)$st roots of unity and let

$$U_n(f, \theta) = \frac{1}{2}\left(T_n\left(\theta + \frac{\pi}{2n+1}\right) + T_n\left(\theta - \frac{\pi}{2n+1}\right)\right). \tag{4.9}$$

Then this is slightly less robust (i.e., it may exaggerate noise errors by a constant larger than 1) but may converge more quickly for noiseless f: thus if one suspects that the noise is small this algorithm may be superior to (B1) but will always approximate anyway. This algorithm is also easy to implement: the data $f+\eta$ is mapped to the function

$$U_n(f+\eta,\theta)=\sum_{m=-n}^{n} z^m(1/(2n+1)) \times \sum_{k=0}^{2n} (f+\eta)(z_k)\, z_k^{-m} \cos(\pi m/(2n+1)), \tag{4.10}$$

where $z_k = \exp(2\pi ik/(2n+1))$. U_n is a trigonometric polynomial of degree n (based on $(2n+1)$ data points) and is closely related to the discrete Fourier transform.

The third and fourth methods are more complicated and it is not clear that they lead to robust algorithms. The reader is referred to [11] for details.

The theoretical error bound of (B2) is $O(\omega(1/n))$ where ω is the modulus of continuity of f. That is,

$$\|U_n(f,\theta)\|_\infty \leqslant K_1 \|f\|_\infty \qquad \text{and} \qquad \|U_n(f,\theta)-f\|_\infty \leqslant K_2\omega(1/n) \tag{4.11}$$

for fixed constants K_1 and K_2 (whose smallest possible values Natanson does not attempt to determine), where

$$\omega(\delta)=\sup\{|f(e^{i\theta})-f(e^{i\phi})| : |\theta-\phi| \leqslant \delta\}. \tag{4.12}$$

Note that for a function which is continuously differentiable on the circle, one has

$$\omega(\delta) \leqslant \delta \, \|f_\theta\|_\infty. \tag{4.13}$$

Thus the error in using (B2) will be $O(1/n)$ if $\partial f/\partial\theta$ is bounded on the circle. As above, one can combine methods (B1) and (B2) with a Nehari extension as in (4.1), which we shall still refer to as methods (B1) and (B2): we thus deduce the following result.

THEOREM 4.3. *Using technique* (B2) *above produces an identification algorithm which approximates any $f \in A(D)$ yielding an approximation F_n with an error bound*

$$\|F_n - f\|_\infty \leqslant O(\|\eta\|_\infty) + O(\omega(1/n)). \tag{4.14}$$

This algorithm is therefore robustly convergent over $H_\infty(D_\rho, M)$ for any $\rho > 1$ and $M > 0$.

Proof. This follows from (4.11) and (4.3), together with easy bounds on f_θ for $f \in H_\infty(D_\rho, M)$. ▮

One important application here is to the identification of the class of retarded delay systems. The most general SISO retarded delay system has the transfer function

$$G(s) = h_2(s)/h_1(s),$$

where

$$h_1(s) = \sum_{o}^{n_1} p_i(s)\, e^{-\gamma_i s},$$

and

$$h_2(s) = \sum_{o}^{n_2} q_i(s)\, e^{-\beta_i s},$$

with $0 = \gamma_o < \gamma_1 < \cdots < \gamma_{n_1}$, $0 \leqslant \beta_o < \cdots < \beta_{n_2}$, the p_i being polynomials of degree δ_i, and $\delta_i < \delta_o$ for $i \neq 0$, and the q_i polynomials of degree $d_i < \delta_o$ for each i. Such functions were analysed by Bellman and Cooke [1], who showed that they have only finitely many poles in any right half plane. They have since been considered from the viewpoint of approximation in numerous papers and we begin by summarising some known results about such systems.

It is convenient to define the *index* of a stable delay system, $I(G)$, to be the unique integer $r \geqslant 1$ such that one may write

$$G(s) = R(s) + \sum_{1}^{P} a_i e^{-\alpha_i s}/(s+1)^r + O(s^{-r-1}),$$

with R rational, $P \geqslant 1$, the (a_i) nonzero constants, and the (α_i) positive coefficients. For the case $G(s) = e^{-sT}R(s)$ with R rational (i.e., $n_1 = 0$), the index $I(G)$ is just the relative degree of R.

It is then the case that the Hankel singular values σ_k of $G(s)$ are bounded above and below by constant multiples of $k^{-I(G)}$, and that the minimum achievable H_∞ error in approximating G by rationals of degree k has the same property, whereas the minimum achievable H_2 error behaves as $k^{-I(G)+1/2}$ (see [2–4]).

Moreover, when one transforms to the disc, the Fourier coefficients of the transformed function $G((1-z)/(1+z))$ are bounded above by a multiple of $k^{-3/4-I(G)/2}$ (see [10, 13]). The transformed function will in all cases have an essential singularity at the point $z = -1$, and hence not lie in $H_\infty(D_\rho)$ for any $\rho > 1$.

LEMMA 4.4. *Let $G(s)$ be a stable retarded delay system, and $g(z) = G((1-z)/(1+z))$ the transformed function in $A(D)$. Then the modulus of continuity $\omega(\delta)$ is $O(\delta)$ unless G has index 1, in which case it is $O(\delta^{1/2})$.*

Proof. By virtue of standard techniques of expanding delay systems about infinity, given a significant part plus a well-behaved remainder term (as in [2, 3, 13]), it is sufficient to verify this assertion for functions of the form $G_m(s) = \exp(-\lambda s)/(s+1)^m$, which give $g_m(z) = ((1+z)/2)^m \exp(\lambda(z-1)/(z+1))$. It is now straightforward to verify that for $m \geq 2$, $\partial g/\partial\theta$ is bounded on the circle, whereas for θ small, $g_1(\exp(i(\pi+\theta)))$ behaves as $\theta \sin(1/\theta)$ does near zero, from which it is easy to estimate its modulus of continuity. ∎

THEOREM 4.5. *The following error bounds hold for the identification of stable retarded delay systems in the presence of noise satisfying $\|\eta\|_\infty \leq \varepsilon$.*

(i) *If $I(G)=1$, then method* (B1) *has an error of $O(n^{-1/4}) + O(\varepsilon)$ and* (B2) *has an error of $O(n^{-1/2}) + O(\varepsilon)$.*

(ii) *If $I(G)=2$, then method* (B1) *has an error of $O(n^{-3/4}) + O(\varepsilon)$ and* (B2) *has an error of $O(n^{-1}) + O(\varepsilon)$.*

(iii) *If $I(G) \geq 3$, then both methods have an error $O(n^{-1}) + O(\varepsilon)$.*

Proof. The bounds for (B1) follow from Theorem 4.1 and the estimates of Fourier coefficients given above. Those for (B2) follow from Lemma 4.4 and Theorem 4.3. ∎

5. APPLICATION TO MODEL REDUCTION

Suppose that $g(z) \in A(D)$ and hence is uniformly approximable not only by trigonometric polynomials but by stable rational functions. Let

$$e_n(g) = \inf\{\|g(e^{i\theta}) - T(e^{i\theta})\|_\infty : T \text{ a trigonometric polynomial of degree at most } n\} \tag{5.1}$$

and

$$E_n(g) = \inf\{\|g(e^{i\theta}) - R(e^{i\theta})\|_\infty : R \text{ a stable rational function of degree at most } n\}. \tag{5.2}$$

The following result is an immediate consequence of the technique above, cf. (4.1), (4.2), (4.3), which involve approximation followed by a Nehari extension.

PROPOSITION 5.1. *Let $g(z) \in A(D)$. Then $E_{2n-1}(g) \leqslant 2e_n(g)$.*

Proof. Let T be any trigonometric polynomial of degree n and R its closest approximant in H_∞, a rational function of degree at most $(2n-1)$. Then $\|T-R\| \leqslant \|T-g\|$ so that $\|g-R\| \leqslant \|g-T\| + \|T-R\| \leqslant 2\,\|g-T\|$. Hence the result. ∎

This can be used to obtain results on distributed systems using the Jackson theorems [11, Vol. I, Chap. IV], notably the version for functions on the circle with p continuous derivatives: if $\omega_p(\delta)$ is the modulus of continuity of $f^{(p)}(\theta)$, then

$$e_n \leqslant \frac{12^{p+1}\omega_p(1/n)}{n^p}. \tag{5.3}$$

For delay systems this gives achievable approximation rates which are not optimal: for example, for $G(s) = e^{-s}/(s+1)$, $G(\mathscr{M}z) = ((z+1)/2)\exp((z-1)/(z+1))$, one obtains an error bound of order $n^{-1/2}$ for E_n where the optimal error is known to be of order n^{-1} (see [3]).

However, for a transfer function such as $G(s) = \exp(-\sqrt{s})$ occurring in the theory of transmission lines, where

$$g(z) = G(\mathscr{M}z) = \exp\left(-\sqrt{\frac{1-z}{1+z}}\right),$$

although, as is easily verified, g is not differentiable along the circle, one can still show that $e_n(g)$ goes to zero. This is despite the fact that g is not in $H_\infty(D_\rho)$ for any $\rho > 1$. We expect Proposition 5.1 to be of use in other contexts.

6. NUMERICAL EXAMPLES

Two functions are considered by Gu and Khargonekar [5] for identification, and we use them as a basis for comparing methods (B1) and (B2). Taking $n = 20$ and the noise to be randomly determined but of modulus 0.1 the following quantities are of interest.

1. E_1, the error in the trigonometric interpolant to the data;

2. E_2, the final error in the identified model after the Nehari extension (4.1).

The two functions in question are $f_1(z) = 3(z^2+1)/(z^2+2z+5)$, which is exponentially stable and in $H_\infty(D_\rho)$ for $\rho < \sqrt{5}$; and $f_2(z) = e^{-s}/(s^2+s+1)$, with $s = \mathscr{M}z$, which is a delay system of index 2.

The following table summarizes the L_∞ errors obtained; so as to compare the errors caused by noise and errors caused by finite sampling, we include for reference the errors obtained when identifying with zero noise.

Function	Method	Noise Level	E_1	E_2
f_1	B1	0.1	0.16	0.14
	B2	0.1	0.12	0.15
	B1	0	0.10	0.10
	B2	0	0.02	0.02
f_2	B1	0.1	0.20	0.17
	B2	0.1	0.12	0.14
	B1	0	0.15	0.15
	B2	0	0.04	0.03

These results reflect the error bounds given earlier. Two remarks may be made. Firstly, that the second function is clearly more difficult to identify correctly than the first one, owing to its lack of exponential stability, and that the effects of noise are complemented by the effects of taking a limited number of interpolation points. Secondly, that the Nehari extension step rarely seems to increase the L_∞ error in the identified model, and frequently reduces it.

REFERENCES

1. R. BELLMAN AND K. L. COOKE, "Differential-Difference Equations," Academic Press, Orlando, FL, 1963.
2. K. GLOVER, J. LAM, AND J. R. PARTINGTON, Rational approximation of a class of infinite-dimensional systems. I. Singular values of Hankel operators, *Math. Control Circ. Systems* **3** (1990), 325–344.
3. K. GLOVER, J. LAM, AND J. R. PARTINGTON, Rational approximation of a class of infinite-dimensional systems. II. Optimal convergence rates of L_∞ approximants, *Math. Control. Circ. Systems* **4** (1991), 233–246.
4. K. GLOVER, J. LAM, AND J. R. PARTINGTON, Rational approximations of a class of infinite-dimensional systems: The L_2 case, *in* "Progress in Approximation Theory" (P. Nevai and A. Pinkus, Eds.), Academic Press, San Diego, 1991.
5. G. GU AND P. P. KHARGONEKAR, Linear and nonlinear algorithms for identification in H^∞ with error bounds, *in* "Proceedings, 1991 Amer. Control Conf."
6. A. J. HELMICKI, C. A. JACOBSON, AND C. N. NETT, Identification in H_∞: A robustly convergent, nonlinear algorithm, *IEEE Trans. Automat. Control* **36** (1991), 1163–1176.
7. A. J. HELMICKI, C. A. JACOBSON, AND C. N. NETT, Identification in H_∞: Linear algorithms, *in* "Proceedings, 1990 Amer. Control Conf."
8. A. J. HELMICKI, C. A. JACOBSON, AND C. N. NETT, Identification in H_∞: The continuous-time case, *in* "Proceedings, 1990 Amer. Control Conf."
9. K. HOFFMAN, "Banach Spaces of Analytic Functions," Dover, New York, 1962.
10. P. M. MÄKILÄ, Laguerre series approximation of infinite dimensional systems, *Automatica* **26** (1990), 985–995.

11. I. P. NATANSON, "Constructive Function Theory," Ungar, New York, 1965.
12. J. R. PARTINGTON, "An Introduction to Hankel Operators," Cambridge Univ. Press, London/New York, 1988.
13. J. R. PARTINGTON, Approximation of delay systems by Fourier–Laguerre series, *Automatica* **27** (1991), 569–572.
14. W. RUDIN, "Functional Analysis," McGraw-Hill, New York, 1974.
15. N. J. YOUNG, "An Introduction to Hilbert Space," Cambridge Univ. Press, London/New York, 1988.
16. A. ZYGMUND, "Trigonometric Series," Cambirdge Univ. Press, London/New York, 1959.