Systems & Control Letters 7 (1986) 247–255
North-Holland

July 1986

Weighted Hankel-norm approximation: Calculation of bounds

Brian D.O. ANDERSON

Department of Systems Engineering, Research School of Physical Sciences, Australian National University, Canberra, A.C.T. 2601, Australia

Received 18 December 1985

Abstract: Frequency-domain bounds are computed for the approximation error involved in the optimum Hankel-norm approximation procedure. The bounds differ from those arising in the unweighted optimal Hankel-norm approximation procedure by a multiplicative factor depending solely on the weighting function. This factor becomes unity when the weighting function is constant.

Keywords: Approximation theory, Model reduction, System-order reduction.

1. Introduction

In [1], a method was given for optimal Hankel-norm approximation with frequency weighting. In a broad sense, this method stands in the same relation to frequency-weighted balanced approximation [2] as unweighted optimal Hankel-norm approximation stands in relation to unweighted balanced approximation, see, e.g. [3].

No bounds were given in [1] on the frequency-domain error associated with frequency-weighted optimal Hankel-norm approximation; the derivation and statement of such bounds is the main task of the present work.

In [3], frequency-domain bounds are given for unweighted optimal Hankel-norm approximations. The bounds given here for the weighted case essentially are obtained by introducing to the bounds of [3] an extra multiplicative factor which depends on the weighting function, but not on the object being approximated. This extra factor will be unity if the weighting function is constant. (Thus, as one would hope, the unweighted results are obtained.) The extra factor will tend to be large when there is significant variation between the maximum and minimum amplitudes of the weighting function along the frequency axis.

2. Review of frequency-weighted optimal Hankel-norm approximation

Let $C(s)$ be a real rational transfer function of degree n. We shall restrict attention to the scalar case for convenience: the calculation of the approximation bounds will easily extend to the matrix case. Let us suppose further that $C(s)$ is stable, i.e. all poles lie in $\operatorname{Re}(s) < 0$, and suppose $\hat{G}(s)$ is a stable, minimum-phase (strict left-half-plane zeros) rational transfer function with $G(\infty) \neq 0$. Let us set

$$\check{G}(s) = \hat{G}(-s). \tag{2.1}$$

In [1], a procedure is given for finding a stable $\hat{C}(s)$ of degree k where $k < n$, such that

$$\|(C - \hat{C})\check{G}\|_{\mathrm{H}}$$

is minimized. The procedure can be summarized as follows.

Let X be a degree-k Hankel-norm approximation of $C\check{G}$. (We permit X to be nonstrictly proper, i.e. $X(\infty) \neq 0$, if this makes $\| C\check{G} - X \|_\infty$ smaller. Note that [3] also explains the construction of an algorithm for finding X.) Define

$$\hat{C} = [X\check{G}^{-1}]_- \tag{2.2}$$

where $[Z]_-$ denotes taking the constant part and strictly proper, stable part of Z. Then, as shown in [1],

$$\hat{C} \text{ has degree } k \text{ and is stable,} \tag{2.3a}$$

$$\|[C - \hat{C}]\check{G}\|_{\mathrm{H}} \text{ is minimized over all stable } \hat{C} \text{ of degree } k. \tag{2.3b}$$

3. Construction of a bound

Let $[Z]_+$ denote the strictly proper, unstable part of Z. Let $E = [C\check{G}]_- - X = [(C - C)\check{G}]_-$. Bounds on $\| E \|_\infty$ are available from [3] in terms of the Hankel singular values σ_i of $[C\check{G}]_-$, thus

$$\| E \|_\infty \leqslant \sigma_{k+1} + \cdots + \sigma_n. \tag{3.1}$$

Now

$$\|[C - \hat{C}]\hat{G}\|_\infty = \|[C - \hat{C}]\check{G}\|_\infty \leqslant \| E \|_\infty + \|[(C - \hat{C})\check{G}]_+\|_\infty = \| E \|_\infty + \|[(E\check{G}^{-1})_-\check{G}]_+\|_\infty. \tag{3.2}$$

To simplify the calculation bounding the second term in (3.2), let us assume temporarily that we are working in the discrete-time domain, so that stability corresponds to poles lying in $|z| < 1$. (By setting $z = (1 + s)/(1 - s)$, we can map s-domain quantities into z-domain quantities.) Then we can establish several simple results:

Lemma 1. *Let $E(z) = \Sigma_{i=0}^{\infty} E_i z^{-i}$ be rational with poles in $|z| < 1$, let $\check{G} = \Sigma_{i=-\infty}^{0} H_i z^{-i}$ with poles and zeros in $|z| > 1$. Then*

$$\|[E\check{G}^{-1}]_-\|_2 \leqslant \|z^{-1}E\|_{\mathrm{H}} \|\check{G}^{-1}\|_2 \leqslant \| E \|_\infty \|\check{G}^{-1}\|_2. \tag{3.3}$$

Proof. Evidently $\|[E\check{G}^{-1}]_-\|_2 = \| z^{-1}[E\check{G}^{-1}]_- \|_2 = \|[(z^{-1}E)\check{G}^{-1}]_-\|_2$. By a standard property of the Hankel norm [4, eq. (2.6)], $\|[(z^{-1}E)\check{G}^{-1}]_-\|_2 \leqslant \| z^{-1}E \|_{\mathrm{H}} \| \check{G}^{-1} \|_2$. The second inequality follows because $\| z^{-1}E \|_{\mathrm{H}} \leqslant \| z^{-1}E \|_\infty = \| E \|_\infty$, see [4].

Lemma 2. *With quantities as above,*

$$\|[(E\check{G}^{-1})_-\check{G}]_-\|_2 \leqslant \|\check{G}\|_{\mathrm{H}} \|[E\check{G}^{-1}]_-\|_2. \tag{3.4}$$

Proof. Let $(E\check{G}^{-1})_- = \Sigma_{i=0}^{\infty} A_i z^{-i}$. Let $\check{G} = \Sigma_{i=0}^{\infty} G_i z^i$ and observe that $\check{G}(z) = \Sigma_{i=0}^{\infty} G_i z^{-i}$, since $|\check{G}| = |\hat{G}|$, and poles and zeros of $\check{G}$ and $\hat{G}$ are the reciprocals of one another. Now

$$[(E\check{G}^{-1})_-\check{G}]_+ = \left[\sum_{i=0}^{\infty} A_i z^{-i} \sum_{k=0}^{\infty} G_k z^k \right]_+$$

$$= \text{strictly positive powers of} \left[\sum_{i=0}^{\infty} A_i z^{-i} \sum_{k=0}^{\infty} G_k z^k \right].$$

Evidently

$$\|[(E\check{G}^{-1})_-\check{G}]_+\|_2 = \left\| \text{strictly negative powers of} \left[\sum_{i=0}^{\infty} A_i z^{i} \sum_{k=0}^{\infty} G_k z^{-k} \right] \right\|_2$$

and by the standard property of the Hankel norm, we then have

$$\left\| [(E\check{G}^{-1})_{-}\check{G}]_{+} \right\|_2 \leqslant \left\| \sum_{k=0}^{\infty} G_k z^{-k} \right\|_{\mathrm{H}} \left\| \sum_{i=0}^{\infty} A_i z^i \right\|_2 = \|\hat{G}\|_{\mathrm{H}} \|[E\check{G}^{-1}]_{-}\|_2.$$

Observe that Lemmas 1, 2 together imply that

$$\left\| [(E\check{G}^{-1})_{-}\check{G}]_{+} \right\|_2 \leqslant \|z^{-1}E\|_{\mathrm{H}} \|\check{G}^{-1}\|_2 \|\hat{G}\|_{\mathrm{H}} \leqslant \|E\|_{\infty} \|\check{G}^{-1}\|_2 \|\hat{G}\|_{\mathrm{H}}. \tag{3.5}$$

However, this is of no immediate help in (3.2), since there is no obvious relation between L^2 and L^∞ bounds. In pursuit of such a relation, we shall use the following lemma, a proof of which is in the Appendix.

Lemma 3. *Let* $M(z) = \alpha_0 + \sum_{i=1}^{N} \alpha_i z(1-\beta_i z)^{-1}$, $|\beta_i| < 1$. *Suppose that the* β_i *are prescribed and distinct, the* α_i *are unknown,* $M(z)$ *is real rational and that* $\|M\|_2^2 \leqslant k$. *Then*

$$\|M\|_{\infty}^2 \leqslant k\left[1 + \|m'(\mathrm{e}^{\mathrm{j}\omega})B^{-1}m(\mathrm{e}^{-\mathrm{j}\omega})\|_{\infty}\right] \tag{3.6}$$

where

$$B = \left[\frac{1}{1-\beta_i^*\beta_j}\right] m'(\mathrm{e}^{\mathrm{j}\omega}) = \left[\frac{1}{\mathrm{e}^{-\mathrm{j}\omega}-\beta_1}, \ldots, \frac{1}{\mathrm{e}^{-\mathrm{j}\omega}-\beta_N}\right]. \tag{3.7}$$

In case $\alpha_0 = 0$, (3.5) *is replaced by*

$$\|M\|_{\infty}^2 \leqslant k\|m'(\mathrm{e}^{\mathrm{j}\omega})B^{-1}m(\mathrm{e}^{-\mathrm{j}\omega})\|_{\infty}. \tag{3.8}$$

The lemma applies as follows. The poles of $\check{G}$ are known, and it is not hard to see that they are identical with the poles of $[(E\check{G}^{-1})_{-}\check{G}]_{+}$. These poles all lie in $|z| > 1$. Provided they are distinct, Lemma 3 applies with α_0 zero to yield

$$\left\| [(E\check{G}^{-1})_{-}\check{G}]_{+} \right\|_{\infty} \leqslant \mu(\check{G}) \left\| (E\check{G}^{-1})_{-}\check{G}]_{+} \right\|_2 \tag{3.9}$$

where μ^2 is the overbound on the ratio $\|\check{G}\|_{\infty}^2 / \|\check{G}\|_2^2$, computable for $\check{G}$ and depending only on the poles of $\check{G}$, as described in Lemma 3. There follows from (3.2), (3.5) and (3.9) the overbounds we have been seeking:

$$\|[C-\hat{C}]\hat{G}\|_{\infty} \leqslant \|E\|_{\infty} + \mu\|z^{-1}E\|_{\mathrm{H}}\|\hat{G}^{-1}\|_2\|\hat{G}\|_{\mathrm{H}} \tag{3.10a}$$

$$\leqslant \left[1 + \mu(\check{G})\|\hat{G}^{-1}\|_2\|\hat{G}\|_{\mathrm{H}}\right]\|E\|_{\infty}. \tag{3.10b}$$

(Observe that $\|\check{G}\|_2 = \|\hat{G}\|_2$.) While (3.10a) is the tighter bound, (3.10b) may be more useful in that it separates the error bound into a product of two parts, the 'normal' bound which is the bound on $\|E\|_{\infty}$, and the second part which depends *solely* on the weighting. Notice that when the weighting is constant, $\|\hat{G}\|_{\mathrm{H}}$ will be zero, and there is no additional cost in the approximation procedure which gets reflected in the larger bound.

Notice also that the product $\|\hat{G}^{-1}\|_2\|\hat{G}\|_{\mathrm{H}}$ is a form of condition number; the greater the frequency variation in $|\hat{G}|$, the greater this product is likely to be.

Lemma 3 is restricted in the requirement that $M(z)$, equivalently $\check{G}(z)$, have distinct poles. Doubtless it could be extended to cope with multiple poles. Alternatively, should one encounter $\check{G}(z)$ with multiple poles, one could vary one very slightly.

Finally, we remark that it is possible to characterize $\mu(\check{G})$, $\|\hat{G}^{-1}\|_2$ and $\|\hat{G}\|_{\mathrm{H}}$ in terms of s-domain quantities, rather than z-domain quantities. There seems little numerical benefit in doing so.

4. Examples

To illustrate the computation of the quantity $[1+\mu(\check{G})\,\|\hat{G}^{-1}\|_2\,\|\hat{G}\|_H]$, let us take the s-domain

$$\hat{G}(s)=(s^2+2s+1)(s^2+2\alpha s+1)^{-1},\quad \alpha<1,$$

which under the bilinear transformation $s=(z-1)(z+1)^{-1}$, becomes

$$\hat{G}(z)=\frac{1}{\frac{1}{2}(1+\alpha)+\frac{1}{2}(1-\alpha)z^{-2}}.$$

In Appendix B, we set out calculation leading to

$$\mu(\check{G})=\sqrt{\frac{2}{\alpha}},\qquad \|\hat{G}^{-1}\|_2=\sqrt{\frac{1+\alpha^2}{2}},\qquad \|\hat{G}\|_H=\frac{1-\alpha}{2\alpha}.$$

With $\alpha=0.1$, observe that $G(j0)=G(j\infty)=1$ while $G(j1)=10$. Thus there is a weighting in the vicinity of $\omega=1$ of a factor of 10. With this α, one has $\mu(\check{G})\,\|\hat{G}^{-1}\|_2\,\|\hat{G}\|_H=14.3$. With $\alpha=0.01$, the corresponding figure is 500.

Consider now

$$C(s)=\frac{(s+1)(s+3)}{(s+2)(s+4)}.$$

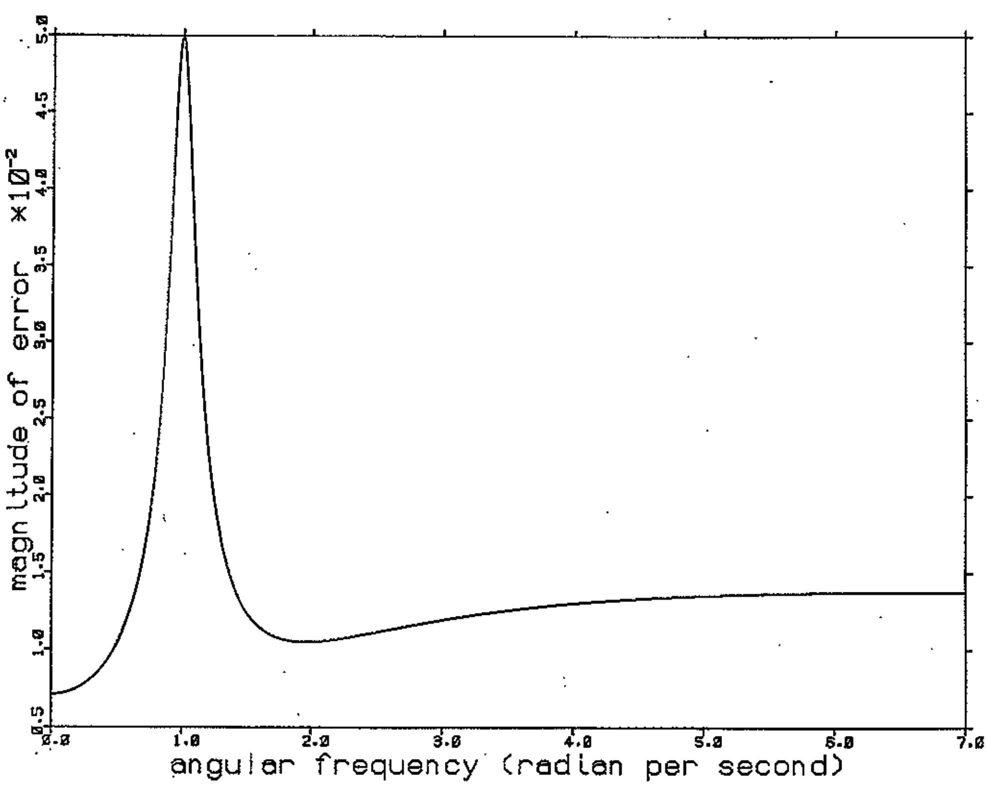

Fig. 1. Weighted error for first example.

With $\hat{G}(s) = (s^2 + 2s + 1)(s^2 + 2\alpha s + 1)^{-1}$, $\alpha = 0.1$, the procedure of Section 3 leads to a first-order weighted approximation

$$\hat{C}(s) = \frac{0.9867s + 1.1262}{s + 2.9475}.$$

The smallest singular value of $(C\check{G})_-$ is 0.0133 and the bound from the previous section is given by $\|(C - \hat{C})\hat{G}\|_\infty \leqslant 14.3 \times 0.0133 = 0.1902$. Actual evaluation for the functions themselves shows that $\|(C - \hat{C})\hat{G}\|_\infty = 0.050$.

Figures 1 and 2 show as a function of frequency the magnitude of the weighted error $(C - \hat{C})\hat{G}$ and unweighted error $C - \hat{C}$. We remark that if $\hat{C}$ were determined with no weighting, i.e. $\hat{G} = 1$, then $C - \hat{C}$ will turn out to have constant magnitude. The effect of the higher weighting around $\omega = 1$ is clear from Figure 2.

As a second example, adopt the same $\hat{G}$, but

$$C(s) = \frac{(s^2 + 0.2s + 1.01)(s^2 + 0.2s + 9.01)}{(s^2 + 0.2s + 4.04)(s^2 + 0.2s + 16.02)}.$$

With the same weighting function $\hat{G}$, the third-order approximation is

$$C(s) = \frac{3.5317s^3 - 3.6926s^2 + 51.9880s - 79.2582}{s^3 + 4.9209s^2 + 16.3429s + 77.1334}$$
$$= \frac{3.5317(s^2 + 0.4182s + 15.332)(s - 1.4637)}{(s^2 + 0.0828s + 15.9425)(s + 4.8382)}.$$

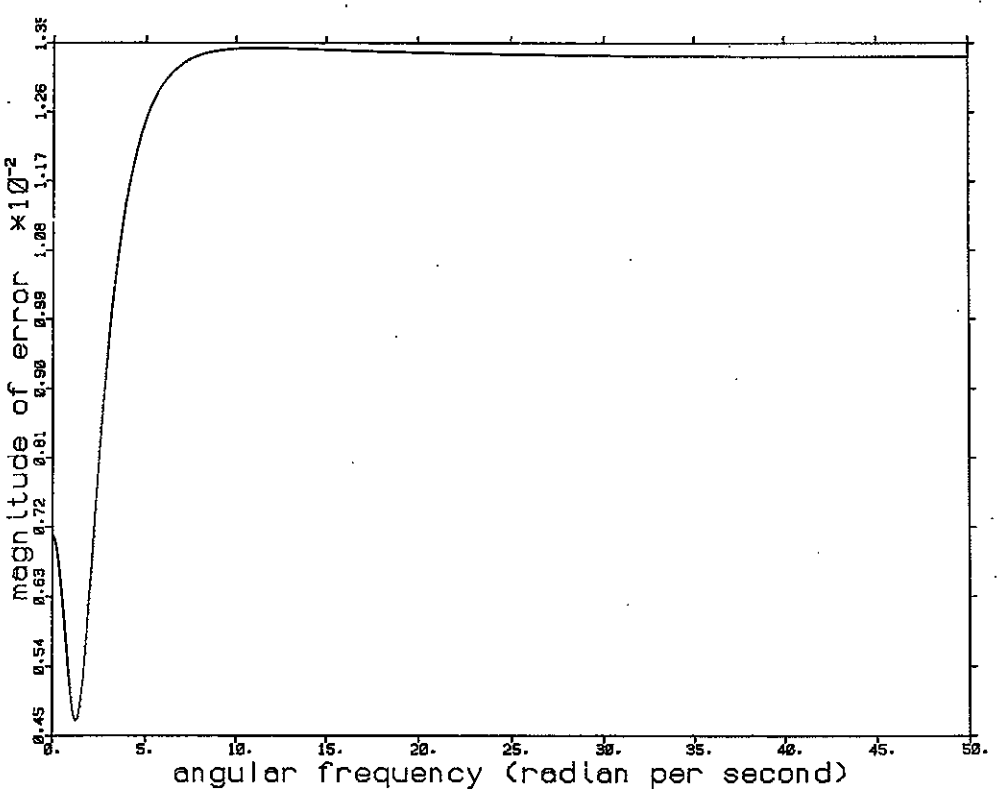

Fig. 2. Unweighted error for first example.

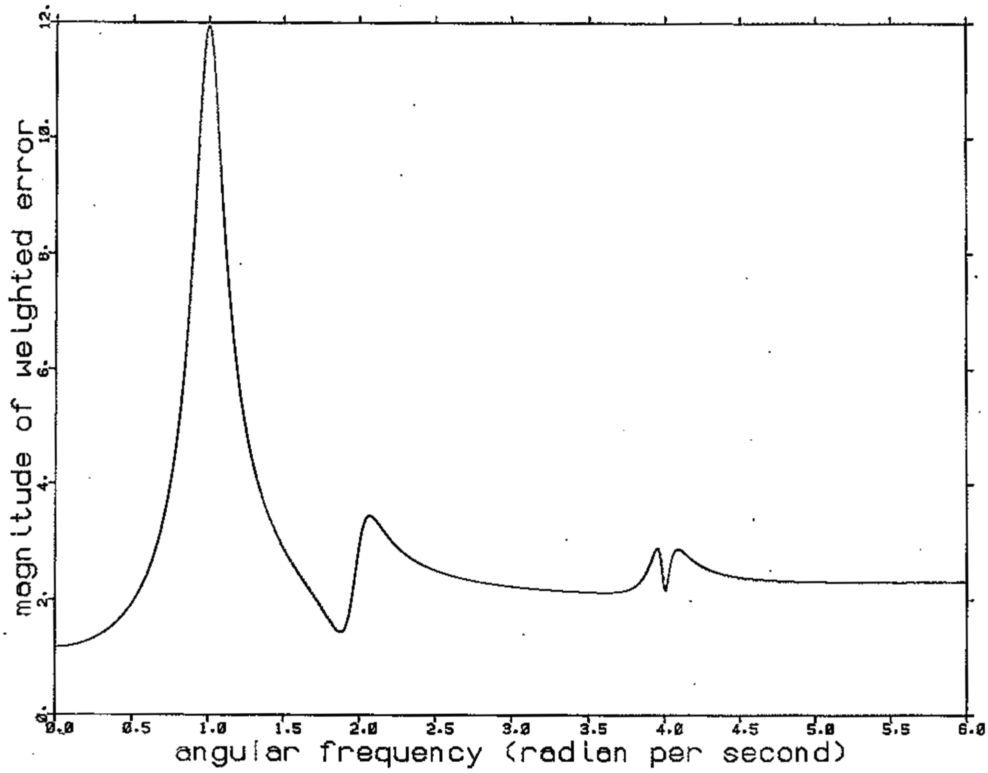

Fig. 3. Weighted error for second example.

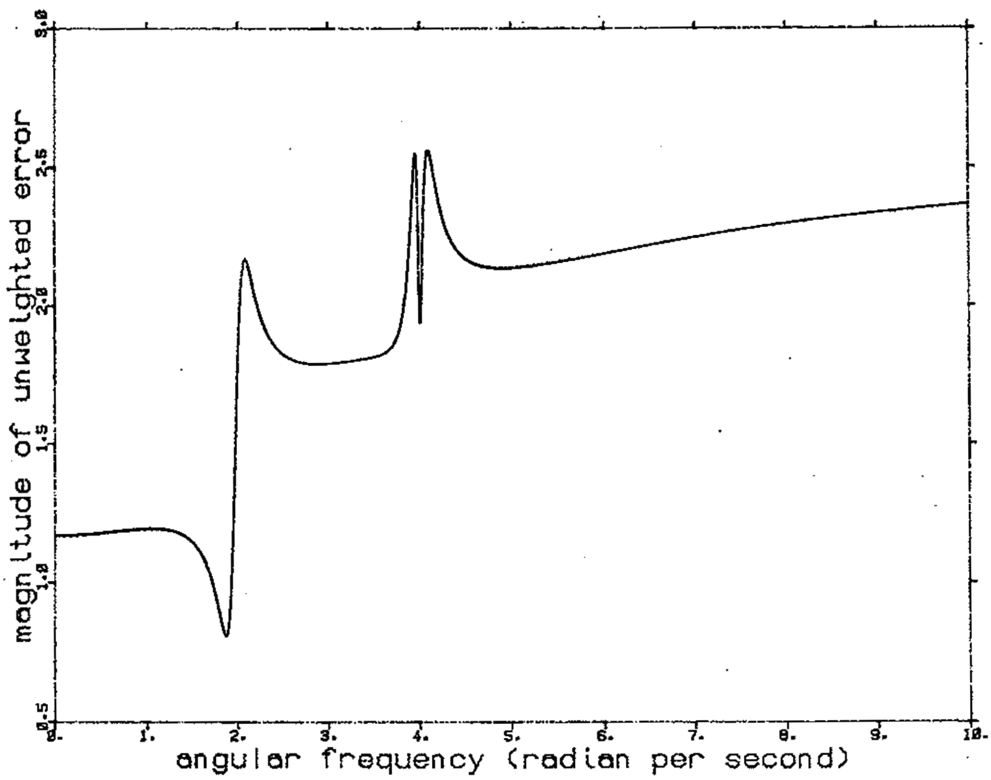

Fig. 4. Unweighted error for second example.

The smallest singular value of $(C\check{G})_-$ is 2.5317 and the bound becomes $\|(C-\hat{C})\hat{G}\|_\infty \leqslant 14.3 \times 2.5317 = 36.203$. Direct evaluation yields $\|(C-\hat{C})\hat{G}\|_\infty = 11.9411$.

Figures 3 and 4 show the weighted and unweighted errors as a function of frequency.

5. Conclusions and remarks

The prime task of this work has been to state frequency-domain bounds for frequency-weighted optimal Hankel-norm approximation. What of the utility of these results? It is argued in [2] that the practically important problem of approximating a high-order feedback controller by a low-order controller is best viewed as a weighted L_∞ approximation problem. Because of the difficulty of solving an optimum L_∞ approximation problem, the alternative is advanced in [2] of doing frequency-weighted balanced approximation and in [1] of doing frequency-weighted Hankel-norm approximation. Now when one examines the particular weighting function appropriate to the controller reduction problem, one will generally find that it is strictly proper, i.e. takes the value zero at infinite frequency. This in turn means that the L_2-norm of its inverse will be infinite, and the bounds of this paper will be useless. One way out of this difficulty would be to modify the weighting function at high frequencies so that it approaches a nonzero value for infinite frequency.

Frequency-domain bounds are not available for the error in frequency-weighted balanced approximation. It is possible to connect the Hankel-norm and balanced approximation bounds in the unweighted case [3], and it is interesting to speculate whether this relation could be extended to the weighted case, to allow derivation of bounds for frequency-weighted balanced approximation.

Appendix A

Proof of Lemma 3. Before proving Lemma 3 itself, we require a preliminary result.

Lemma A. *Let x, y be vectors in C^n and let A be a positive definite hermitian matrix. Let k be positive real. Then*

$$\operatorname*{Max}_x |x^*y|^2 \quad \text{subject to } x^*Ax \leqslant k$$

is solved by $x = [A^{-1}y/\sqrt{(y^*A^{-1}y)}]\sqrt{k}$ *and is* $ky^*A^{-1}y$.

Proof. It is obvious that at the optimum, $x^*Ax = k$. Set $\bar{x} = A^{1/2}x$. Then we are required to maximize

$$x^*yy^*x = \bar{x}^*(A^{-1/2})^*yy^*A^{-1/2}\bar{x}$$

subject to $\bar{x}^*\bar{x} = k$. The maximum is

$$k\lambda_{\max}\left[(A^{-1/2})^*yy^*A^{-1/2}\right] = ky^*A^{-1/2}A^{-1/2*}y = ky^*A^{-1}y.$$

The associated $\bar{x}$ is

$$\bar{x} = \frac{(A^{-1/2})^*y}{\sqrt{(y^*A^{-1}y)}}\sqrt{k} \quad \text{or} \quad x = \frac{A^{-1}y}{\sqrt{(y^*A^{-1}y)}}\sqrt{k}.$$

Now we turn to the proof of Lemma 3. With

$$M(z) = \alpha_0 + \sum_{i=1}^{N} \alpha_i z(1-\beta_i z)^{-1} = \alpha_0 + \sum_{l=1}^{\infty}\left(\sum_{i=1}^{N}\alpha_i\beta_i^{l-1}\right)z^l,$$

it follows that

$$\| M \|_2^2 = |\alpha_0|^2 + \sum_{i,j} \sum_{l=1}^{\infty} \alpha_i^* \alpha_j (\beta_i^*)^{l-1} \beta_j^{l-1}$$

$$= |\alpha_0|^2 + \sum_{i,j} \alpha_i^* \alpha_j \frac{1}{1 - \beta_i^* \beta_j} = [\alpha_0^*, \alpha_1^*, \ldots, \alpha_N^*][1 + B][\alpha_0, \alpha_1, \ldots, \alpha_N]'.$$

Also

$$\| M \|_\infty^2 = \sup_\omega \left| [\alpha_0^*, \alpha_1^*, \ldots, \alpha_N^*] \left[1, \frac{1}{e^{j\omega} - \beta_1^*}, \ldots, \frac{1}{e^{j\omega} - \beta_N^*} \right]' \right|^2.$$

Set $a = [\alpha_0, \alpha_1, \ldots, \alpha_N]'$, and $p(j\omega) = [1, m'(e^{j\omega})]'$. Let us seek

$$\max_a \max_\omega a^* p(e^{-j\omega}) p'(e^{j\omega}) a$$

subject to $a^*[1 + B]a \leqslant k$. It is easily verified using Lemma A that for any fixed ω,

$$\max_a a^* p(e^{-j\omega}) p'(e^{j\omega}) a$$

subject to $a^*[1 + B]a \leqslant k$ is given by

$$kp'(e^{j\omega})(1 + B)^{-1} p(e^{-j\omega}) = k[1 + m'(e^{j\omega}) B^{-1} m(e^{-j\omega})].$$

It then follows that with $a^*[1 + B]a \leqslant k$,

$$\| M \|_\infty^2 \leqslant \max_a \max_\omega a^* p(e^{-j\omega}) p'(e^{j\omega}) a \leqslant k \left[1 + \max_\omega m'(e^{j\omega}) B^{-1} m(e^{-j\omega}) \right].$$

The case $\alpha_0 = 0$ follows with trivial variation.

Appendix B

Evaluation of $\mu(\check{G})$. There holds

$$\hat{G}(z) = \frac{1}{\frac{1}{2}(1 + \omega) + \frac{1}{2}(1 + \omega) z^{-2}}, \qquad \check{G}(z) = \frac{1}{\frac{1}{2}(1 + \alpha) + \frac{1}{2}(1 - \alpha) z^2}$$

and so $\beta_1, \beta_2 = \pm jr$ where $r^2 = (1 - \alpha)(1 + \alpha)^{-1}$. The matrices B and B^{-1} become

$$B = \begin{bmatrix} \frac{1}{1 - r^2} & \frac{1}{1 + r^2} \\ \frac{1}{1 + r^2} & \frac{1}{1 - r^2} \end{bmatrix}, \qquad B^{-1} = \frac{(1 - r^4)^2}{4r^2} \begin{bmatrix} \frac{1}{1 - r^2} & -\frac{1}{1 + r^2} \\ -\frac{1}{1 + r^2} & \frac{1}{1 - r^2} \end{bmatrix}$$

and thus

$$m'(e^{j\omega}) B^{-1} m(e^{-j\omega}) = \frac{(1 - r^4)^2}{4r^2} \begin{bmatrix} \frac{1}{e^{-j\omega} - jr} & \frac{1}{e^{-j\omega} + jr} \end{bmatrix} \begin{bmatrix} \frac{1}{1 - r^2} & -\frac{1}{1 + r^2} \\ -\frac{1}{1 + r^2} & \frac{1}{1 - r^2} \end{bmatrix} \begin{bmatrix} \frac{1}{e^{j\omega} + jr} \\ \frac{1}{e^{j\omega} - jr} \end{bmatrix}$$

$$= \frac{2(1 - r^2)}{1 + r^2 - 2r \cos 2\theta} = \frac{2(1 - r^4)}{1 + 2r^2 \cos 2\omega + r^4}.$$

and

$$\left\| m'(e^{j\omega}) B^{-1} m(e^{-j\omega}) \right\|_\infty = 2\frac{1+r^2}{1-r^2} = \frac{2}{\alpha}.$$

Evaluation of $\| \hat{G}^{-1} \|_2$. Observe that $\hat{G}^{-1} = \frac{1}{2}(1+\alpha) + \frac{1}{2}(1-\alpha)z^{-2}$. Hence

$$\| \hat{G}^{-1} \|_2^2 = \tfrac{1}{4}(1+\alpha)^2 + \tfrac{1}{4}(1-\alpha)^2 = \tfrac{1}{2}(1+\alpha^2).$$

Evaluation of $\| \hat{G} \|_{\mathrm{H}}$. Since $\hat{G}(z) = (2/(1+\alpha))z^2/(z^2+(1-\alpha)/(1+\alpha))$ the following is a minimal state-variable realization:

$$A = \begin{bmatrix} 0 & 1 \\ -\dfrac{1-\alpha}{1+\alpha} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} -\dfrac{2(1-\alpha)}{(1+\alpha)^2} \\ 0 \end{bmatrix}, \quad D = \frac{2}{1+\alpha}.$$

The solutions, P, Q of $P - APA' = BB'$ and $Q - A'QA = CC'$ are

$$P = \frac{(1+\alpha)^2}{4\alpha} I, \qquad Q = \frac{(1-\alpha)^2}{\alpha(1+\alpha)^2} I.$$

Then we have $\| \hat{G} \|_{\mathrm{H}} = \lambda_{\max}^{1/2}(PQ) = (1-\alpha)/2\alpha$.

References

[1] G. Latham and B.D.O. Anderson, Frequency-weighted optimal Hankel-norm approximation of stable transfer functions, *Systems Control Lett.* **5** (1985) 229–236.

[2] D. Enns, Model reduction for control system design, PhD. Thesis, Department of Electrical Engineering, Stanford University (1984).

[3] K. Glover, All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds, *Internat. J. Control* **39** (1984) 1115.

[4] S.-Y. Kung and D.W. Lin, Optimal Hankel-norm model reductions: multivariable systems, *IEEE Trans. Automat. Control* **26** (1981) 832–852.