INT. J. CONTROL, 1987, VOL. 46, NO. 5, 1783–1801

Unstable rational function approximation

PHILIP J. PARKER† and BRIAN D. O. ANDERSON†

The problem is considered of approximating a transfer function with stable and unstable poles by a lower-order transfer function with the same number of unstable poles. A method is suggested and compared with alternative known methods based on Hankel-norm approximation. Examples suggest that no one method always outperforms other methods in terms of minimizing the approximation error.

Notation

$\lVert G(s) \rVert_\infty$	L^∞-norm of $G(s)$
$\sigma_i(G)$	the ith Hankel singular value (in descending order of magnitude) of (the stable part of) $G(s)$
$\delta(G)$	McMillan degree of $G(s)$ if $G(s)$ is a rational function; or degree of $G(s)$ if $G(s)$ is a polynomial
$[G(s)]_+$	the strictly proper stable projection of $G(s)$, i.e. the stable part of a partial fraction expansion of $G(s)$
$[(G(s)]_-$	$G(s) - [G(s)]_+$, the unstable projection of $G(s)$

1. Introduction

In this paper, we examine a problem of model reduction. Our focus is on systems which have both stable and unstable poles, and we seek to use methods such as optimal Hankel-norm reduction, and balanced realization truncation (Glover 1984).

To keep the ideas simple, we shall restrict the discussion to scalar transfer functions. Thus, let $G(s)$ be a scalar transfer function with $n > 0$ poles in $\mathrm{Re}(s) < 0$ and $\alpha > 0$ poles in $\mathrm{Re}(s) > 0$. Our interest is in obtaining an approximation $\hat{G}(s)$ to $G(s)$ with $m < n$ poles in $\mathrm{Re}(s) < 0$, with α poles in $\mathrm{Re}(s) > 0$, and with $\lVert G(j\omega) - \hat{G}(j\omega) \rVert_\infty$ small, if not as small as possible. (Subsequently, we shall make remarks about the approximation of those $G(s)$ with poles on the imaginary axis.)

There is an important reason for maintaing the number of unstable poles in the approximant. By way of illustration, suppose that $G(s)$ can be made stable with unity negative feedback. Then the Nyquist diagram of $G(j\omega)$ encircles the -1 point precisely α times in a counterclockwise direction. Now, a good approximation $\hat{G}(s)$ to $G(s)$ will be one for which the Nyquist diagrams are close, and for which unity negative feedback will continue to stabilize the loop. Close Nyquist diagrams will imply that the number of encirclements of the -1 point remains at α; retention of closed-loop stability then implies $\hat{G}(s)$ must have α unstable poles.

Another constraint that may be imposed in the approximation process is that if $G(\infty) = 0$, then also $\hat{G}(\infty) = 0$. For suppose that $G(s)$ is the transfer function of a higher-order controller, and one seeks to approximate $G(s)$ by a lower-order $\hat{G}(s)$. As

Received 21 January 1987.

† Department of Systems Engineering, Research School of Physical Sciences, Australian National University, G.P.O. Box 4, Canberra, A.C.T., 2601, Australia.

noted by Vidyasagar (1985 a), stabilization is robust against singular perturbations in the plant just when $G(s)$ is strictly proper.

Before outlining the contents of the paper, let us note two approaches suggested by Glover (1984) to the approximation of an unstable $G(s)$. The first method ('Glover 1') proceeds by decomposing $G(s)$ as

$$G(s) = G_+(s) + G_-(s) \tag{1.1}$$

with all poles of $G_+(s)$ in Re $[s] < 0$ and poles of $G_-(s)$ in Re $[s] > 0$. Then one performs an optimal Hankel-norm approximation of $G_+(s)$, to obtain $\hat{G}_+(s)$. Finally, one sets

$$\hat{G}_1(s) = \hat{G}_+(s) + G_-(s) \tag{1.2}$$

Thus $\hat{G}(s)$ approximates the stable part and copies the unstable part of $G(s)$. The second method ('Glover 2') involves a two-step procedure. First, one constructs $\hat{G}_+(s)$ as above. It follows from Theorem 7.2 of Glover (1984) that

$$G_+(s) = \hat{G}_+(s) + F(s) + \sigma U(s) \tag{1.3}$$

where $F(\cdot)$ has all poles in Re $(s) > 0$, $|U(j\omega)| = 1$ and σ is a Hankel singular value of G. Hence

$$G(s) = \hat{G}_+(s) + [F(s) + G_-(s)] + \sigma U(s) \tag{1.4}$$

Next, one calculates $\hat{G}_-(-s)$, of degree α, an optimal Hankel-norm approximation of $F(-s) + G_-(-s)$. The Glover 2 approximation is defined as

$$\hat{G}_2(s) = \hat{G}_+(s) + \hat{G}_-(s) \tag{1.5}$$

In comparison with Glover 1, where no change is made to the unstable part, the Glover 2 method modifies the unstable part by taking into account the error made in approximating the stable part—while not changing the number of unstable poles.

Note that when $\hat{G}_+(s)$ has degree one less than $G_+(s)$, $F(s)$ turns out to be constant, and then Glover 1 and Glover 2 become the same.

We have defined the Glover 1 and Glover 2 methods; comparing their error bounds is the next task to hand. Rearranging (1.4) and using (1.2) gives

$$G(s) = \hat{G}_1(s) + F(s) + \sigma U(s) \tag{1.6}$$

for Glover 1, while for Glover 2, adding and subtracting $\hat{G}_-(s)$ to the right-hand side of (1.4), then using (1.5) yields

$$G(s) = \hat{G}_2(s) + [F(s) + G_-(s) - \hat{G}_-(s)] + \sigma U(s) \tag{1.7}$$

It has been shown by Glover (1984) that

$$\|F(s)\|_\infty \leqslant \sum_{i=1}^{l} \sigma_i[F(-s)], \quad l = \delta(F) \tag{1.8}$$

and

$$\|F(s) + G_-(s) - \hat{G}_-(s)\|_\infty \leqslant \sum_{i=1}^{l} \sigma_{i+\alpha}[F(-s) + G_-(-s)]$$

$$\leqslant \sum_{i=1}^{l} \sigma_i[F(-s)] \tag{1.9}$$

and so the approximation error in (1.6) on Glover 1 can be bounded in at least as favourable a manner as the error in (1.5) for Glover 2. (This does not guarantee that the actual Glover 2 error will be less than the actual Glover 1 error.)

The paper is structured as follows. In the next section, we suggest a procedure for approximating $G(s)$ by representing $G(s)$ as a ratio of stable proper transfer functions. The question of denominator selection for these transfer functions is important. Section 3 contains examples, which, we believe, illustrate the possibility of achieving improvements to the methods of Glover 1 and 2. In § 4, we discuss in theoretical detail the error properties of the approximation procedure of § 2. Section 5 contains concluding remarks.

2. New approximation method

Approximation by considering $G(s)$ as the ratio of two stable rational transfer functions is the subject of this section. We first define the approximation method, then consider how to satisfy the necessary constraints on the approximation.

2.1. *Representation as the ratio of rational transfer functions*

A rational, proper transfer function

$$G(s)=\frac{n(s)}{d(s)} \tag{2.1}$$

where $n(s)$ and $d(s)$ are coprime polynomials and $d(s)$ has zeros in both $\operatorname{Re}(s)<0$ and $\operatorname{Re}(s)>0$, can be written as the ratio of rational functions:

$$G(s)=\frac{n(s)/p(s)}{d(s)/p(s)} \tag{2.2}$$

The polynomial $p(s)$ is constrained to have zeros only in $\operatorname{Re}(s)<0$ and to have degree $(p(s))$ equal to degree $(d(s))$. Within these constraints, one is free to choose $p(s)$. Such a representation has gained popularity because it is suitable for many problems in control theory (Vidyasagar 1985 b). This is the so-called *factorization* approach.

We shall approximate

$$H(s)=\begin{bmatrix}\dfrac{n(s)}{p(s)}\\[2ex] \dfrac{d(s)}{p(s)}\end{bmatrix}$$

by using optimal Hankel-norm reduction or balanced realization truncation, to give

$$\hat{H}(s)=\begin{bmatrix}\dfrac{\hat{n}(s)}{\hat{p}(s)}\\[2ex] \dfrac{\hat{d}(s)}{\hat{p}(s)}\end{bmatrix}$$

then form

$$\hat{G}(s)=\frac{\hat{n}(s)}{\hat{d}(s)}$$

and demonstrate that $\hat{G}(s)$ is a good approximation to $G(s)$. However, before going further we must consider the choice of the polynomial $p(s)$. The preceeding constraints on $p(s)$ still allow a wide choice for this polynomial, and this paper considers just two:

(*a*) *all-pass factorization*:

$$p(s)p(-s) = d(s)d(-s)$$

(*b*) *normalized factorization*:

$$p(s)p(-s) = n(s)n(-s) + d(s)d(-s)$$

In both cases, $p(s)$ is the Hurwitz spectral factor of $p(s)p(-s)$; that is, it has zeros only in $\mathrm{Re}\,(s) < 0$.

These are not arbitrary choices of $p(s)$. 'Normalized factorization' is a term defined by Vidyasagar (1985 b, § 7.3) in which he establishes a metric and topology for the associated factorization. The all-pass factorization—so called because $d(s)/p(s)$ is a stable all-pass—is, we believe, novel. A justification of its use is in order.

Such a justification comes in part in § 4, where we derive an elegant L^∞-error bound for approximation by the all-pass factorization. Here, we carry the reasoning further. It makes sense to have an L^∞-error bound if white noise is input to the system $G(s) = n(s)/d(s)$. (If coloured noise is the input, *frequency weighted* model reduction is more appropriate—Latham and Anderson 1985, Anderson 1986, Hung and Glover 1986.) Referring to Fig. 1 (*a*)

$$\frac{v(s)}{w(s)} = \frac{d(s)}{p(s)} \tag{2.3}$$

and we know that $d(s)/p(s)$ is an all-pass. Therefore, if w is white noise, v is also white noise, at least if a stabilizing compensator is connected around $G(s)$. We can shift our perspective of Fig. 1 (*a*) and consider v as a white noise input, and y and x as the outputs. The result is Fig. 1 (*b*).

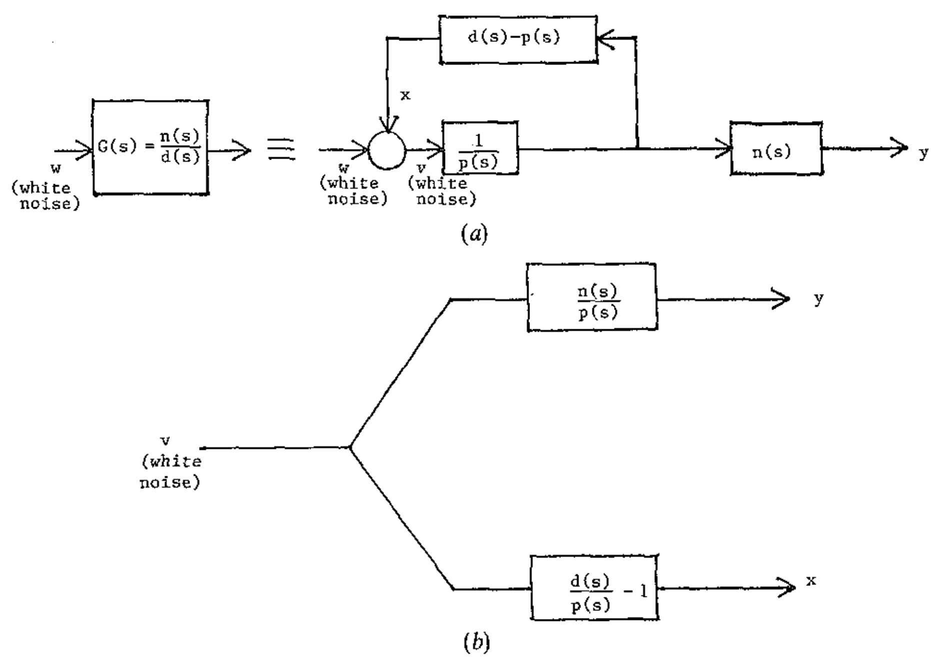

Figure 1. Justification of the all-pass factorization.

It is then valid to consider approximating

$$\begin{bmatrix} \dfrac{n(s)}{p(s)} \\ \dfrac{d(s)}{p(s)} - 1 \end{bmatrix}$$

or, alternatively, to approximate

$$H(s) = \begin{bmatrix} \dfrac{n(s)}{p(s)} \\ \dfrac{d(s)}{p(s)} \end{bmatrix}$$

and to bound the approximation $H - \hat{H}$ in the L^∞-norm because v, the input to $H(s)$, is white. This is, indeed, the path we pursue in § 4.

There is a second way in which the approximation of H might be considered as an appropriate tool for approximating G. Observe that $|H_1| = |G|$, $|H_2| = 1$. Approximating H_1 approximates the magnitude of G; approximating H_2 approximates the all-pass which equalizes the phase of H_1 to match the phase of G.

In summary, both factorization methods are motivated: the normalized factorization from the topological arguments of Vidyasagar (1985 b); and the all-pass method from the preceding arguments, the test of which better lies in simulations and the error bounds achieved.

2.2. *Scaling effects*

Introducing a gain k to $G(s)$ has an effect on the Glover approximation methods different from that on the factorization methods. If $\hat{G}(s)$ is a Glover 1 or Glover 2 approximation to $G(s)$, then $k\hat{G}(s)$ will be the corresponding approximation to $kG(s)$. However, for the factorization methods the effect is more complicated. Start with

$$kG(s) = \frac{kn(s)}{d(s)}$$

Form

$$H_k(s) = \begin{bmatrix} \dfrac{kn(s)}{p(s)} \\ \dfrac{d(s)}{p(s)} \end{bmatrix} \tag{2.4}$$

Then perform model reduction to yield

$$H_k(s) = \begin{bmatrix} \dfrac{k\hat{n}_k(s)}{\hat{p}_k(s)} \\ \dfrac{\hat{d}_k(s)}{\hat{p}_k(s)} \end{bmatrix} \tag{2.5}$$

and finally obtain

$$k\hat{G}_k(s) = \frac{k\hat{n}_k(s)}{\hat{d}_k(s)} \tag{2.6 a}$$

or

$$\hat{G}_k(s) = \frac{\hat{n}_k(s)}{\hat{d}_k(s)} \tag{2.6 b}$$

While (2.6 *a*) indicates that $k\hat{G}_k(s)$ approximates $kG(s)$, (2.6 *b*) demonstrates another link: $\hat{G}_k(s)$ approximates $G(s)$, with the particular approximation depending on the choice of k. That is, the gain k is a parameter we may select freely, giving a whole class of approximations—$\{\hat{G}_k(s):k>0\}$. We shall use this property repeatedly in this tractate.

A simple first application of this property is to show that the normalized and all-pass factorizations become the same in the limit as $k\to 0$. The normalized factorization becomes, with scaling,

$$p(s)p(-s) = k^2 n(s)n(-s) + d(s)d(-s)$$

and, as $k\to 0$, this $p(s)$ approaches that given by the all-pass method

$$p(s)p(-s) = d(s)d(-s)$$

More applications of scaling will appear in the following sections, but we now turn to one of the constraints of approximating unstable systems.

2.3. *Preservation of the number of right half-plane poles*

An approximation to an unstable system must, as we have already stated, keep the same number of right half-plane poles as the original system. For the factorization methods, the number of poles of $G(s)$ in $\mathrm{Re}\,(s)>0$ is equal to the number of zeros of $d(s)/p(s)$ in $\mathrm{Re}\,(s)>0$. Therefore, the constraint can be reformulated: $d(s)/p(s)$ and $\hat{d}_k(s)/\hat{p}_k(s)$ must have the same number of right half-plane zeros.

Using Rouché's theorem (Copson 1935), we can derive a suitable condition for this to hold. Consider the contour C in Fig. 2. By choosing r large enough, all of the right half-plane zeros of $d(s)/p(s)$ and $\hat{d}_k(s)/\hat{p}_k(s)$ lie within C. Furthermore, since $p(s)$ and $\hat{p}_k(s)$ are Hurwitz, both d/p and $\hat{d}_k/\hat{p}_k$ will be analytic on and within C. If

$$\left|\frac{\hat{d}_k(s)}{\hat{p}_k(s)} - \frac{d(s)}{p(s)}\right| < \left|\frac{d(s)}{p(s)}\right| \quad \forall\, s \text{ on } C \tag{2.7}$$

then Rouché's theorem states that $d(s)/p(s)$ and $\hat{d}_k(s)/\hat{p}_k(s)$ have the same number of

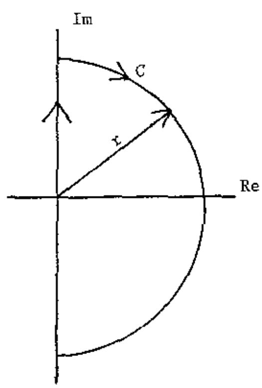

Figure 2. The contour C.

zeros within C. Since $d(s)/p(s)$ and $\hat{d}_k(s)/\hat{p}_k(s)$ are proper, they will remain bounded on C if we allow $r \to \infty$. In this case, (2.7) simplifies to

$$\left|\frac{d(j\omega)}{p(j\omega)}\right| - \left|\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)} - \frac{d(j\omega)}{p(j\omega)}\right| > 0 \quad \forall \omega \in \mathbb{R} \tag{2.8}$$

For the all-pass factorization, we can refine this condition further. Since, for the all-pass factorization

$$|d(j\omega)/p(j\omega)| = 1 \quad \forall \omega \in \mathbb{R}$$

we have

$$1 - \left|\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)} - \frac{d(j\omega)}{p(j\omega)}\right| > 0 \quad \forall \omega \in \mathbb{R} \tag{2.9}$$

Therefore, (2.8) for the normalized factorization and (2.9) for the all-pass factorization are the desired conditions for preserving the number of unstable poles.

Scaling again plays an important part here. If the model reduction violates condition (2.8) or (2.9), a new attempt with a smaller gain k could ensure that (2.8) or (2.9) is satisfied. The reasoning in outline for the all-pass case is as follows. A small k will cause kn/p to be small, but d/p will not depend on k. In performing model reduction on $H = [kn/p \quad d/p]^{\mathrm{T}}$, we should expect comparable errors to occur in approximating kn/p compared with approximating d/p. However, kn/p is small in magnitude. Therefore, we expect $|\hat{d}_k(j\omega)/\hat{p}_k(j\omega) - d(j\omega)/p(j\omega)|$ to be small, and the conditions (2.8) and (2.9) will be more easily satisfied. For the normalized factorization, the argument is not so simple, as $p(s)$ depends on k. However, we may appeal to the convergence of the two methods as $k \to 0$, as discussed in § 2.2, and the same argument will hold.

We now consider the other constraint that our approximation may have to satisfy.

2.4. *Strictly proper constraint*

As mentioned in the introduction, we shall frequently have $G(\infty) = 0$, and require that $\hat{G}(\infty) = 0$. That is, if $G(s)$ is strictly proper, $\hat{G}(s)$ must also be strictly proper.

Strict properness is not ensured by performing an optimal Hankel-norm reduction. Indeed, an optimal Hankel-norm reduction leaves one free to choose the feedthrough term, and Glover (1984) has given a particular choice of feedthrough term which reduces the L^∞-error bound. For Glover 1 and Glover 2 approximations, we achieve strict properness (at the expense of an increased L^∞-error bound) by simply setting the feedthrough term to zero. For the factorization methods, the process is slightly more detailed.

Step 1. From $G(s) = n(s)/d(s)$, form $H(s) = [n(s)/p(s) \quad d(s)/p(s)]^{\mathrm{T}}$. Since $G(s)$ is strictly proper, $H(s)$ will have a feedthrough term of $H_{\mathrm{f}} = [0 \pm 1]^{\mathrm{T}}$.

Step 2. Form $H_{\mathrm{sp}}(s) =$ strictly proper part of $H(s)$.

Step 3. Approximate $H_{\mathrm{sp}}(s)$, by optimal Hankel-norm approximation or balanced realization truncation, to yield $\hat{H}_{\mathrm{sp}}(s)$ which is constrained to be strictly proper.

Step 4. Copy the original feedthrough term H_{f} to the approximation, giving

$$H_{\mathrm{f}} + \hat{H}_{\mathrm{sp}}(s) = [\hat{n}(s)/\hat{p}(s) \quad \hat{d}(s)/\hat{p}(s)]^{\mathrm{T}}$$

Step 5. $\hat{G}(s) = \hat{n}(s)/\hat{d}(s)$.

Incorporation of a scaling gain k into this method is analogous to the non-strictly proper case outlined in § 2.2.

2.5. *Approximation of a transfer function with $j\omega$-axis poles*

To this point, we have assumed that $G(s)$ has no poles with zero real part. If this is not the case, one can write $G(s)$ as $K_1(s) + K_2(s)$ where $K_1(s)$ has no poles with zero real part, and all poles of $K_2(s)$ have zero real part. An approximation $\hat{K}_1(s)$ is formed for $K_1(s)$, and then one can define $\hat{G}(s) = \hat{K}_1(s) + K_2(s)$.

A possible alternative is to seek to use the normalized factorization method. This, however, cannot be guaranteed to yield a $\hat{G}_k(s) = \hat{n}_k(s)/\hat{d}_k(s)$ in which $\hat{d}_k(s)$ and $d(s)$ have the same number of $j\omega$-axis and right half-plane poles, and so this alternative is probably unappealing.

In this section, we have defined approximation by factorization methods—all-pass and normalized—together with extensions into scaling, preservation of right half-plane poles, strict properness and $j\omega$-axis poles. Before approaching further theoretical issues, it is appropriate to consider some examples.

3. Examples

The first example approximates a plant with transfer function

$$G(s) = \frac{500s + 3400}{s^2 + 505s + 2500} + \frac{1000}{s - 50} \tag{3.1}$$

(The poles are at -5, -500 and 50.) The magnitude and phase of $G(s)$ are plotted in Fig. 3.

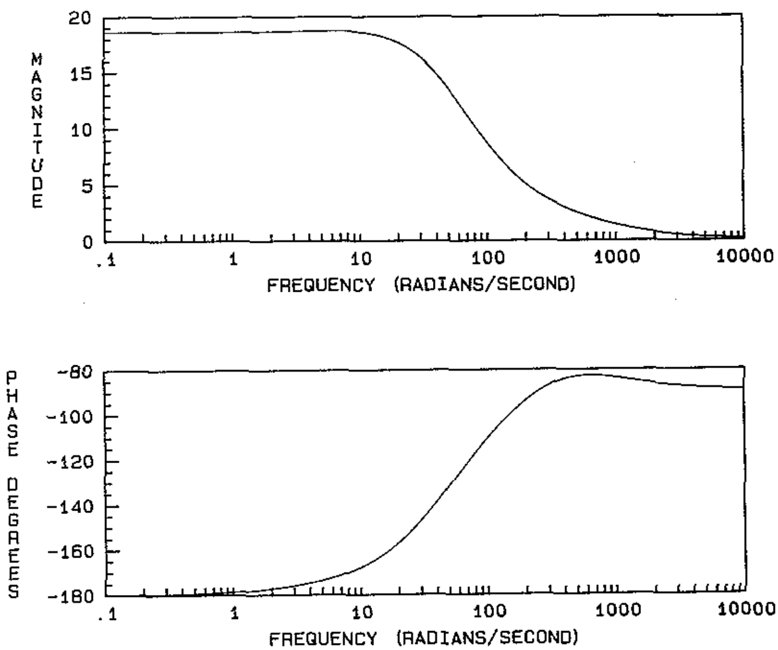

Figure 3. Example 1: original system $G(s) = \dfrac{500s + 3400}{s^2 + 505s + 2500} + \dfrac{1000}{s - 50}$.

Figure 4 compares the error magnitude (as a function of frequency) in approximating $G(s)$ by the three methods, Glover 1, all-pass and normalized factorization. In all cases, $G(s)$ is approximated by a system with one stable and one unstable pole, with strictly proper transfer functions. [Because the stable degree reduction here is one, Glover 2 and Glover 1 do not differ.] For all-pass and normalized factorizations, a value $k = 0{\cdot}1$ was used. The approximating transfer functions were as shown in Table 1, as are the maxima over all frequencies of the error magnitudes.

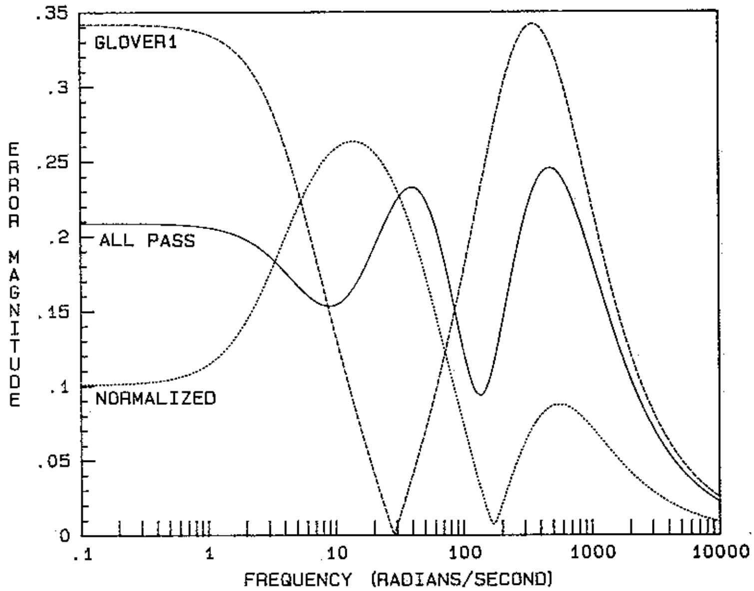

Figure 4. Example 1: comparison of approximation methods, error magnitude: $\|G(j\omega) - \hat{G}(j\omega)\|$.

Notice how the all-pass and normalized factorization methods adjust the unstable part of the approximation, which results in error reduction.

In both the all-pass and factorization methods, it is necessary that the approximation of H_k be sufficiently accurate that the number of unstable poles of G is preserved during approximation. Figure 5 depicts, for the all-pass and normalized

Method	Approximating transfer function	Error $\|G - \hat{G}\|_\infty$
Glover	$\frac{245}{s+241} + \frac{1000}{s-50}$	0·34
All-pass	$\frac{306}{s+327} + \frac{973}{s-49{\cdot}2}$	0·25
Normalized	$\frac{420}{s+403} + \frac{990}{s-50{\cdot}1}$	0·26

Table 1. Approximation of third-order transfer function with one unstable pole.

factorizations methods, the quantities

$$1-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}(j\omega)}{\hat{p}(j\omega)}\right| \quad \text{and} \quad \left|\frac{d(j\omega)}{p(j\omega)}\right|-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}(j\omega)}{\hat{p}(j\omega)}\right|$$

Their positivity for all ω is a sufficient condition for the preservation of the number of unstable poles.

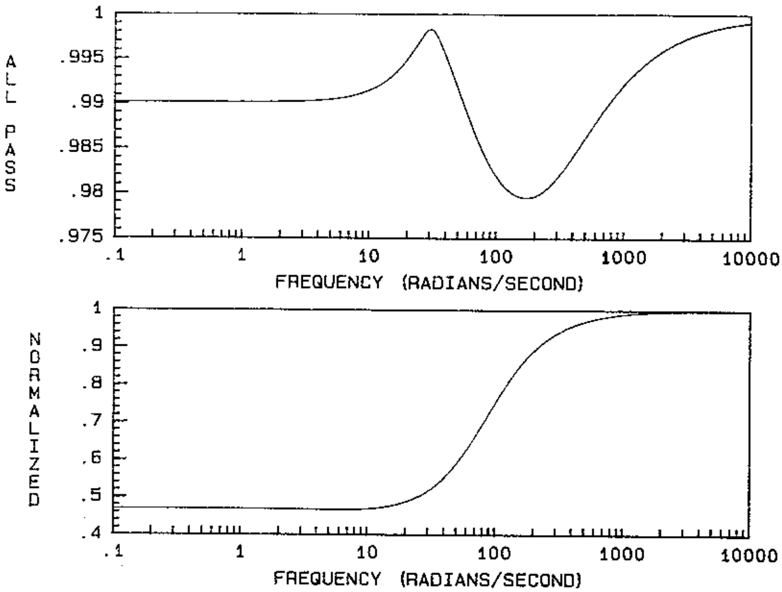

Figure 5. Example 1: preservation of unstable pole count in factorization methods: all-pass $1-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$; normalized $\left|\frac{d(j\omega)}{p(j\omega)}\right|-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$.

Figure 6 plots separately the errors

$$\left|\frac{n}{p}-\frac{\hat{n}_k}{\hat{p}_k}\right| \quad \text{and} \quad \left|\frac{d}{p}-\frac{\hat{d}_k}{\hat{p}_k}\right|$$

as a function of frequency, for the all-pass approximation method. The fact that (at least roughly) the errors track one another may explain why the error in $\|G-\hat{G}_k\|$ is kept small, smaller in fact than the maximum error of either quantity individually.

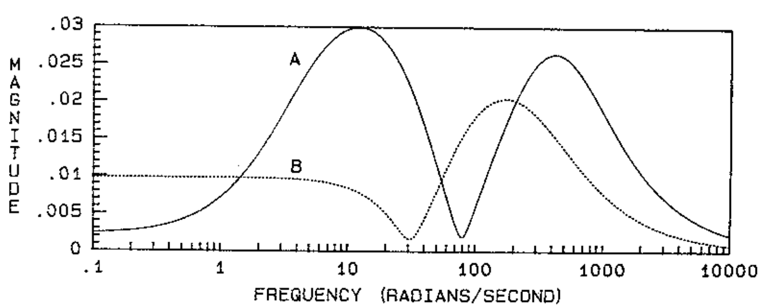

Figure 6. Example 1: factor errors in the all-pass approximation: $(A)\left|\frac{kn(j\omega)}{p(j\omega)}-\frac{k\hat{n}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$; $(B)\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$.

In a second example, we exhibit a situation where Glover 1 outperforms the all-pass and factorization methods.

The transfer function is

$$G(s) = \frac{20\,000}{s^2 + 100s + 10\,000} + \frac{50}{s + 10} + \frac{10}{s - 5} \tag{3.2}$$

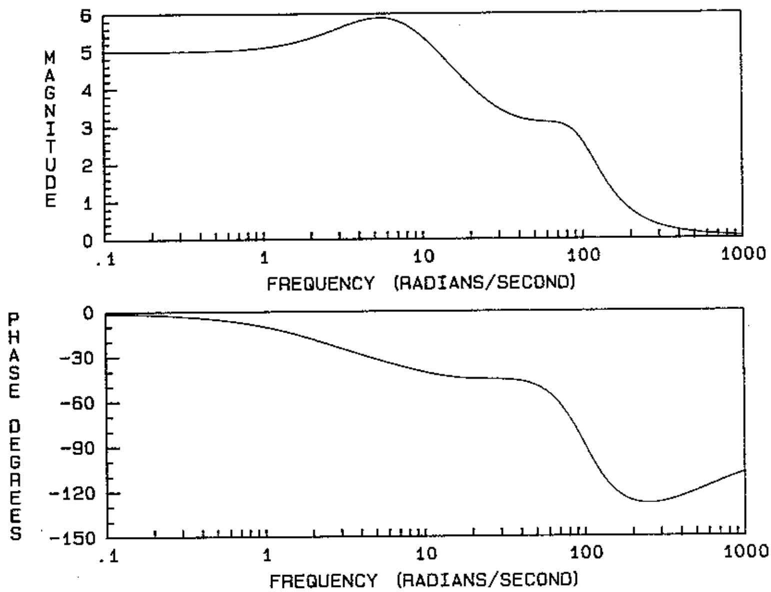

Figure 7. Example 2: original system: $G(s) = \dfrac{20\,000}{s^2 + 100s + 100\,000} + \dfrac{50}{s + 10} + \dfrac{10}{s - 5}$.

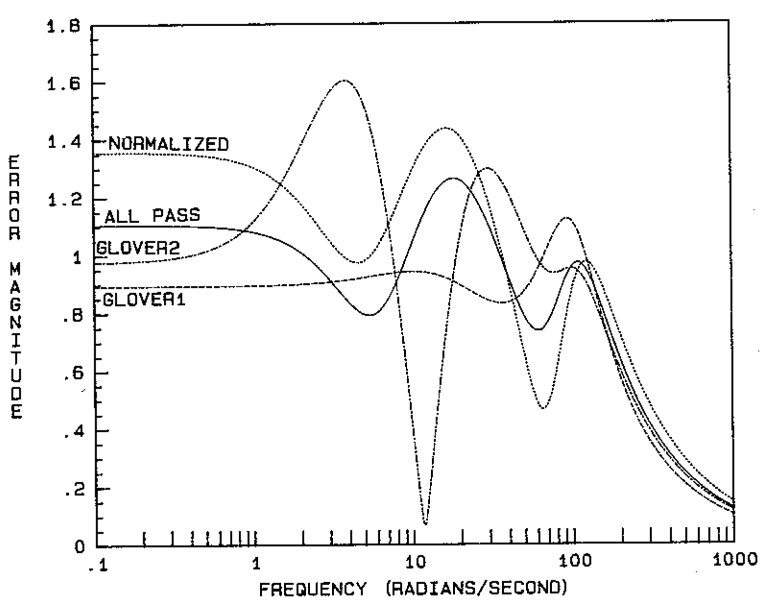

Figure 8. Example 2: comparison of approximation methods error magnitude: $|G(j\omega) - \hat{G}(j\omega)|$.

which has poles at $-50 \pm j86{\cdot}6$, -10, 5. The magnitude and phase of this transfer function are plotted in Fig. 7. A second-order approximation is determined, using Glover 1, Glover 2, all-pass and normalized factorization. Error magnitude plots are shown in Fig. 8 and Table 2 indicates both the approximation and the maximum error magnitude for each approximation; a value of $k = 0{\cdot}2$ was assumed for the two factorization methods.

Method	Transfer function $\hat{G}(s)$	$\|G-\hat{G}\|_\infty$
Glover 1	$\dfrac{151}{s+24{\cdot}7}+\dfrac{10}{s-5}$	1·12
Glover 2	$\dfrac{151}{s+24{\cdot}7}+\dfrac{26.6}{s-12{\cdot}8}$	1·60
All-pass	$\dfrac{171}{s+27{\cdot}1}+\dfrac{12{\cdot}5}{s-5{\cdot}21}$	1·26
Normalized	$\dfrac{194}{s+31{\cdot}9}+\dfrac{11{\cdot}5}{s-4{\cdot}73}$	1·44

Table 2. Approximation of a fourth-order transfer function with one unstable pole.

Figure 9 plots the quantities

$$1-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}(j\omega)}{\hat{p}(j\omega)}\right| \quad \text{and} \quad \left|\frac{d(j\omega)}{p(j\omega)}\right|-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}(j\omega)}{\hat{p}(j\omega)}\right|$$

and their positivity for all ω guarantees preservation of the unstable pole count for the new methods.

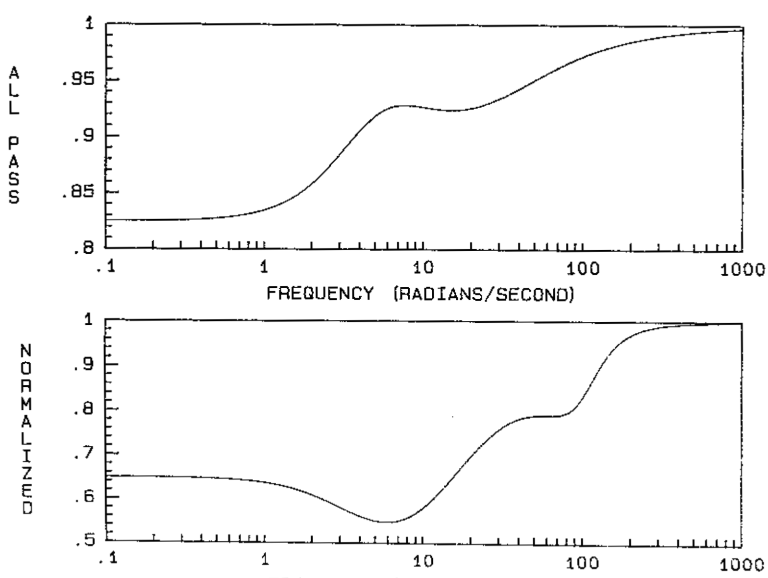

Figure 9. Example 2: preservation of unstable pole count in factorization methods: all-pass $1-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$; normalized $\left|\frac{d(j\omega)}{p(j\omega)}\right|-\left|\frac{d(j\omega)}{p(j\omega)}-\frac{\hat{d}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$.

Figure 10, illustrating the factor errors for the all-pass approximation, shows that the two errors do *not* track one another; the error in the quotient is greater than both errors.

As foreshadowed, Glover 1 outperforms the two new methods. It also outperforms Glover 2. As explained elsewhere in the paper, Glover 2 offers better error bounds than Glover 1, at least when direct feedthrough terms can be freely chosen. Error bounds are not the same as actual errors, and the direct feedthrough term cannot be freely chosen here. So one cannot expect that Glover 2 will always outperform Glover 1.

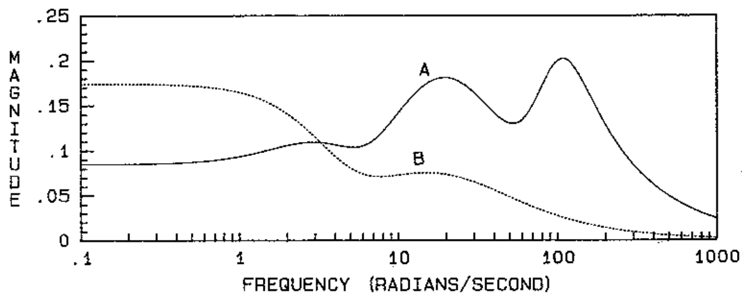

Figure 10. Example 2: factor error in the all-pass approximation: $(A)\left|\frac{kn(j\omega)}{p(j\omega)}-\frac{k\hat{n}_k(j\omega)}{\hat{p}_k(j\omega)}\right|$; $(B)\left|\frac{d(j\omega)}{p(j\omega)}-\frac{d_k(j\omega)}{p_k(j\omega)}\right|$.

4. Discussion and theoretical issues

This section derives error bounds for the all-pass factorization, Glover 1, and Glover 2 approximation methods. Within each approximation method, there is a model reduction step: in Glover 1, one must find $\hat{G}_+(s)$, a reduced-order approximation to $G_+(s)$; in Glover 2, one must find $\hat{G}_+(s)$, and then from $G_-(s)+F(s)$ find $\hat{G}_-(s)$; in the factorization methods one reduces the degree of $H_k(s)$ to get $\hat{H}_k(s)$. Before considering the error bounds, it is necessary to decide on the method of model reduction in these steps. For strictly proper approximations of a strictly proper $G(s)$, the choice is between

(1) optimal Hankel-norm reduction with a feedthrough term of zero; and

(2) balanced realization truncation.

If the strictly proper constraint does not apply, one may choose between:

(*a*) optimal Hankel-norm approximation with an optimal feedthrough term; and

(*b*) balanced realization truncation.

For Glover 2, balanced truncation is not an option because, from the optimal Hankel-norm approximation of $G_+(s)$, the associated optimal anti-causal approximation $F(s)$ forms part of the calculation of $\hat{G}_-(s)$. The balanced truncation of $G_+(s)$ gives no corresponding anti-causal $F(s)$ to use in calculating $\hat{G}_-(s)$. The choice of model reduction method affects the error bounds.

In fact, we can easily accommodate the different model reduction methods by including a multiplicative constant c in the error bounds: we define

$c = 1$ for optimal Hankel-norm approximation with an L^∞-optimal feedthrough term

$c = 2$ for optimal Hankel-norm approximation with a feedthrough term of zero (when approximating strictly proper transfer functions) or for balanced realization truncation

We are now ready to derive L^∞-error bounds for the approximation error in all methods except normalized factorization. The derivations which follow assume that neither $G(s)$ nor $H_k(s) = [kn(s)/p(s)]^T$ have repeated Hankel-singular values. Because the error bounds can be tightened if one knows that a Hankel-singular value is multiple, this is not a restrictive assumption; it merely states that our error bounds may be looser than is necessary. For the remainder of this section we assume, unless otherwise stated, that $G(s)$ has n stable and α unstable poles, and $\hat{G}(s)$ and $\hat{G}_k(s)$ have m stable and α unstable poles.

Lemma 4.1

If $\hat{G}(s)$ is a Glover 1 or Glover 2 approximation to $G(s)$, then

$$\|G - \hat{G}\|_\infty \leqslant c \sum_{i>m} \sigma_i(G) \tag{4.1}$$

Proof

For balanced realization truncation, see Glover (1984, Theorem 9.6); for optimal Hankel-norm approximation with an optimal feedthrough term, see Glover (1984, Theorem 9.7); and for strictly proper optimal Hankel-norm approximation of a strictly proper system, see Glover (1984, Theorem 9.7 and Corollary 9.3). □

Having bounded the error for Glover 1 and Glover 2, we now turn to the all-pass factorization.

Lemma 4.2

For the all-pass factorization, with the notation of (2.1) and (2.4)–(2.6 *b*), we obtain

$$\|H_k - \hat{H}_k\|_\infty \leqslant c \sum_{i>m+\alpha} \sigma_i(H_k) \tag{4.2}$$

Proof

Noting that $\delta(\hat{H}_k) = m + \alpha$, the proof is the same as for Lemma 4.1. □

Theorem 4.1

With the notation given in § 2, for the all-pass factorization:

$$\sigma_{i+\alpha}(H_k) \leqslant k\sigma_i(G) \quad i = 1, 2, \ldots, \delta[G]_+ \tag{4.3}$$

Proof

With

$$G(s) = \frac{n(s)}{d(s)}$$

we can write

$$d(s) = d_+(s)d_-(s) \tag{4.4}$$

where $d_+(s)$ is Hurwitz and $d_-(-s)$ is Hurwitz. Therefore

$$\delta(d_+) = n = \text{number of stable poles of } G(s)$$

$$\delta(d_-) = \alpha = \text{number of unstable poles of } G(s)$$

Then, for the all-pass factorization

$$p(s) = d_+(s)d_-(-s) \tag{4.5}$$

and

$$\frac{d(s)}{p(s)} = \frac{d_-(s)}{d_-(-s)} \tag{4.6}$$

which is a stable all-pass. For use below we define

$$A(s) = \frac{n(s)}{p(s)} \tag{4.7}$$

We can write, for certain polynomials $m_+(s)$ and $m_-(s)$.

$$G(s) = \frac{m_+(s)}{d_+(s)} + \frac{m_-(s)}{d_-(s)}$$

and then form the Glover 1 approximation of degree $i+\alpha-1$:

$$\hat{G}_1(s) = \frac{\hat{m}_+(s)}{\hat{d}_+(s)} + \frac{m_-(s)}{d_-(s)} \tag{4.8}$$

where

$$\frac{\hat{m}_+(s)}{\hat{d}_+(s)}$$

is an optimal Hankel-norm approximation to

$$\frac{m_+(s)}{d_+(s)}$$

with

$$\delta\left[\frac{\hat{m}_+(s)}{\hat{d}_+(s)}\right] \leqslant i-1$$

Using an extension of Nehari's theorem (see Glover 1984, Theorem 7.2), there exists $F(s)$, anti-causal, such that

$$\|G - \hat{G}_1 - F\|_\infty = \sigma_i(G) \tag{4.9}$$

Now, from (4.8), we can write

$$\hat{G}_1(s) = \frac{\tilde{n}(s)}{\hat{d}_+(s)d_-(s)}$$

and we define $\hat{A}(s)$, a particular approximation to $A(s)$:

$$\hat{A}(s) = \frac{\tilde{n}(s)}{\hat{d}_+(s)d_-(-s)} \tag{4.10}$$

Therefore

$$A(s) - \hat{A}(s) - F(s)\frac{d_-(s)}{d_-(-s)} = \left[\frac{n(s)}{d_+(s)d_-(s)} - \frac{\tilde{n}(s)}{\hat{d}_+(s)d_-(s)} - F(s)\right]\frac{d_-(s)}{d_-(-s)}$$

$$= [G(s) - \hat{G}_1(s) - F(s)]\frac{d_-(s)}{d_-(-s)} \tag{4.11}$$

Hence

$$\left\|A(s) - \hat{A}(s) - F(s)\frac{d_-(s)}{d_-(-s)}\right\|_\infty = \|G(s) - \hat{G}_1(s) - F(s)\|_\infty = \sigma_i(G) \tag{4.12}$$

where the last result comes from (4.9).

Now

$$\hat{A}(s) + F(s)\frac{d_-(s)}{d_-(-s)} = \frac{1}{d_-(-s)}\left[\frac{\tilde{n}(s)}{\hat{d}_+(s)} + F(s)d_-(s)\right]$$

and $F(s)d_-(s)$ has no poles in the left half-plane; therefore,

$$\delta\left(\left[\hat{A}(s) + F(s)\frac{d_-(s)}{d_-(-s)}\right]_+\right) = \delta(d_-(-s)) + \delta(\hat{d}_+(s)) \leqslant i + \alpha - 1 \tag{4.13}$$

To complete the proof, we must first make three definitions:

$$\hat{d}(s) = \hat{d}_+(s)d_-(s) \tag{4.14}$$

$$\hat{p}(s) = \hat{d}_+(s)d_-(-s) \tag{4.15}$$

$$\hat{H}_k(s) = \begin{bmatrix} k\hat{A}(s) + kF(s)\dfrac{d_-(s)}{d_-(-s)} \\ \dfrac{\hat{d}(s)}{\hat{p}(s)} \end{bmatrix} \tag{4.16}$$

Now $\hat{H}_k(s)$ is a particular approximation to

$$H_k(s) = \begin{bmatrix} kn(s)/p(s) \\ d(s)/p(s) \end{bmatrix}$$

and we wish to determine $\delta([\hat{H}_k]_+)$. Rewriting (4.16) gives

$$\hat{H}_k(s) = \begin{bmatrix} \dfrac{k\tilde{n}(s)}{\hat{d}_+(s)d_-(-s)} + kF(s)\dfrac{d_-(s)}{d_-(-s)} \\ \dfrac{d_-(s)}{d_-(-s)} \end{bmatrix} \tag{4.17}$$

while (4.13) and (4.17) combined show that

$$\delta[\hat{H}_k]_+ \leqslant i + \alpha - 1$$

Again using the extention of Nehari's theorem (Glover 1984, Theorem 7.2) gives

$$\sigma_{i+\alpha}(H_k) \leqslant \|H_k - \hat{H}_k\|_\infty = \left\| \begin{bmatrix} kA(s) \\ \dfrac{d(s)}{p(s)} \end{bmatrix} - \begin{bmatrix} k\hat{A}(s) + kF(s)\dfrac{d_-(s)}{d_-(-s)} \\ \dfrac{\hat{d}(s)}{\hat{p}(s)} \end{bmatrix} \right\|_\infty$$

and since

$$\frac{d(s)}{p(s)} = \frac{\hat{d}(s)}{\hat{p}(s)}$$

we get

$$\sigma_i(H_k) \leqslant \left\| \begin{bmatrix} kA(s) - k\hat{A}(s) - kF(s)\dfrac{d_-(s)}{d_-(-s)} \\ 0 \end{bmatrix} \right\|_\infty$$

Therefore, using (4.12),

$$\sigma_{i+\alpha}(H_k) \leqslant k\sigma_i(G)$$ □

The transfer function matrices $H_k(s)$ and $\hat{H}_k(s)$ are but intermediate quantities introduced in the process of approximating $G(s)$, and our main interest is in the error in the latter approximation. The following theorem relates the approximation error of $H_k(s)$ to that for $G(s)$.

Theorem 4.2

With the notation given in § 2, for the all-pass factorization, if

$$\|H_k - \hat{H}_k\|_\infty \leqslant k\varepsilon$$

then

$$\|G - \hat{G}_k\|_\infty \leqslant \frac{\varepsilon}{1 - k\varepsilon}[k\|G\|_\infty + 1] \tag{4.18}$$

Proof

$$\|H_k - \hat{H}_k\|_\infty \leqslant k\varepsilon \Rightarrow \left\| \begin{bmatrix} \dfrac{kn}{p} \\ \dfrac{d}{p} \end{bmatrix} - \begin{bmatrix} \dfrac{k\hat{n}_k}{\hat{p}_k} \\ \dfrac{\hat{d}_k}{\hat{p}_k} \end{bmatrix} \right\|_\infty \leqslant k\varepsilon \Rightarrow \left\| \frac{n}{p} - \frac{\hat{n}_k}{\hat{p}_k} \right\|_\infty \leqslant \varepsilon \tag{4.19}$$

and

$$\left\| \frac{d}{p} - \frac{\hat{d}_k}{\hat{p}_k} \right\|_\infty \leqslant k\varepsilon \tag{4.20}$$

Since $d(s)/p(s)$ is all-pass we can get, using the triangle inequality,

$$\left\| \frac{\hat{d}_k}{\hat{p}_k} \right\|_\infty \geqslant 1 - k\varepsilon \tag{4.21}$$

Now

$$\|G-\hat{G}_k\|_\infty=\left\|\frac{n}{d}-\frac{\hat{n}_k}{\hat{d}_k}\right\|_\infty=\left\|\frac{n/p}{d/p}-\frac{\hat{n}_k/\hat{p}_k}{\hat{d}_k/\hat{p}_k}\right\|_\infty=\left\|\frac{\frac{n}{p}\left(\frac{\hat{d}_k}{\hat{p}_k}-\frac{d}{p}\right)+\frac{d}{p}\left(\frac{n}{p}-\frac{\hat{n}_k}{\hat{p}_k}\right)}{\frac{d}{p}\frac{\hat{d}_k}{\hat{p}_k}}\right\|_\infty$$

Using the triangle inequality gives

$$\|G-\hat{G}_k\|_\infty\leqslant\frac{1}{\|\hat{d}_k/\hat{p}_k\|_\infty}\left\{\left\|\frac{n}{d}\right\|_\infty\left\|\frac{\hat{d}_k}{\hat{p}_k}-\frac{d}{p}\right\|_\infty+\left\|\frac{n}{p}-\frac{\hat{n}_k}{\hat{p}_k}\right\|_\infty\right\}$$

Now applying (4.19)–(4.21) gives

$$\|G-\hat{G}_k\|_\infty\leqslant\frac{1}{1-k\varepsilon}\left\{\left\|\frac{n}{d}\right\|_\infty k\varepsilon+\varepsilon\right\}$$

Using the fact that $G(s)=n(s)/d(s)$, the result is evident. □

Now we tie together Lemma 4.2 and the last two theorems, to relate the error in approximating G to its Hankel singular values.

Corollary 4.1

For the all-pass factorization method, as $k\to 0$, we obtain

$$\|G-\hat{G}_k\|_\infty\leqslant c\sum_{i>m}\sigma_i(G) \tag{4.22}$$

Proof

Lemma 4.2 and Theorem 4.1 give,

$$\|H_k-\hat{H}_k\|_\infty\leqslant kc\sum_{i>m}\sigma_i(G)$$

so we can set, in Theorem 4.2,

$$\varepsilon=\sum_{i>m}\sigma_i(G)$$

As $k\to 0$ in Theorem 4.2,

$$\|G-\hat{G}_k\|_\infty\leqslant\varepsilon=c\sum_{i>m}\sigma_i(G)$$ □

Inspection of (4.22) and (4.1) shows a remarkable fact: the error bound for the all-pass factorization (in the limit as $k\to 0$), and the error bound for the Glover 1 and Glover 2 methods are the same. Section 2.2 included a similar result, showing that as $k\to 0$ the normalized and all-pass factorizations become the same. Therefore, all four approximation methods have the same error bound as $k\to 0$. (One should remember that Glover 1 and Glover 2 are independent of k; furthermore, as k increases the error bound for the all-pass method increases.) Although the error *bounds* coincide as $k\to 0$, the *actual* approximations are, in general, different for Glover 1, Glover 2 and the factorization methods in the limit as $k\to 0$—simulations demonstrate this difference. Further, when $k>0$, even the error bounds are not the same, so that one can hardly expect the approximation error to be the same. Moreover, examples show that no one method always outperforms the other methods, in terms of securing a lower L^∞ approximation error. Therefore, for any particular $G(s)$ one cannot select *a priori* the approximation method with the least L^∞ error bound or actual L^∞ error.

5. Conclusion

Four different methods for unstable rational function approximation have formed the subject of this paper. We have shown that each gives the same error bound (in the limit as the scaling $k \to 0$), although each method yields a different approximation. Apart from considerations of L^∞ error, is there any intrinsic merit in choosing one of these four methods above the others? We believe that on occasion there is. The factorization methods of approximation fit clearly into the framework of factorization in control systems synthesis (Vidyasagar 1985 b): this theory provides results on closed-loop stabilization, sensitivity, filtering and robustness of stabilization. Among the two factorization methods, the all-pass method is the more promising because we have derived its error bound, not only in the limit as $k \to 0$, but for all $k > 0$.

Extensions to the work of this paper exist. The frequency weighted approximation (Latham and Anderson 1985, Anderson 1986, Hung and Glover 1986) is one such extension, and can easily be accommodated. Generalization of the factorization methods to multi-input–multi-output systems is also without difficulty: the polynomials $n(s)$, $d(s)$ and $p(s)$ become polynomial matrices $N(s)$, $D(s)$ and $P(s)$, with ND^{-1} a right matrix coprime factorization of $G(s)$ (Vidyasagar 1985 b). Approximation by factorization techniques has wide applicability.

References

Anderson, B. D. O., 1986, *Syst. Control Lett.*, **7**, 247.

Copson, E., 1935, *An Introduction to the Theory of Functions of a Complex Variable* (London: Oxford University Press).

Glover, K., 1984, *Int. J. Control*, **39**, 1115.

Hung, Y. S., and Glover, K., 1986, *Syst. Control Lett.*, **7**, 165.

Latham, G. A., and Anderson, B. D. O., 1985, *Syst. Control Lett.*, **5**, 229.

Vidyasagar, M., 1985 a, *Syst. Control Lett.*, **5**, 413; 1985 b, *Control Systems Synthesis: a Factorization Approach* (Cambridge, Mass.: MIT).