Stable Rational Approximation in the Context of Interpolation and Convex Optimization

Johan Karlsson and Anders Lindquist

***Abstract*— A quite comprehensive theory of analytic interpolation with degree constraint, dealing with rational interpolants with an a priori bound, has been developed in recent years. In this paper we consider the limit case when this bound is removed, and only stable interpolants with a prescribed maximum degree are sought. This leads to weighted H_2 minimization, where the interpolants are parameterized by the weights. The inverse problem of determining the weight and the interpolation points given a desired interpolant profile is considered, and a rational approximation procedure based on the theory is proposed. This provides a tool for tuning the solution to specifications. The basic idea could also be applied to the case with bounded interpolants.**

I. INTRODUCTION

Let f be a function in $H(\mathbb{D})$, the space of functions analytic in the unit disc $\mathbb{D} = \{z : |z| < 1\}$, satisfying

(i) the interpolation condition

$$f(z_k) = w_k, \quad k = 0, \dots, n, \tag{1}$$

(ii) the a priori bound $\|f\|_\infty \le \gamma$, and

(iii) the condition that f be rational of degree at most n,

where $z_0, z_1, \dots, z_n \in \mathbb{D}$ are distinct (for simplicity) and $w_0, w_1, \dots, w_n \in \mathbb{C}$. It was shown in [4] that, for each such f, there is a unique rational function $\sigma(z)$ of the form

$$\sigma(z) = \frac{p(z)}{\tau(z)}, \qquad \tau(z) := \prod_{k=0}^{n} (1 - \bar{z}_k z),$$

where $p(z)$ is a polynomial of degree n with $p(0) > 0$ and $p(z) \neq 0$ for $z \in \mathbb{D}$ such that f is the unique minimizer of the generalized entropy functional

$$-\int_{-\pi}^{\pi} |\sigma(e^{i\theta})|^2 \gamma^2 \log(1 - \gamma^{-2} |f(e^{i\theta})|^2) \frac{d\theta}{2\pi}$$

subject to the interpolation conditions (1). In fact, there is a complete parameterization of the class of all interpolants satisfying (i)–(iii) in terms of the zeros of σ, which also are the spectral zeros of f; i.e., the zeros of $\gamma^2 - f(z)f^*(z)$ located in the complement of the unit disc. It can also be shown that this parameterization is smooth, in fact a diffeomorphism [5].

This smooth parameterization in terms of spectral zeros is the center piece in the theory of analytic interpolation with degree constraints; see [3], [4] and reference therein. By tuning the spectral zeros one can obtain an interpolant that better fulfills additional design specifications. However, one of the stumbling-blocks in the application of this theory has been the lack of a systematic procedure for achieving this tuning. In fact, the relation between the spectral zeros of f and f itself is nontrivial, and how to choose the spectral zeros in order to obtain an interpolant which satisfy the given design specifications is a partly open problem.

In order to understand this problem better, we will in this paper focus on the limit case as $\gamma \to \infty$; i.e., the case when condition (ii) is removed. We shall refer to this problem – which is of considerable interest in its own right – as *stable interpolation with degree constraint*. Note that, as $\gamma \to \infty$,

$$-\gamma^2 \log(1 - \gamma^{-2} |f|^2) \to |f|^2,$$

and hence (see Proposition 2),

$$-\int_{-\pi}^{\pi} |\sigma|^2 \gamma^2 \log(1 - \gamma^{-2} |f|^2) \frac{d\theta}{2\pi} \to \int_{-\pi}^{\pi} |\sigma f|^2 \frac{d\theta}{2\pi}.$$

Consequently, the stable interpolants with degree constraint turn out to be minimizers of weighted H_2 norms. Indeed, the H_2 norm plays the same role in stable interpolation as the entropy functional does in bounded interpolation. Stable interpolation and H_2 norms are considerably easier to work with than bounded analytic interpolation and entropy functionals, but many of the concepts and ideas are similar.

In many applications, no interpolation conditions are given a priori. This allows us to use the interpolation points as additional tuning variables, available for satisfying design specifications. Such a situation occurs in a recent method of passive model reduction based on interpolation proposed in [1], [10], where however only the central solution corresponding to the choice $\sigma \equiv 1$ is considered; for a more general approach see [6]. Here the interpolation conditions should be chosen so that the approximation is as good as possible. How to do this in a systematic way is an open problem.

In this paper we address the problem on how to choose both the spectral zeros and the interpolation points in a systematic way, thus answering the more general question posed in [6] in the context when the a priori bound condition (ii) is removed. In fact, although the procedures presented in this paper are in the setting of stable interpolation, they will also give insight into both bounded analytic interpolation [4] and positive real interpolation [3].

The paper is outlined as follows. In Section II we show that the problem of stable interpolation is the limit, as the bound tend to infinity, of the bounded analytic interpolation problem stated above. In Section III we derive the basic theory for

Supported by the Swedish Research Council.

Johan Karlsson and Anders Lindquist are with the Department of Mathematics, Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden, johan.karlsson@math.kth.se, alq@math.kth.se

how all stable interpolants with a degree bound may be obtained as weighted H_2-norm minimizers. In Section IV we consider the inverse problem of H_2 minimization, and in Section V the inverse problem is used for model reduction of interpolants. The inverse problem and the model reduction procedure are closely related to the theory in [7]. A model reduction procedure where no a priori interpolation conditions are required are derived in Section VI. This is motivated by a weighed relative error bound of the approximant and gives a systematic way to choose the interpolation points. This approximation procedure is also tunable so as to give small error in selected regions. In the Appendix we describe how the corresponding quasi-convex optimization problems can be solved. Finally, in Section VII we illustrate our new approximation procedures by applying them to a simple example.

II. Bounded Interpolation and Stable Interpolation

In this section we show that the H_2 norm is the limit of a sequence of entropy functionals. From this limit, the relation between stable interpolation and bounded interpolation is established, and it is shown that some of the important concepts in the two different frameworks match.

First consider one of the main results of bounded interpolation: a complete parameterization of all interpolants with a degree bound [4]. For this, we will need two key concepts in that theory; the entropy functional

$$\mathbb{K}^{\gamma}_{|\sigma|^2}(f) = -\int_{-\pi}^{\pi} \gamma^2 |\sigma(e^{i\theta})|^2 \log(1-\gamma^{-2}|f(e^{i\theta})|^2)\frac{d\theta}{2\pi},$$

where we take $\mathbb{K}^{\gamma}_{|\sigma|^2}(f) := \infty$ for $\|f\|_\infty > \gamma$, and the co-invariant subspace

$$\mathcal{K} = \left\{ \frac{p(z)}{\tau(z)} \,:\, \tau(z) = \prod_{k=0}^{n}(1-\bar{z}_k z), p \in \mathrm{Pol}(n) \right\}. \quad (2)$$

Here $\mathrm{Pol}(n)$ denotes the set of polynomials of degree at most n, and $\{z_k\}_{k=0}^n$ are the interpolation points.

In fact, any interpolant f of degree at most n with $\|f\|_\infty \leq \gamma$ is a minimizer of $\mathbb{K}^{\gamma}_{|\sigma|^2}(f)$ subject to (1) for some $\sigma \in \mathcal{K}_0$, where

$$\mathcal{K}_0 = \{\sigma \in \mathcal{K} \,:\, \sigma(0) > 0, \sigma \text{ outer}\}.$$

Furthermore, all such interpolants are parameterized by $\sigma \in \mathcal{K}_0$. This is one of the main results for bounded interpolation in [4] and is stated more precisely as follows.

Theorem 1: Let $\{z_k\}_{k=0}^n \subset \mathbb{D}, \{w_k\}_{k=0}^n \subset \mathbb{C}$, and $\gamma \in \mathbb{R}_+$. Suppose that the Pick matrix

$$P = \left[\frac{\gamma^2 - w_k \bar{w}_\ell}{1 - z_k \bar{z}_\ell}\right]_{k,\ell=0}^{n} \quad (3)$$

is positive definite, and let σ be an arbitrary function in $\mathcal{K}_0$. Then there exists a unique pair of elements $(a,b) \in \mathcal{K}_0 \times \mathcal{K}$ such that

(i) $f(z) = b(z)/a(z) \in H^\infty$ with $\|f\|_\infty \leq \gamma$
(ii) $f(z_k) = w_k, \quad k = 0, 1, \ldots, n$, and
(iii) $|a(z)|^2 - \gamma^{-2}|b(z)|^2 = |\sigma(z)|^2$ for $z \in \mathbb{T}$,

where $\mathbb{T} := \{z : |z| = 1\}$. Conversely, any pair $(a,b) \in \mathcal{K}_0 \times \mathcal{K}$ satisfying (i) and (ii) determines, via (iii), a unique $\sigma \in \mathcal{K}_0$. Moreover, the optimization problem

$$\min \mathbb{K}^{\gamma}_{|\sigma|^2}(f) \ \text{ s.t. } \ f(z_k) = w_k, \ k = 0, \ldots, n$$

has a unique solution f that is precisely the unique f satisfying conditions (i), (ii) and (iii).

The essential content of this theorem is that the class of interpolants satisfying $\|f\|_\infty \leq \gamma$ may be parameterized in terms of the zeros of σ, and that these zeros are the same as the *spectral zeros* of f; i.e., the zeros of the spectral outer factor $w(z)$ of $w(z)w^*(z) = \gamma^2 - f(z)f^*(z)$, where $f^*(z) = \overline{f(\bar{z}^{-1})}$.

Let $\|f\| = \sqrt{< f, f >}$ denote the norm in the Hilbert space $H_2(\mathbb{D})$ with inner product

$$< f, g >= \int_{-\pi}^{\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\frac{d\theta}{2\pi},$$

and let $RH(\mathbb{D})$ denote the rational functions analytic in $\mathbb{D}$. As the bound γ tend to infinity,

$$-\gamma^2 \log(1 - \gamma^{-2}|f|) \to |f|^2.$$

Therefore, the entropy functional $\mathbb{K}^{\gamma}_{|\sigma|^2}(f)$ converge to the weighted H_2 norm $\|\sigma f\|^2$.

Proposition 2: Let $f, \sigma \in RH(\mathbb{D})$ with σ outer and $\|f\|_\infty < \infty$. Then

(i) $\mathbb{K}^{\gamma}_{|\sigma|^2}(f)$ is a non-increasing function of γ, and,
(ii) $\mathbb{K}^{\gamma}_{|\sigma|^2}(f) \to \|\sigma f\|^2$ as $\gamma \to \infty$.

Proof: It clearly suffices to consider only $\gamma \geq \|f\|_\infty$. Then the derivative of $-\gamma^2\log(1-\gamma^{-2}|f|^2)$ with respect to γ is non-positive for $|f| \leq \gamma$, and hence $\mathbb{K}^{\gamma}_{|\sigma|^2}(f)$ is non-increasing. To establish (ii), note that

$$-\gamma^2 \log(1-\gamma^{-2}|f|^2) = |f|^2 + O(\gamma^{-2}|f|^2),$$

and therefore

$$-|\sigma|^2\gamma^2 \log(1-\gamma^{-2}|f|^2) \to |\sigma f|^2$$

pointwise in $\mathbb{T}$ except for σ with poles in $\mathbb{T}$. There are two cases of importance. First, if σ has no poles in $\mathbb{T}$, or if a pole of σ coincided with a zero of f of at least the same multiplicity, then $-|\sigma|^2\gamma^2\log(1-\gamma^{-2}|f|^2)$ is bounded, and (ii) follows from bounded convergence. Secondly, if σ has a pole in $\mathbb{T}$ at a point in which f does not have a zero, then both $\mathbb{K}^{\gamma}_{|\sigma|^2}(f)$, and $\|\sigma f\|^2$ are infinite for any γ. ■

The condition $\|f\|_\infty < \infty$ is needed in Proposition 2. Otherwise, if $\|f\|_\infty = \infty$, then $\mathbb{K}^{\gamma}_{|\sigma|^2}(f)$ is infinite for any γ, while $\|\sigma f\|^2$ may be finite if σ has zeros in the poles of f on $\mathbb{T}$.

The next proposition shows that stable interpolation may be seen as the limit case of bounded interpolation when the bound γ tend to infinity.

Proposition 3: Let σ be any outer function such that the minimizer f of

$$\min \|\sigma f\| \text{ such that } f(z_k) = w_k, k = 0, \ldots, n \quad (4)$$

satisfies $\|f\|_\infty < \infty$. Let f_γ be the minimizer of

$$\min \mathbb{K}^\gamma_{|\sigma|^2}(f_\gamma) \text{ such that } f_\gamma(z_k) = w_k, k = 0, \dots, n$$

for $\gamma \in \mathbb{R}_+$ large enough so that the Pick matrix (3) is positive definite. Then $\|\sigma(f - f_\gamma)\| \to 0$ as $\gamma \to \infty$.

Proof: By Proposition 2, and since f and f_γ are minimizers of the respective functional, we have

$$\mathbb{K}^\gamma_{|\sigma|^2}(f) \geq \mathbb{K}^\gamma_{|\sigma|^2}(f_\gamma) \geq \|\sigma f_\gamma\|^2 \geq \|\sigma f\|^2.$$

Moreover, since $\mathbb{K}^\gamma_{|\sigma|^2}(f) \to \|\sigma f\|^2$ as $\gamma \to \infty$ it follows that $\|\sigma f_\gamma\|^2 \to \|\sigma f\|^2$, and hence, by Lemma 8, we have $\|\sigma(f - f_\gamma)\| \to 0$ as $\gamma \to \infty$, as claimed. ■

Note that Proposition 3 holds for any σ which is outer and not only for $\sigma \in \mathcal{K}_0$. However, if $\sigma \in \mathcal{K}_0$, then $\deg f_\gamma \leq n$ for any γ. Therefore, since $\|\sigma(f - f_\gamma)\| \to 0$ as $\gamma \to \infty$, for $\sigma \in \mathcal{K}_0$ the minimizer f of (4) will be a stable interpolant of degree at most n. We will return to this in the next section.

It is interesting to note how concepts in the two types of interpolation are related. First of all, the weighted H_2 norm plays the same role in stable interpolation as the entropy functional does in bounded interpolation. Secondly, the spectral zeros, which play an major role in degree constrained bounded interpolation, simply correspond to the poles in stable interpolation. This may be seen from (iii) in Theorem 1.

III. Rational Interpolation and H_2 minimization

In the previous section we have seen that minimizers of a specific class of H_2 norms are stable interpolants of degree at most n. This, and also the fact that this class may be parameterized by $\sigma \in \mathcal{K}_0$ can be proved using basic Hilbert space concepts. This will be done in this section.

To this end, first consider the minimization problem

$$\min \|f\| \text{ s. t. } f(z_k) = w_k, k = 0, \dots, n, \tag{5}$$

without any weight σ. Let $f_0 \in H_2(\mathbb{D})$ satisfy the interpolation condition (1). Then any $f \in H_2(\mathbb{D})$ satisfying (1) can be written as $f = f_0 + v$, where $B = \prod_{k=0}^n \frac{z_k - z}{1 - \bar{z_k} z}$ and $v \in BH_2$. Therefore, (5) is equivalent to

$$\min_{v \in BH_2} \|f_0 + v\|.$$

By the Projection Theorem (see, e.g., [8]), there exists a unique solution $f = f_0 + v$ to this optimization problem, which is orthogonal to BH_2, i.e. $f \in \mathcal{K} := H_2 \ominus BH_2$.

Conversely, if $f \in \mathcal{K}$ and $f(z_k) = w_k$, for $k = 0, \dots, n$, then f is the unique solution of (5). To see this, note that any interpolant in $H_2(\mathbb{D})$ may be written as $f + v$ where $v \in BH_2$. However, since $v \in BH_2 \perp \mathcal{K} \ni f$, we have $\|f + v\|^2 = \|f\|^2 + \|v\|^2$, and hence the minimizer is f, obtained by setting $v = 0$.

We summarize this in the following proposition.

Proposition 4: The unique minimizer of (5) belongs to $\mathcal{K}$. Conversely, if $f \in \mathcal{K}$ and $f(z_k) = w_k$, for $k = 0, \dots, n$, then f is the minimizer of (5).

Consequently, in view of (2), f is a rational function with its poles fixed in the mirror images (with respect to the unit circle) of the interpolation points. By introducing weighted norms, any interpolant with poles in prespecified points may be constructed in a similar way. In fact, the set of interpolants f of degree $\leq n$ may be parameterized in this way. One way to see this is by considering

$$\min \|\sigma f\| \text{ s. t. } f(z_k) = w_k, k = 0, \dots, n, \tag{6}$$

where $\sigma \in \mathcal{K}_0$. Since σ is invertible in $H(\mathbb{D})$, (6) is equivalent to

$$\min \|\sigma f\| \text{ s. t. } (\sigma f)(z_k) = \sigma(z_k) w_k, k = 0, \dots, n.$$

According to Proposition 4, this has the optimal solution $\sigma f = b \in \mathcal{K}$, and hence the solution of (6), $f = \frac{b}{\sigma}$, is rational of degree at most n. To see that any solution of degree at most n can be obtained in this way, note that any such interpolant f is of the form $f = \frac{b}{\sigma}, b \in \mathcal{K}, \sigma \in \mathcal{K}_0$. Since $\sigma f = b \in \mathcal{K}$ holds together with the interpolation condition (1) if and only if $\sigma(z_k) f(z_k) = \sigma(z_k) w_k$ for $k = 0, \dots, n$, f is the unique solution of (6), by Proposition 4. This proves the following proposition.

Theorem 5: Let $\sigma \in \mathcal{K}_0$. Then the unique minimizer of

$$\min \|\sigma f\| \text{ s. t. } f(z_k) = w_k,\ k = 0, \dots, n, \tag{7}$$

belong to $H(\mathbb{D})$ and is rational of a degree at most n. More precisely,

$$f = \frac{b}{\sigma} \tag{8}$$

where $b \in \mathcal{K}$ is the unique solution of the linear system of equations

$$b(z_k) = \sigma(z_k) w_k, \quad k = 0, 1, \dots, n. \tag{9}$$

Conversely, if f satisfies (8) for some $b \in \mathcal{K}$ and the interpolation condition (1), then f is the unique minimizer of (7).

In other words, the set of interpolants in $H(\mathbb{D})$ of degree at most n may be parameterized in terms of weights $\sigma \in \mathcal{K}_0$. Another way to look at this is that the poles of the minimizer (8) are specified by the zeros of σ and that the numerator $b = \beta/\tau$ is determined from the interpolation condition by solving the linear system of equations

$$\beta(z_k) = \tau(z_k)\sigma(z_k) w_k, \quad k = 0, 1, \dots, n \tag{10}$$

for the $n + 1$ coefficients $\beta_0, \beta_1, \dots, \beta_n$ of the polynomial $\beta(z)$. This is a Vandermonde system that is known to have a unique solution (as long as the interpolation point $z_o, z_1, \dots, z_n$ are distinct as here).

Note that this parameterization is not necessarily complete; i.e. it may not be injective. If, for example, $w_k = 1$ for $k = 0, \dots, n$, then there is a unique function f of degree at most n that satisfies $f(z_k) = w_k, k = 0, \dots, n$. No matter how $\sigma \in \mathcal{K}_0$ is chosen, $b = \sigma$, and hence the minimizer of (6) will be $f \equiv 1$.

IV. The Inverse Problem

In [7] we considered the *inverse problem of analytic interpolation*; i.e., the problem of choosing an entropy functional whose unique minimizer is a prespecified interpolant. In this section we will consider the counterpart of this problem for stable interpolation.

Suppose $f \in RH(\mathbb{D})$ satisfies the interpolation condition (1). Then, when does there exist $\sigma \in RH(\mathbb{D})$ which is outer such that f is the minimizer of

$$\min \|\sigma f\| \text{ s. t. } f(z_k) = w_k, k = 0, \ldots, n?$$

We refer to this as the *inverse problem of H_2 minimization*, and its solution is given in the following proposition.

Theorem 6: Let $f \in RH(\mathbb{D})$ satisfy the interpolation condition $f(z_k) = w_k, k = 0, \ldots, n$. Then f is the minimizer of

$$\min \|\sigma f\| \text{ s. t. } f(z_k) = w_k,\ k = 0, \ldots, n, \tag{11}$$

where σ is outer if and only if $\sigma f \in \mathcal{K}$. Such a σ exists if and only if f has no more than n zeros in $\mathbb{D}$.

Proof: The function f is the minimizer of (11) if and only if $g = \sigma f$ is the minimizer of

$$\min \|g\| \text{ s. t. } g(z_k) := w_k \sigma(z_k),\ k = 0, \ldots, n,$$

which, by Proposition 4, holds if and only if $\sigma f = g \in \mathcal{K}$. Such a σ only exists if f has less or equal to n zeros inside $\mathbb{D}$. To see this, first note that, if f has more than n zeros in $\mathbb{D}$, then σf has more than n zeros in $\mathbb{D}$ and can therefore not be of the form p/τ with $p \in \text{Pol}(n)$. On the other hand, if f has less or equal to n zeros in $\mathbb{D}$, then let $p = \prod(z - p_k)$ where p_k are the zeros of f, and set $\sigma := \frac{p}{f\tau}$. Then σ is outer and satisfies $\sigma f \in \mathcal{K}$. ■

Let W_f denote the set of weights σ that give f as a minimizer of (11). By Theorem 6,

$$\begin{aligned} W_f &= \{\sigma \text{ outer} : \sigma f \in \mathcal{K}\} \\ &= \left\{\sigma = \frac{p}{f\tau} : p \in \text{Pol}(n) \setminus \{0\}, \frac{p}{f} \text{ outer}\right\}, \end{aligned} \tag{12}$$

i.e., W_f may be parameterized in terms of the polynomials $p \in \text{Pol}(n)$. For the condition that pf^{-1} is outer to hold for some $p \in \text{Pol}(n)$, it is necessary that f has at most n zeros in $\mathbb{D}$. This is in accordance with Theorem 6. It is interesting to note that the dimension of W_f depends on the number of zeros of f inside $\mathbb{D}$. The more zeros f has inside $\mathbb{D}$, the more restricted is the class W_f. One extreme case is when f has no zeros inside $\mathbb{D}$. Then p could be any stable polynomial of degree n. The other extreme is when f has n zeros in $\mathbb{D}$, in which case p is uniquely determined up to a multiplicative constant.

V. Rational Approximation with Interpolation Constraints

In this section the solution of the inverse problem (Theorem 6) will be used to develop an approximation procedure for interpolants. Let $f \in RH(\mathbb{D})$ be a function satisfying the interpolation condition (1). We want to construct another function $g \in RH(\mathbb{D})$ of degree at most n satisfying the same interpolation condition such that g is as close as possible to f.

Let $\sigma \in W_f$; i.e., let σ be a weight and such that f is the minimizer of (11), and let ρ be close to σ. Then it seems reasonable that the minimizer g of the optimization problem

$$\min \|\rho g\| \text{ s. t. } g(z_k) = w_k,\ k = 0, \ldots, n, \tag{13}$$

is close to f. This is the statement of the following theorem.

Theorem 7: Let $f \in RH(\mathbb{D})$ satisfy the interpolation condition $f(z_k) = w_k, k = 0, \ldots, n$, and let $\sigma \in W_f$. Moreover, let ρ be an outer function such that

$$\left\| 1 - \left|\frac{\rho}{\sigma}\right|^2 \right\|_\infty = \epsilon, \tag{14}$$

and let g be the corresponding minimizer of (13). Then

$$\|\sigma(f - g)\|^2 \leq \frac{4\epsilon}{1 - \epsilon}\|\sigma f\|^2. \tag{15}$$

For the proof we need the following useful lemma.

Lemma 8: Let f be the minimizer of (11), and let $g \in RH(\mathbb{D})$ satisfy $g(z_k) = w_k$, for $k = 0, \ldots, n$. If $\|\sigma g\|^2 \leq (1 + \epsilon)\|\sigma f\|^2$, then

$$\|\sigma(f - g)\|^2 \leq 2\epsilon\|\sigma f\|^2.$$

Proof: From the parallelogram law we have,

$$\frac{1}{2}\left(\|\sigma f\|^2 + \|\sigma g\|^2\right) = \left\|\sigma\frac{f + g}{2}\right\|^2 + \left\|\sigma\frac{f - g}{2}\right\|^2.$$

Therefore, since $\|\sigma f\| \leq \|\sigma(f + g)/2\|$, it follows that

$$\|\sigma(f - g)\|^2 \leq 2(\|\sigma g\|^2 - \|\sigma f\|^2) \leq 2\epsilon\|\sigma f\|^2,$$

which concludes the proof of the lemma. ■

Proof of Theorem 7: Since f and g are the minimizers of the respective weighted H_2 norms, by (14), we have

$$\begin{aligned} \|\sigma f\|^2 &\leq \|\sigma g\|^2 \leq \frac{1}{1 - \epsilon}\|\rho g\| \\ &\leq \frac{1}{1 - \epsilon}\|\rho f\|^2 \leq \frac{1 + \epsilon}{1 - \epsilon}\|\sigma f\|^2. \end{aligned}$$

Therefore

$$\|\sigma(f - g)\|^2 \leq \frac{4\epsilon}{1 - \epsilon}\|\sigma f\|^2.$$

follows from Lemma 8. ■

We have shown that if $\left|\frac{\rho(z)}{\sigma(z)}\right|$ is close to 1 for $z \in \mathbb{T}$, then $\|\sigma(f - g)\|$ is small. In order to obtain a g which is both a good approximation of f and has desired degree bound there are several things to consider.

First, to achieve degree reduction, ρ should be chosen so that g has the appropriate degree bound. By Theorem 5, selecting a ρ in $\mathcal{K}_0$ will guarantee that g satisfies the condition $\deg g \leq n$. Secondly, Theorem 7 suggests various strategies for choosing the functions $\rho \in \mathcal{K}_0$ and $\sigma \in W_f$ depending on the design preferences. If a small error bound for $\|\sigma(f - g)\|$ is desired for a particular $\sigma \in W_f$, this σ should be used together with the $\rho \in \mathcal{K}_0$ that minimizes (14).

However, obtaining a small value of (14) is often more important than the choice of σ. Therefore, in general it is more natural to choose the pair $(\sigma, \rho) \in (W_f, \mathcal{K}_0)$ that minimizes ϵ. For such a pair, setting $q := \tau\rho$, we can be see from (2) and (12) that

$$\epsilon = \left\| 1 - \left| \frac{\rho}{\sigma} \right|^2 \right\|_\infty = \left\| 1 - \left| \frac{qf}{p} \right|^2 \right\|_\infty , \tag{16}$$

where $q \in \mathrm{Pol}(n)$ and $p \in \mathrm{Pol}(n) \setminus \{0\}$ needs to be chosen so that p/f is outer. It is interesting to note that (16) is independent of $\tau(z) := \prod_{k=0}^n (1 - \bar{z}_k z)$ and hence of the interpolation points $z_0, z_1, \ldots, z_n$.

Now suppose that f has ν zeros in $\mathbb{D}$; i.e., ν nonminimum-phase zeros. Then $f = \pi f_0$, where f_0 is outer (minimum phase) and π is an unstable polynomial of degree $\nu \le n$. Setting $p = \pi p_0$, our optimization problem to minimize ϵ reduces to the problem to find a pair $(p_0, q) \in \mathrm{Pol}(n-\nu) \times \mathrm{Pol}(n)$ that minimizes

$$\epsilon = \left\| 1 - \left| \frac{qf_0}{p_0} \right|^2 \right\|_\infty \tag{17}$$

for a given nonminimum-phase f_0. This is a quasi-convex optimization problem, which can be solved as described in the Appendix (see also [11]). The optimal q yields the optimal $\rho = q/\tau$. The approximant g is then obtained by solving the optimization problem (13) as described in Theorem 5.

One should note that, the more zeros f has inside $\mathbb{D}$, the smaller is the choice of p. Therefore one expects approximations of non-minimum phase plants to be worse than approximations of plants without unstable zeros.

VI. Rational Approximation

In applications where there are no a priori interpolation constraints, the choice of interpolation points serve as additional design parameters. It is then important to choose them so that a good approximation is obtained. As far as the authors know, there has previously not been any systematic procedures for this. There are some general guidelines that one could use for manual tuning. The main strategy is to chose interpolation points close to the regions of the unit circle where good fit is desired. The closer to the unit circle the points are placed, the better fit, but the smaller is the region where good fit is ensured; see [6] for further discussions on this.

As we have seen in the previous section the choice of interpolation points does not affect ϵ given by (16). However, since $\sigma = \frac{p}{f\tau}$, the weighted H_2 error bound (15) in Theorem 7 becomes

$$\left\| \frac{p}{\tau} \frac{f-g}{f} \right\|^2 \le \frac{4\epsilon}{1-\epsilon} \left\| \frac{p}{\tau} \right\|^2 ,$$

which depends on τ and hence on the choice of interpolation points. In fact, this is a weighed H_2 bound on the relative error. If a specific part of the unit circle is of particular interest, interpolation points may be placed close to that part, which gives a bound on the weighted relative error with high emphasis on that specific region. If no particular part is more important than the rest, we suggest to select τ as the outer part of p; i.e., $|\tau(z)| = |p(z)|$ for $z \in \mathbb{T}$. This gives a natural choice of interpolation points that are the mirror images of the roots of τ. Furthermore, this choice gives the relative error bound $\|(f-g)/f\| \le 4\epsilon/(1-\epsilon)$. This is summarized in the following theorem.

Theorem 9: Let p and q be polynomials of degrees at most n such that pf^{-1} is outer, and set

$$\epsilon := \left\| 1 - \left| \frac{qf}{p} \right|^2 \right\|_\infty . \tag{18}$$

Let $z_0, z_1, \ldots, z_n \in \mathbb{D}$ and let

$$g = \arg\min \|\rho g\| \text{ s. t. } g(z_k) = f(z_k), k = 0, \ldots, n,$$

where $\rho = q/\tau$ and $\tau = \prod_{k=0}^n (1 - \bar{z}_k z)$. Then

$$\left\| \frac{p}{\tau} \frac{f-g}{f} \right\|^2 \le \frac{4\epsilon}{1-\epsilon} \left\| \frac{p}{\tau} \right\|^2 . \tag{19}$$

In particular, if the interpolation points $z_0, z_1, \ldots, z_n$ are chosen so that $|\tau(z)| = |p(z)|$ for $z \in \mathbb{T}$, then

$$\left\| \frac{f-g}{f} \right\|^2_{L_2(\mathbb{T})} \le \frac{4\epsilon}{1-\epsilon}. \tag{20}$$

VII. The Computational Procedure

Next we summarize the computational procedure suggested by the theory presented above and apply it to some numerical problems.

Given a function $f \in RH(\mathbb{D})$ with at most n zeros in $\mathbb{D}$, we want to construct a function $g \in RH(\mathbb{D})$ of degree at most n that approximates f as closely as possible. We consider two versions of this problem. First we assume that f satisfies the interpolation condition (1), and we require g to satisfy the same interpolation conditions. Secondly, we relax the problem by removing the interpolation constraints.

Suppose that f has $\nu \le n$ zeros in $\mathbb{D}$. Then $f = \pi f_0$, where f_0 is minimum-phase, and π is a polynomial of degree ν with zeros in $\mathbb{D}$. The approximant g can then be determined in two steps:

(i) Solve the quasi-convex optimization problem to find a pair $(p_0, q) \in \mathrm{Pol}(n-\nu) \times \mathrm{Pol}(n)$ that minimizes (17), as outlined in the Appendix. This yields optimal ϵ, p_0 and q. Set $p := \pi p_0$.

(ii) Solve the optimization problem (13) with $\rho = q/\tau$, as described in Theorem 5. Exchanging σ for ρ in (10) we solve the Vandermonde system

$$\beta(z_k) = q(z_k) w_k, \quad k = 0, 1, \ldots, n,$$

for the $\beta \in \mathrm{Pol}(n)$, which yields

$$g = \frac{\beta}{q} \tag{21}$$

and the bound (19), where $\tau(z) := \prod_{k=0}^n (1 - \bar{z}_k z)$.

For the problem without interpolation condition, we replace step (ii) by one of the following steps.

(ii)$'$ Choose $z_0, z_1, \ldots, z_n$ arbitrarily. This yields a solution (21) and a bound (19).

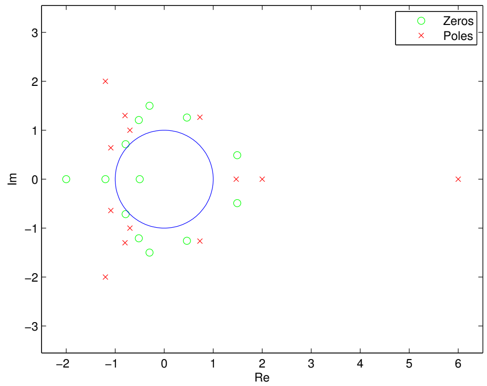

Fig. 1. Poles and zeros of f in Examples 1, 2, and 3.

(ii)$''$ Choose $z_0, z_1, \ldots, z_n$ so that τ is the outer (minimum-phase) factor of p. This yields a solution (21) and the bound (20) for the relative H_2 error.

We apply these procedures to some numerical examples.

Example 1: Let

$$f(z) = \frac{b(z)}{a(z)}$$

be the stable system of order 13 given by

$$\begin{aligned} b(z) &= 30z^{13} + 90z^{12} + 128.6z^{11} + 114.6z^{10} \\ &- 137.4z^9 - 322.3z^8 - 371.4z^7 + 10.8z^6 \\ &+ 1005.8z^5 + 2428.7z^4 + 3967.0z^3 + 4189.7z^2 \\ &+ 2800.6z + 726.2, \\ a(z) &= 4.0z^{13} - 13.4z^{12} - 44.2z^{11} - 144.5z^{10} \\ &+ 83.5z^9 + 363.7z^8 + 791.4z^7 + 340.1z^6 \\ &+ 770.7z^5 + 877.3z^4 - 93.6z^3 - 4767.8z^2 \\ &- 6349.3z - 4532.7. \end{aligned}$$

This system has one minimum-phase zero. The poles and zeros are given in Figure 1.

Consider the problem to approximate f by a function g of degree six while preserving the values in the points $(z_0, z_1, \ldots, z_n) = (0, 0.3, 0.5, -0.1, -0.7, -0.3 \pm 0.3i)$. Such an interpolation condition occurs in certain applications.

Step (i) to solve the quasi-convex optimization problem to minimize (17) yields optimal ϵ, p and q, and Step (ii) the approximant g, the Bode plot of which is depicted in Figure 2 together with that of f. The third subplot in the picture shows the relative error

$$\left| \frac{f(e^{i\theta}) - g(e^{i\theta})}{f(e^{i\theta})} \right| \quad \text{for } \theta \in [0, \pi].$$

It is important to note that the function g, which is guaranteed to be stable, satisfies the prespecified interpolation conditions and the error bound (19). Figure 2 shows that g matches f quite well.

Example 2: Next we approximate the function f in Example 1 without imposing any interpolation condition. For $n = 6$ and $n = 8$, we determine an approximant g_n of degree n via Steps (i) and (ii)$''$ that satisfies the relative error bound (20). Then we compare g_n to an approximant $\hat{f}_n$ of the same degree obtained by balanced truncation [9], [12]. The respective Bode plots and relative errors are depicted in Figures 3 and 4.

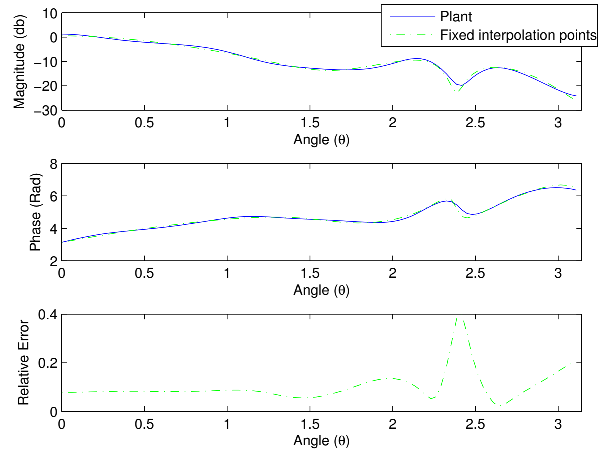

Fig. 2. Bode plots of f and g together with the relative error.

Our procedure performs better in the valleys of f (on the unit circle) than in the peaks. This is not surprising since balanced truncation comes with error bounds on the Hankel singular values and hence also on the H_∞ norm, while the bound of the proposed procedure is based on the relative error. For this reason we expect the approximants obtained by the proposed method to have a smaller relative H_2 error, but a larger absolute H_2 error, than approximants obtained by balanced truncation. The following tables show the relative and absolute H_2 errors.

Relative H_2 Error	Degree	
Approximation method	6	8
Proposed method	0.0764	0.0194
Balanced truncation	0.0785	0.0220
Error bound on g_n	0.8765	0.3994

H_2 Error	Degree	
Approximation method	6	8
Proposed method	0.0422	0.0100
Balanced truncation	0.0451	0.0057

For $n = 6$ the methods match the system with about the same error in both cases. On the other hand, for $n = 8$ the proposed method gives a lower relative H_2 error, and the balanced truncation gives a lower absolute H_2 error. From our experience, the case $n = 8$ is more representative. In fact, as expected, in many cases the H_2 error of the balanced truncation is lower than that of the proposed method, while the proposed method will deliver a smaller relative error.

Concerning the error bounds, it is obvious from the table

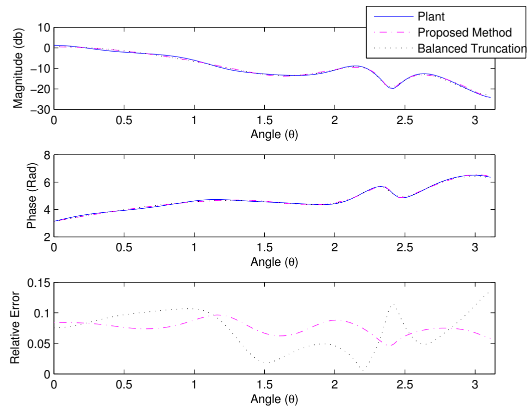

Fig. 3. *B*ode plot of f, g_6, and $\hat{f}_6$ together with the relative errors.

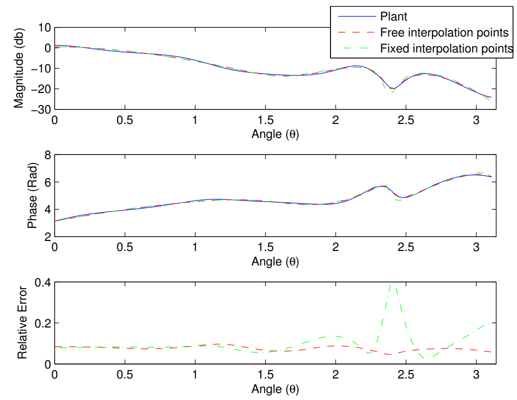

Fig. 5. *B*ode plot of f, g_6, and g together with the relative errors.

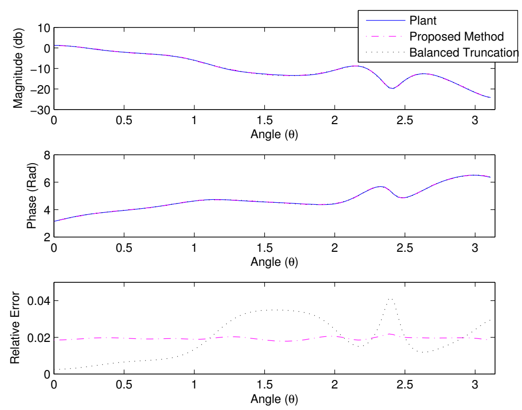

Fig. 4. *B*ode plot of f, g_8, and $\hat{f}_8$ together with the relative errors.

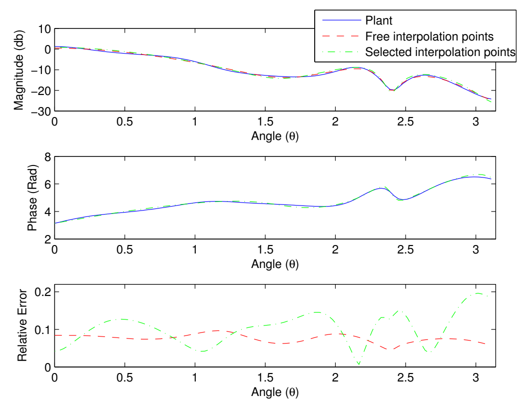

Fig. 6. *B*ode plot of f, g_6, and $\hat{g}_6$ together with the relative errors.

that they are conservative. How to improve them will be the subject to further studies.

In Figure 5 the approximant g from Example 1 is compared to g_6. The interpolation points for g_6 are chosen according to (ii)$''$, and the interpolation condition of g is prespecified. It can be seen from Figure 5 that g_6 matches f better than does g. This is because the interpolation points could be chosen freely for g_6.

Example 3: We continue to approximate the function in Example 1, but this time we move the interpolation points to get a better fit in a selected frequency band. In computing g_6 the interpolation points were determined via (ii)$''$ to be

$$(0, -0.5, -0.8841, -0.0380 \pm 0.7221i, -0.7021 \pm 0.6488i),$$

thus yielding the weight $|p(e^{i\theta})/\tau(e^{i\theta})| = 1$ for $\theta \in [0, \pi]$. In order to get a better fit close to 1 (i.e. at $\theta = 0$) we replace the interpolation point -0.5 with the point 0.9, thus producing the weight

$$\left|\frac{p(e^{i\theta})}{\tau(e^{i\theta})}\right| = \left|\frac{1+0.5e^{i\theta}}{1-0.9e^{i\theta}}\right| \text{ for } \theta \in [0, \pi].$$

Denote by $\hat{g}_6$ the minimizer (13) corresponding to the interpolation points $(0, 0.9, -0.8841, -0.0380 \pm 0.7221i, -0.7021 \pm 0.6488i)$. The functions g_6 and $\hat{g}_6$ are depicted in Figure 6. In the selected region close to 1, $\hat{g}_6$ approximates the original system better than does g_6, but this is at the expense of the approximation in other regions of the unit circle.

VIII. Conclusions and Further Work

In this paper, we propose a method for degree reduction of stable systems. The method is based on weighted H_2 minimization under interpolation constraints. By choosing weights appropriately, the minimizer will both be of low degree and match the original system. This gives a model reduction procedure for the case that both the original system and the degree-reduced system satisfy prespecified interpolation conditions (Section V). In the case where no such interpolation conditions are required, we provide a systematic procedure which utilizes the extra freedom of choosing the interpolation points (Section VI). The model reduction procedure without a priori interpolation conditions is compared with the balanced truncation method, and the quality of the approximants are found to be comparable.

The study of the H_2 minimization problem is motivated by the relation between the H_2 norm and the entropy functional used in bounded interpolation. Therefore, new concepts derived in this framework are useful for understanding entropy minimization. In fact, both the degree reduction methods proposed in this paper easily generalize to the bounded case; see [7] for the method which preserves interpolation conditions. We are currently working on similar bounds for the positive real case; also, see [6].

APPENDIX

A quasi-convex optimization problem is an optimization problem for which each sublevel set is convex. The optimization problem to minimize (18), where p and q are polynomials of fixed degree is quasi-convex. For simplicity, we assume that f is real and hence that p and q are real as well.

As a first step, consider the *feasibility problem* of finding a pair (p, q) of polynomials satisfying

$$\left\| 1 - \left| \frac{qf}{p} \right|^2 \right\|_\infty \leq \epsilon \tag{22}$$

for a given ϵ, or, equivalently,

$$-\epsilon |p(e^{i\theta})|^2 \leq |p(e^{i\theta})|^2 - |q(e^{i\theta}) f(e^{i\theta})|^2 \leq \epsilon |p(e^{i\theta})|^2$$

for all $\theta \in [-\pi, \pi]$. Since $|p|^2$ and $|q|^2$ are pseudo-polynomials, they have representations

$$\begin{aligned} |p(e^{i\theta})|^2 &= 1 + \sum_{k=1}^{n_p} p_k \cos(k\theta), \\ |q(e^{i\theta})|^2 &= \sum_{k=0}^{n_q} q_k \cos(k\theta), \end{aligned}$$

where n_p and n_q are the degree bounds on p and q respectively, and the first coefficient in $|p|^2$ is chosen to be one without loss of generality. Hence (22) is equivalent to

$$\begin{aligned} -(1+\epsilon) \leq (1+\epsilon) \sum_{k=1}^{n_p} p_k \cos(k\theta) - |f(e^{i\theta})|^2 \sum_{k=0}^{n_q} q_k \cos(k\theta), \\ (1-\epsilon) \leq (\epsilon - 1) \sum_{k=1}^{n_p} p_k \cos(k\theta) + |f(e^{i\theta})|^2 \sum_{k=0}^{n_q} q_k \cos(k\theta), \end{aligned}$$

for all $\theta \in [-\pi, \pi]$. There is also a requirement on $1 + \sum_{k=1}^{n_p} p_k \cos(k\theta)$ and $\sum_{k=0}^{n_q} q_k \cos(k\theta)$ to be positive. However, if $\epsilon \in (0, 1)$, then the above constraints will imply positivity. The set of $p_1, p_2, \ldots, p_{n_p}, q_0, q_1, \ldots, q_{n_q}$ satisfying this infinite number of linear constraints is convex.

The most straightforward way to solve this feasibility problem is to relax the infinite number of constraints to a finite grid, which is dense enough to yield an appropriate solution. Here one must be carefully to check the positivity of $1 + \sum_{k=1}^{n_p} p_k \cos(k\theta)$ and $\sum_{k=0}^{n_q} q_k \cos(k\theta)$ in the regions between the grid points. Another method is the Ellipsoid Algorithm, described in detail in [2].

Minimizing (18) then amounts to finding the smallest ϵ for which the feasibility problem has a solution. This can be done by the the bisection algorithm, as described in [2]. Note that for $\epsilon = 1$, the trivial solution $q = 0$ is always feasible.

REFERENCES

[1] A.C. Antoulas, A new result on passivity preserving model reduction, *Systems and Control Letters*, vol. 54, 2005, pp. 361-374.

[2] S. Boyd, L. E. Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, 1994.

[3] C.I. Byrnes, T.T. Georgiou and A. Lindquist, A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint, *IEEE Trans. Automatic Control*, vol. 46, June 2001, pp. 822-839.

[4] C.I. Byrnes, T.T. Georgiou, A. Lindquist and A. Megretski, Generalized interpolation in H^∞ with a complexity constraint, *Trans. American Mathematical Society*, vol. 358, 2006, pp. 965-987.

[5] C. I. Byrnes and A. Lindquist, *On the duality between filtering and Nevanlinna-Pick interpolation*, SIAM J. Control and Optimization **39** (2000), 757–775.

[6] G. Fanizza, J. Karlsson, A. Lindquist and R. Nagamune, Passivity-preserving model reduction by analytic interpolation, accepted for publication in *Linear Algebra and its Applications*.

[7] J. Karlsson, T.T. Georgiou and A. Lindquist, The Inverse Problem of Analytic Interpolation with Degree Constraint, Proc. CDC 2006.

[8] D. G. Luenberger, **Optimization by Vector Space Methods**, John Wiley & Sons, 1969.

[9] S. Skogestad and I. Postlethwaite, **Multivariate Feedback Control**, John Wiley & Sons, 1996.

[10] D.C. Sorensen, Passivity preserving model reduction via interpolation of spectral zeros, *Systems and Control Letters*, vol. 54, 2005, pp. 347-360.

[11] K. C. Sou, A. Megretski, and L. Daniel, “A quasi-convex optimization approach to parameterized model order reduction,” *IEEE Proc. on Design Automation Conference*, 933-938, Jun. 2005.

[12] K. Zhou, **Essentials of Robust Control**, Prentice-Hall, Inc, 1998.