An Interpolation-Based Approach to the Optimal $\mathcal{H}_\infty$ Model Reduction

Garret M. Flagg

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master's of Science
in
Mathematics

Serkan Gugercin, Chair
Chris Beattie
Jeff Borggaard

May 5, 2009
Blacksburg, Virginia

Keywords: Rational Interpolation, Model Reduction, Optimization

An Interpolation-Based Approach to the Optimal H_∞ Model Reduction

Garret M. Flagg

(ABSTRACT)

A model reduction technique that is optimal in the $\mathcal{H}_\infty$-norm has long been pursued due to its theoretical and practical importance. We consider the optimal $\mathcal{H}_\infty$ model reduction problem broadly from an interpolation-based approach, and give a method for finding the approximation to a state-space symmetric dynamical system which is optimal over a family of interpolants to the full order system. This family of interpolants has a simple parameterization that simplifies a direct search for the optimal interpolant. Several numerical examples show that the interpolation points satisfying the Meier-Luenberger conditions for $\mathcal{H}_2$ -optimal approximations are a good starting point for minimizing the $\mathcal{H}_\infty$-norm of the approximation error. Interpolation points satisfying the Meier-Luenberger conditions can be computed iteratively using the IRKA algorithm [12]. We consider the special case of state-space symmetric systems and show that simple sufficient conditions can be derived for minimizing the approximation error when starting from the interpolation points found by the IRKA algorithm. We then explore the relationship between potential theory in the complex plane and the optimal $\mathcal{H}_\infty$-norm interpolation points through several numerical experiments. The results of these experiments suggest that the optimal $\mathcal{H}_\infty$ approximation of order r yields an error system for which significant pole-zero cancellation occurs, effectively reducing an order $n+r$ error system to an order $2r+1$ system. These observations lead to a heuristic method for choosing interpolation points that involves solving a rational Zolatarev problem over a discrete set of points in the complex plane.

Contents

1 Background **1**

1.1 Introduction 1

1.2 System Norms: Measuring Error 3

1.3 Model Reduction in the Projection Framework 6

1.4 Interpolation-Based Model Reduction 7

1.5 SVD-Based Approach To Model Reduction 10

2 Interpolation and the system D-term **15**

2.1 Emancipating D 15

2.1.1 Numerical Results 20

2.2 The optimal d_r-term for state-space symmetric systems 22

3 How Then Shall We Balance Our Charges? **39**

3.1 Potential Theory and Rational Interpolation in the Complex Plane 39

3.2 Observations on Where the Ancillary Zeros of the Error System Do Not Occur 48

3.3 Interpolation Points from Penzl's Heuristic Method 51

4 Conclusions and Future Work **55**

List of Figures

2.1 Comparison of the Error Systems 37

2.2 Comparison of the Error Systems For PEEC Model 38

3.1 Equipotentials of the Error System 41

3.2 Poles of 2-D beam model and ancillary zeros of order 20 optimal Hankel norm approximation 43

3.3 Poles of CD Player model and ancillary zeros of order 8 optimal $H_r^{d_r}(s)$ approximation 44

3.4 Poles of International Space-Station Model Order 270 and Ancillary Zeros of Order 4 Optimal Hankel Norm Approximation 45

3.5 Poles of Eady Model Order 548 and Ancillary Zeros of Order 10 Optimal Hankel Norm Approximation 46

3.6 Average Distance of CD-Player Poles to Nearest Ancillary Zeros of Order 6 Approximations 49

3.7 Magnitude Bode Plot of FOM Model and Equivalent Charges Approximation of Order 12 53

List of Tables

2.1 Comparison of Order 10 Approximations to Order 120 CD Player Model . . 21

2.2 Comparison of Order 4 Approximations to Order 1434 Spiral-Inductor Model 21

3.1 Contribution of Ancillary Zeros to Error for International Space Station Model 48

3.2 Comparison of error norms for approximations of table 3.1 48

3.3 Correspondence of Unmatched Poles to IRKA Interpolation Points For Order 4 Approximations to Various Models 50

3.4 Comparison of Method Performance For Heat Model of Order 197 52

3.5 Comparison of Method Performance For Eady Model Order 598 54

3.6 Comparison of Method Performance For FOM Model Order 1006 54

Chapter 1

Background

1.1 Introduction

Consider a single-input-single-output (SISO) linear time invariant dynamical system Σ described by a system of first order differential equations with state space representation:

$$\Sigma := \begin{cases} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}u(t) \\ y(t) &= \boldsymbol{c}\boldsymbol{x}(t) + du(t). \end{cases} \tag{1.1}$$

Where $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\boldsymbol{b}$ and $\boldsymbol{c}^T \in \mathbb{C}^n$, $d \in \mathbb{C}$, $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$. The order of Σ is n, the dimension of $\boldsymbol{A}$. In an ever increasing number of applications, n may become so large that Σ as the solution to a simulation or control problem becomes computationally intractable. The goal of model reduction is to search for a reduced-order system Σ_r of order r with state-space-representation:

$$\Sigma_r := \begin{cases} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}_r \boldsymbol{x}_r(t) + \boldsymbol{b}_r u(t) \\ y_r(t) &= \boldsymbol{c}_r \boldsymbol{x}_r(t) + d_r u(t) \end{cases} \tag{1.2}$$

having the same inputs as Σ, and satisfying $\|y(t) - y_r(t)\|$ small for some norm. As we will see in the next section, one method of achieving this goal is to minimize the difference between the full and reduced-order transfer functions in the $\mathcal{H}_\infty$- norm. The problem of achieving this minimization is called the optimal $\mathcal{H}_\infty$ model reduction problem.

The basic objects involved in the $\mathcal{H}_\infty$ model reduction problem for (SISO) systems are complex-valued rational functions (the system's transfer function) and their corresponding zeros and poles. The system to be approximated can be thought of as a rational function of very large degree, and the problem of finding a good reduced-order model of order r essentially becomes the problem of finding a rational function of degree r so that their difference is uniformly small in $\mathbb{C}_+$. The image one might start with involves hundreds of positive and negative charges (poles and zeros of the error system respectively) distributed throughout the complex plane. Loosely speaking, the goal is to choose a reduced-order model so that the imaginary axis is an equipotential of this distribution of charges. In this sense, the problem is an electrostatics problem, and so many of the tools from potential theory in the complex plane may be applied to the $\mathcal{H}_\infty$ model reduction problem. The fact that a connection between model reduction and potential theory even exists is interesting in its own right, and a continual effort to appreciate the import of this conceptual image eventually led to the results we present here. The goal of this thesis is to present a simple and effective method for reducing the $\mathcal{H}_\infty$ error, and to give a conceptual account for its performance in terms of this electrostatic analogy. This work is laid out as follows: In Chapter 1, we give an overview of system norms and briefly survey the interpolation and SVD-based methods of model reduction. In Chapter 2, we present a method for reducing the $\mathcal{H}_\infty$-norm of the

error when starting from the $\mathcal{H}_2$ -optimal reduced order model. In Chapter 3 we consider the results of this method in potential theoretic terms and present a number of numerical experiments that suggest the optimal $\mathcal{H}_\infty$ problem can effectively be reduced to a rational Zolatarev problem involving specific poles of the full order system.

1.2 System Norms: Measuring Error

For any reduced-order model to be worth its salt, it should be close to the full-order model in some meaningful way. Naturally then, we should expect a system norm to measure properties of the system which are invariant under representation. To obtain such a norm, we first consider the external representation of the system 1.1 as the convolution operator $\mathcal{S}$ that maps inputs $u(t)$ to outputs $y(t)$ by convolution with the impulse-response of the system:

$$(\mathcal{S}u)(t) = y(t) = \int_0^\infty \boldsymbol{h}(t-\tau)u(\tau)\,\mathrm{d}\tau \tag{1.3}$$

[1]. The kernel $\boldsymbol{h}(t)$ is the systems impulse-response and is defined as:

$$\boldsymbol{h}(t) = \begin{cases} \boldsymbol{c}e^{\boldsymbol{A}t}\boldsymbol{b} + \delta(t)d & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{1.4}$$

Taking the Laplacian of $\mathcal{S}$ and making use of the well-known result that convolution in the time-domain is the same as multiplication in the frequency domain, we have:

$$Y(s) = H(s)U(s) \tag{1.5}$$

and solving for $H(s)$ gives

$$H(s) = \boldsymbol{c}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} + d. \tag{1.6}$$

$H(s)$ is the system's transfer function and is independent of the internal state-space representation of the system. Unless otherwise stated, we will always associate the system Σ with its transfer function $H(s)$ and the reduced order system Σ_r by with its transfer function $H_r(s)$. For SISO systems the transfer function $H(s)$ is simply a rational function, and if the system Σ is asymptotically stable all the poles of $H(s)$ lie in $\mathbb{C}_-$. Thus, $H(s)$ is holomorphic in $\mathbb{C}_+$, and we may therefore apply standard Hardy space theory to $H(s)$ [1]. The system norms in the frequency domain then become the Hardy space p-norms of the transfer function $H(s)$, which are defined in the following way:

$$\|H(s)\|_{\mathcal{H}_p} = \left(\int_{-\infty}^{\infty} |H(i\omega)|^p \, \mathrm{d}\omega\right)^{1/p} \tag{1.7}$$

The values of p we are interested in are $p = 2$ and $p = \infty$. When $p = \infty$ we define the $\mathcal{H}_\infty$-norm by

$$\|H(s)\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} |H(i\omega)| \tag{1.8}$$

Thus the $\mathcal{H}_\infty$-norm of a SISO system is the maximum value of the modulus of the transfer function, where the maximum is taken over the imaginary axis. The $\mathcal{H}_\infty$-norm has a useful interpretation in terms of the systems outputs $y(t)$ in the time domain due to the following result [1].

Theorem 1.2.1. *Let $H(s) = \mathcal{L}(h)(s)$, and $U(s) = \mathcal{L}(u)(s)$ and suppose $H(s) \in \mathcal{H}_\infty(\mathbb{C}_+)$. Then*

$$\|H(s)\|_{\mathcal{H}_\infty} = \sup_{U \neq 0} \frac{||HU||_{\mathcal{H}_2}}{||U||_{\mathcal{H}_2}} = \sup_{u \neq 0} \frac{||h * u||_{\mathcal{L}_2}}{||u||_{\mathcal{L}_2}}. \tag{1.9}$$

This means that the $\mathcal{H}_\infty$-norm of $H(s)$ gives us an upper bound on the magnitude of system outputs in terms of the magnitude of the system inputs. If we consider the error system for an order r approximation, its transfer function is $E(s) = H(s) - H_r(s)$, and we therefore have an upper bound on the magnitude of the error for all inputs of a certain magnitude. Such an upper bound is useful in applications where the error cannot exceed a certain tolerance at any point in the evolution of the system. We can now formally define the optimal $\mathcal{H}_\infty$ model reduction problem.

Optimal $\mathcal{H}_\infty$Model Reduction Problem 1.2.1. Given a stable system Σ of order n, and $r \ll n$, find an asymptotically stable system Σ_r^* of order r, such that

$$H_r^*(s) = \arg \min_{H_r \in \mathcal{H}_\infty} \|H(s) - H_r(s)\|_{\mathcal{H}_\infty} . \tag{1.10}$$

Returning to the functional representation of Σ given by the convolution operator $\mathcal{S}$ we find that the spectrum of $\mathcal{S}$ $= \Lambda(\mathcal{S}) = \{H(i\omega) : \omega \in \mathbb{R}\}$ and the set of singular values of $\mathcal{S}$ is given by

$$\Sigma(\mathcal{S}) = \{|H(i\omega)| : \omega \in \mathbb{R}\}$$

[7]. The largest entry of $\Sigma(\mathcal{S})$ is the $\mathcal{H}_\infty$-norm of $H(s)$. Since the spectrum of $\mathcal{S}$ is a continuum, it does not lend itself to results analogous to the Schmidt-Mirsky-Eckardt-Young result for the optimal approximation of a matrix operator of rank n by a matrix of rank r in the $\mathcal{L}_2$-induced norm defined for that operator [1]. The fact that the set of singular values of $\mathcal{S}$ is a continuum is in some sense the cardinal difficulty in finding a reduced order model which is optimal in the $\mathcal{H}_\infty$-norm, since such a model would in turn minimize the $\mathcal{L}_2$ induced-norm of $\mathcal{S} - \mathcal{S}_r$. In light of this difficulty, the $\mathcal{H}_\infty$ model reduction problem in an interpolation framework avoids the problem of directly approximating an infinite dimensional operator all together.

1.3 Model Reduction in the Projection Framework

The model reduction problem in its most general form can be divided into two relatively independent subproblems. The first problem is to develop a method for constructing reduced-order models that preserve certain representational and intrinsic properties of the full-order system. For example, if the $H(s)$ has a realization $H(s)$=($\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$, d) where $\boldsymbol{A}$ is Hermitian, we might search for method of model reduction which preserves this property. If in addition the original system is passive, then ideally the method would preserve this intrinsic property of the system. The second problem involves characterizing the reduced-order models constructed by a particular method which are optimal, in the sense that they minimize the error in some norm.

Given a system Σ, reduction in the projection framework is a method of constructing a system Σ_r as in (1.2) from a Galerkin approximation to Σ [12]. In general, this can be accomplished by choosing $\boldsymbol{W}, \boldsymbol{V} \in \mathbb{R}^{n\times r}$, satisfying $\boldsymbol{W}^*\boldsymbol{V} = \boldsymbol{I}_r$, and defining the reduced-order model (1.2) in terms of the projection $\boldsymbol{\Pi} = \boldsymbol{V}\boldsymbol{W}^*$:

$$H_r(s) := \left\{ \begin{array}{rcl} \dot{\boldsymbol{x}}_r(t) & = & \boldsymbol{W}^*\boldsymbol{A}\boldsymbol{V}\boldsymbol{x}_r(t) + \boldsymbol{W}^*\boldsymbol{b}u(t) \\ y_r(t) & = & \boldsymbol{c}\boldsymbol{V}\boldsymbol{x}(t) + du(t). \end{array} \right. \tag{1.11}$$

The trajectory $\boldsymbol{x}_r(t) = \boldsymbol{W}^*\boldsymbol{x}(t)$ evolves in an r-dimensional subspace of the original state-space. The system Σ_r given by (1.11) is a Petrov-Galerkin approximation in the the sense that the error between the original state $\boldsymbol{x}(t)$ and the projected state $\boldsymbol{\Pi}\boldsymbol{x}(t)$ is orthogonal to $\boldsymbol{W}$, and lies in the range of $\boldsymbol{V}$. For more details, see [1], [12].

In this framework, the quality of the reduced order model depends directly on the choice of the matrices $\boldsymbol{V}$and $\boldsymbol{W}$. Broadly speaking, the methods for choosing $\boldsymbol{V}$and $\boldsymbol{W}$can be divided between two conceptual approaches: 1.) $\boldsymbol{\Pi}$ is chosen to guarantee the grammians of

Σ_r satisfy certain properties, or 2.) $\mathbf{\Pi}$ is chosen to guarantee Σ_r satisfies some interpolation conditions. The former we refer to as SVD-based methods, and the latter as interpolation-based methods. We will return to SVD-based methods for optimal $\mathcal{H}_\infty$ model reduction in the next section. First we will consider model reduction by interpolation and its explicit solution in the projection framework.

1.4 Interpolation-Based Model Reduction

Given a full-order system Σ, the jth moment of $H(s)$ at σ_i is defined as $\boldsymbol{c}(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-(j+1)}\boldsymbol{b}$, and is equivalent to the jth derivative of $H(s)$ at the point σ_i. Given the points $\sigma_1, \ldots, \sigma_k \in \mathbb{C}$ the problem of model reduction by interpolation, or moment-matching, involves finding a system $H_r(s)$ that matches $H(s)$ at a prescribed number of moments for each point σ_i. This problem can be generalized to matching moments of $H(s)$ in the case where the interpolation point σ is the point at infinity, in terms of the Taylor's series expansion of $H(s)$ about the point at infinity, provided $H(s)$ is analytic there. Trivially, moment matching can be accomplished by computing the necessary moments and then solving for the coefficients of the numerator and denominator polynomials of the reduced order transfer function. The problem with this approach is that it is extremely numerically ill-conditioned, so if interpolation is to be accomplished it must be done by some other means [9].

Rational interpolation by projection was first proposed by Skelton *et al.* [24],[26],[27]. Later, Grimme [11] showed how to obtain the required projection, using the method of Ruhe. Interpolation done by this method is much better conditioned, and avoids the problem of computing higher order moments entirely. The rational interpolation results given here can be extended to the general case of multi-input-multi-output (MIMO) systems, but for our purposes, we will present only the special case of interpolation for SISO systems. Proposition

1.4.1 guarantees Hermite interpolation by the reduced order system, which is an important necessary condition for first-order $\mathcal{H}_2$ -optimal model reduction.

Proposition 1.4.1 ([12]). *Consider the system Σ defined by $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$, a set of distinct shifts given by $\{\sigma_i\}_{k=1}^{r}$, that is closed under conjugation, and subspaces spanned by the columns of $\boldsymbol{V}_r$, and $\boldsymbol{W}_r$ with*

$$Ran(\boldsymbol{V}_r) = span\left\{(\sigma_1\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b},\ldots,(\sigma_r\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}\right\}, \tag{1.12}$$

$$Ran(\boldsymbol{W}_r) = span\left\{(\sigma_1\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{c},\ldots,(\sigma_r\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{c}\right\}. \tag{1.13}$$

Then $\boldsymbol{V}_r$and $\boldsymbol{W}_r$ can be chosen to be real matrices and the reduced order system Σ_r defined by $\boldsymbol{A}_r$=$(\boldsymbol{W}^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{A}\boldsymbol{V}_r$, $\boldsymbol{b}_r$=$(\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{b}$, $\boldsymbol{c}_r$=$\boldsymbol{c}\boldsymbol{V}_r$ is real and matches the first two moments of $H(s)$ at each of the interpolation points σ_k.

The matrices $\boldsymbol{V}_r$and $(\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T$ can easily be seen to correspond to $\boldsymbol{V}$ and $\boldsymbol{W}^T$ as in the projection framework, thus it becomes clear how interpolation in the projection framework is accomplished. The general Rational Krylov Method, due to Grimme, gives us the ability to construct $H_r(s)$ with control over $2r$ distinct interpolation points [11] .

Proposition 1.4.2. *Given $2r$ distinct points $\sigma_1,\ldots,\sigma_{2r}\in\mathbb{C}$, let*

$$\boldsymbol{V}_r = [(\sigma_1\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}\ldots(\sigma_r\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}] \qquad \boldsymbol{W}_r^T = \begin{bmatrix} \boldsymbol{c}^T(\sigma_{r+1}\boldsymbol{I}-\boldsymbol{A})^{-1} \\ \vdots \\ \boldsymbol{c}^T(\sigma_{2r}\boldsymbol{I}-\boldsymbol{A})^{-1} \end{bmatrix}.$$

If $det(\boldsymbol{W}_r^T\boldsymbol{V}_r)\neq 0$,

define $H_r(s)$ by:

$$H_r(s) := \begin{array}{c|c} \boldsymbol{A}_r = (\boldsymbol{W}_r^T \boldsymbol{V}_r)^{-1} \boldsymbol{W}_r^T \boldsymbol{A} \boldsymbol{V}_r & \boldsymbol{b}_r = (\boldsymbol{W}_r^T \boldsymbol{V}_r)^{-1} \boldsymbol{W}_r^T \boldsymbol{b} \\ \hline \boldsymbol{c}_r = \boldsymbol{c} \boldsymbol{V}_r & d_r = d. \end{array}$$

Then $H(\sigma_i) = H_r(\sigma_i)$, $i = 1, \dots 2r$.

We conclude this section with an important result in interpolation-based model reduction that provides the necessary first-order conditions for optimality in the $\mathcal{H}_2$ -norm. These conditions are generally referred to as the Meier-Luenberger conditions [12].

Theorem 1.4.3 ($\mathcal{H}_2$ **-Optimal Interpolation Conditions** [12])**.** *Let* $H(s)$ *be stable and of order* n*. If* $H_r(s)$ *is a stable* $\mathcal{H}_2$ *-optimal reduced order model of order* r*, and the poles of* $H_r(s)$ *are* $\hat{\lambda}_1, \dots, \hat{\lambda}_r$*, then* $H(-\hat{\lambda}_i) = H_r(-\hat{\lambda}_i)$ *for* $i = 1, \dots, r$*, and* $H'(-\hat{\lambda}_i) = H_r'(-\hat{\lambda}_i)$ *for* $i = 1, \dots, r$.

To satisfy these necessary conditions requires us to force interpolation of the full-order transfer function and its derivative at the mirror image of the poles of the reduced-order system. Since we cannot know the reduced order systems poles *a priori*, we compute them iteratively using the Iterative Rational Krylov Algorithm (IRKA) developed by Gugercin *et al.* [12]. This algorithm is a fixed point iteration, outlined as follows:

Algorithm 1.4.1 (**An Iterative Rational Krylov Algorithm**)**.**

1. Make an initial selection of interpolation points σ_i, for $i = 1, \dots, r$ that is closed under conjugation and fix a convergence tolerance *tol*.
2. Choose $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$ satisfying the hypotheses of Proposition 1.4.1.
3. while (relative change in $\{\sigma_i\} < tol$)

   a.) $\boldsymbol{A}_r = \boldsymbol{W}_r^T \boldsymbol{A} \boldsymbol{V}_r$

b.) Assign $\sigma_i \leftarrow -\lambda(\boldsymbol{A}_r)$ for $i = 1, \ldots, r$.

c.) Update $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$ to satisfy hypotheses of Proposition 1.4.1 with new σ_i's.

4. $\boldsymbol{A}_r{=}\boldsymbol{W}_r^T\boldsymbol{A}\boldsymbol{V}_r$, $\boldsymbol{b}_r{=}\boldsymbol{W}_r^T\boldsymbol{b}$ $\boldsymbol{c}_r{=}\boldsymbol{c}_r\boldsymbol{V}_r$

It has been observed that IRKA has good convergence behavior, and tends to converge to at least a local minimum of the $\mathcal{H}_2$ optimal model reduction problem. [12]

1.5 SVD-Based Approach To Model Reduction

Recall that the spectrum of the convolution operator $\mathcal{S}$ associated with the system Σ is a continuum. This presents the cardinal difficulty for approaching the $\mathcal{H}_\infty$ model reduction problem from the position of approximating the spectrum of $\mathcal{S}$: there is simply no good way to order the importance of points in the spectrum [1]. In light of this difficulty, researchers turned their attention to the Hankel operator associated with Σ. The Hankel operator of Σ can be viewed as a map of previous inputs $u_-(t)$ to future outputs $y_+(t)$ and is defined in the following way:

$$\mathcal{H}(u)(t) = \int_{-\infty}^{0} h(t-\tau)u(\tau)\,\mathrm{d}\tau. \tag{1.14}$$

The operator $\mathcal{H}$ is compact and bounded, and therefore has a discrete SVD [1]. The singular values of $\mathcal{H}$ are called the Hankel singular values of Σ. If Σ is asymptotically stable and if $\mathcal{H}$ has domain $\mathcal{L}_2(\mathbb{R}_-)$ and codomain $\mathcal{L}_2(\mathbb{R}_+)$, then the $\mathcal{L}_2$-induced norm of $\mathcal{H}$ is

$$\|\mathcal{H}\|_{\mathcal{L}_{2-\text{induced}}} = \sigma_{max}(\mathcal{H}) = \|\Sigma\|_{\text{H}} \leq \|H(s)\|_{\mathcal{H}_\infty}. \tag{1.15}$$

The quantity $\|\Sigma\|_{\mathrm{H}}$ is the Hankel-norm of Σ. The Hankel singular values give a break-down of the system states that require the most energy to control and the states that produce the smallest amount of observable energy. This can be seen directly by considering the infinite observability and controllability gramians $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{Q}}$ respectively. If Σ=($\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$, d) is stable, then the infinite observability and controllability gramians are given as the solutions of the following Lyapunov equations:

$$\boldsymbol{A}\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\boldsymbol{A}^* + \boldsymbol{b}^*\boldsymbol{b} = 0, \tag{1.16}$$

$$\boldsymbol{A}^*\boldsymbol{\mathcal{Q}} + \boldsymbol{\mathcal{Q}}\boldsymbol{A} + \boldsymbol{c}\boldsymbol{c}^* = 0. \tag{1.17}$$

Furthermore, if Σ is completely reachable and completely observable, then $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{Q}}$ are positive definite. The minimal $\mathcal{L}_2$-energy required for Σ to reach a given state $\boldsymbol{x}$ is then given by

$$\mathcal{E}_r = \boldsymbol{x}^*\boldsymbol{\mathcal{P}}^{-1}\boldsymbol{x},$$

and the largest observable $\mathcal{L}_2$-energy produced by the state $\boldsymbol{x}$ is given by

$$\mathcal{E}_o = \boldsymbol{x}^*\boldsymbol{\mathcal{Q}}\boldsymbol{x}.$$

Therefore the states which are hard to reach or hard to control have significant components in the direction of the eigenvectors of $\boldsymbol{\mathcal{P}}$ with the smallest eigenvalues, and the states which are easiest to observe have components in the direction of the eigenvectors corresponding to the largest eigenvalues of $\boldsymbol{\mathcal{Q}}$. Unfortunately, these states do not necessarily coincide for an arbitrary realization of Σ, so if we throw away states which are very hard to control we might also be throwing away the states which are easiest to observe, and vice-versa.

However if $\boldsymbol{\mathcal{Q}}$=$\boldsymbol{\mathcal{P}}$=$\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is a diagonal matrix, then the states which produce the largest observability energy are also the states which require the smallest energy to control. It turns out that the Hankel singular values of Σ are given by the following equation:

$$\sigma(\mathcal{H}) = \sqrt{\Lambda(\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{Q}})}. \tag{1.18}$$

So if

$$\boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{Q}} = \boldsymbol{\Sigma}, \tag{1.19}$$

then the diagonal entries of $\boldsymbol{\Sigma}$ are precisely the Hankel singular values of $\mathcal{H}$. If the solutions to the Lyapunov equations for a realization of the system Σ satisfy (1.19), then the realization is called a principle-axis balanced realization. If Σ is a stable system, then it is always possible to find a state-space transformation for which the corresponding realization of Σ is balanced. The standard SVD-based approach to model reduction is to transform Σ to its balanced realization and then simply throw away the states corresponding the the smallest Hankel singular values of Σ. This method is called balanced truncation. If $H_{bal}(s)$ is a reduced order model constructed by balanced truncation, then we have the following upper bound on the error given in terms of the discarded Hankel singular values of Σ:

$$\|H(s) - H_{\text{bal}}(s)\|_{\mathcal{H}_\infty} \leq 2 \sum_{i=r+1}^{n} \sigma_i. \tag{1.20}$$

Balanced truncation is roughly analogous to the optimal method of approximating a rank n matrix $\boldsymbol{A}$ by a rank r matrix $\boldsymbol{A}_r$, which is done by simply truncating $\boldsymbol{A}$ at its rth singular value. This analogy breaks down at the error bounds that balanced truncation yields for the $\mathcal{H}_\infty$-norm of the error. However, it is possible to completely extend the approximation analogy to dynamical systems if we instead replace the $\mathcal{H}_\infty$-norm by the Hankel-norm of Σ,

and consider the operator $\mathcal{H}$, rather than $\mathcal{S}$. This is an important and nontrivial result, which we summarize in the theorem below. The interested reader can find the full development of this result and the related results for balanced truncation presented above in the following references [14], [28], [1].

Theorem 1.5.1. *If Σ is a stable system of order n, and Σ_r is stable and of order $r < n$, then*

$$\|\Sigma - \Sigma_r\|_H \geq \sigma_{r+1}.$$

Furthermore, there exists a stable system Σ_{Hank} of order r such that

$$\|\Sigma - \Sigma_{Hank}\|_H = \sigma_{r+1}. \tag{1.21}$$

and

$$\|H(s) - H_{Hank}(s)\|_{\mathcal{H}_\infty} \leq \sum_{i=r+1}^{n} \sigma_i. \tag{1.22}$$

This result also yields the following lower bound on the error due to an order r approximation in the $\mathcal{H}_\infty$-norm.

Corollary 1.5.2. *If Σ is a stable system of order n and Σ_r is a stable system of order r, then*

$$\|H(s) - H_r(s)\|_{\mathcal{H}_\infty} \geq \sigma_{r+1}. \tag{1.23}$$

The approximation $H_{\text{Hank}}(s)$ is called the optimal Hankel norm approximation to Σ, and it serves as the starting point for more complicated approximation methods involving the solution of linear matrix inequalities that can further reduce the $\mathcal{H}_\infty$-norm of the error system. For example, Kavronoglu [13] actually gives a characterization of the solution to the optimal $\mathcal{H}_\infty$-norm model reduction problem using the techniques of matrix dilation developed by Glover *et al.* Unfortunately, as Kavronoglu points out, his method of solution

bottlenecks at the step of determining the exact lower-bound $\gamma \geq \sigma_{r+1}$ on the error for all approximations of order r. Given that our goal is to reduce the order of very large systems, the main disadvantage to the SVD and LMI-based approaches is their dependence on the solution of Lyapunov equations that are on the order of the full system. The work required to solve the Lyapunov equations is on the order of n^3, so for large systems this becomes impractical. The most obvious strength of the SVD-based methods is that they yield tight upper bounds on the approximation error in the $\mathcal{H}_\infty$-norm.

Chapter 2

Interpolation and the system D-term

2.1 Emancipating D

As the chapter heading suggests, the hero of this chapter is the D-term of the reduced-order system. Surprisingly enough, this one parameter does a number of things for us all at once: it will enable us to increase the number of interpolation points we control by one, to bound the behavior of the system at infinity away from zero, and even to avail ourselves of simple expressions for the state-space representation of the transfer function's inverse and zeros. But in order to make use of D, we will first have to emancipate it from its current enthrallment to full-order system. Anyone who doubts its servitude in the interpolation framework need only cast a cursory glance over $H_r(s)$ defined as in Theorem 1.4.2 and simply note that we do not arbitrarily fix $d_r = d$, but that this is necessary for the result to hold. How then do we construct reduced-order models from interpolation with an independent d_r-term? The answer is to look at interpolation of a perturbed version $\tilde{H}(s)$ of $H(s)$ [5]. If the D-term of $\tilde{H}(s)$ is the D-term desired for the reduced order system, and if $\tilde{H}(s)$ takes the same value at the points $\sigma_1, \ldots, \sigma_{2r}$, as $H(s)$, then it is sufficient to apply the Rational Krylov method to

$\tilde{H}(s)$ to guarantee $H_r(s)$ interpolates $H(s)$ [5]. This is precisely the method of interpolation achieved by Beattie and Gugercin in [5], which we will demonstrate below.

First, we begin with $\sigma_1, \ldots, \sigma_{2r} \in \mathbb{C}$ and define $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$ as in the Rational Krylov method with respect to $H(s)$:

$$\boldsymbol{V}_r = [(\sigma_1 \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} \ldots (\sigma_r \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b}] \qquad \boldsymbol{W}_r^T = \begin{bmatrix} \boldsymbol{c}^T(\sigma_{r+1}\boldsymbol{I} - \boldsymbol{A})^{-1} \\ \vdots \\ \boldsymbol{c}^T(\sigma_{2r}\boldsymbol{I} - \boldsymbol{A})^{-1} \end{bmatrix}.$$

Then for any $d_r \in \mathbb{C}$, let

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} - d_r \boldsymbol{g}\boldsymbol{f}^T, \quad \tilde{\boldsymbol{b}} = \boldsymbol{b} - d_r \boldsymbol{g}\boldsymbol{f}^T, \text{ and } \quad \tilde{\boldsymbol{c}} = \boldsymbol{c} - d_r \boldsymbol{f}^T, \tag{2.1}$$

where $\boldsymbol{f}$ and $\boldsymbol{g}$ are solutions to

$$\boldsymbol{f}^T \boldsymbol{V}_r = \boldsymbol{e}^T \text{ and } \quad \boldsymbol{W}_r^T \boldsymbol{g} = \boldsymbol{e}. \tag{2.2}$$

and $\boldsymbol{e}$ is a column vector of r ones. Then we define the perturbed system $\tilde{H}(s)$:

$$\tilde{H}(s) = \tilde{\boldsymbol{c}}^T(s\boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}} + d_r \tag{2.3}$$

Note that because we put no constraints on d_r, $\tilde{H}(s)$ may only be stable, or even unstable, even if $H(s)$ is not, due to the fact that $\tilde{\boldsymbol{A}}$ is a rank one perturbation of $\boldsymbol{A}$. Nevertheless, the following result holds provided $\tilde{H}(s)$ is analytic at the points $\sigma_1, \ldots, \sigma_{2r}$.

Theorem 2.1.1. *For $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{b}}$, $\tilde{\boldsymbol{c}}$ as defined in (2.1),*

$$(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} = (\sigma_i \boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}} \text{ and } \boldsymbol{c}^T(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1} = \tilde{\boldsymbol{c}}^T(\sigma_i \boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1} \tag{2.4}$$

It follows that $H(\sigma_i) = \tilde{H}(\sigma_i),\ i = 1 \dots 2r.$

Proof. Observe that

$$\begin{aligned}
&(\sigma_i \boldsymbol{I} - \tilde{\boldsymbol{A}})(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} \\
&= (\sigma_i \boldsymbol{I} - \boldsymbol{A} - d_r \boldsymbol{g}\boldsymbol{f}^T)(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} \\
&= \boldsymbol{b} - d_r \boldsymbol{g}\boldsymbol{f}^T \boldsymbol{V}_r \boldsymbol{e}_i \\
&= \boldsymbol{b} - d_r \boldsymbol{g} = \tilde{\boldsymbol{b}}
\end{aligned}$$

and result follows similarly for the second equality in (2.4).

Therefore we have

$$\begin{aligned}
H(\sigma_i) - \tilde{H}(\sigma_i) &= \boldsymbol{c}^T(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} - \tilde{\boldsymbol{c}}^T(\sigma_i \boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}} - d_r \\
&= (\boldsymbol{c}^T - \boldsymbol{c}^T + d_r \boldsymbol{f}^T)(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} - d_r \\
&= (d_r \boldsymbol{f}^T)(\sigma_i \boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} - d_r \\
&= (d_r \boldsymbol{f}^T)\boldsymbol{V}_r \boldsymbol{e}_i - d_r \\
&= d_r - d_r = 0.
\end{aligned}$$

□

Corollary 2.1.2. *Let* $H_r(s)$ *correspond to model reduction of* $\tilde{H}(s)$ *as in* (2.3) *via the oblique projection* $\Pi = \tilde{\boldsymbol{V}}_r(\tilde{\boldsymbol{W}}_r{}^T\tilde{\boldsymbol{V}}_r)^{-1}\tilde{\boldsymbol{W}}_r{}^T$ *where*

$$\tilde{\boldsymbol{V}}_r = [(\sigma_1 \boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}} \dots (\sigma_r \boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}}] \qquad \tilde{\boldsymbol{W}}_r^T = \begin{bmatrix} \tilde{\boldsymbol{c}}^T(\sigma_{r+1}\boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1} \\ \vdots \\ \tilde{\boldsymbol{c}}^T(\sigma_{2r}\boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1} \end{bmatrix}.$$

Then $H_r(s)$ interpolates $H(s)$ at σ_i. Furthermore, $H_r(s) = (\boldsymbol{b}_r - d_r\boldsymbol{e}^T)(s\boldsymbol{I}_r - (\boldsymbol{A}_r - (\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{e}\boldsymbol{e}^T)^{-1}(\boldsymbol{b}_r - d_r(\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{e})$, where $\boldsymbol{A}_r$, $\boldsymbol{b}_r$, $\boldsymbol{c}_r$, are defined as in 1.4.2 applied to $H(s)$.

Proof. $H_r(s)$ interpolates $\tilde{H}(s)$ by 1.4.2. From 2.1.1, it follows that $\tilde{H}(\sigma_i) = H(\sigma_i) = H_r(\sigma_i)$ $i = 1, \dots, 2r$. Also from 2.1.1 we have that $\boldsymbol{W}_r = \tilde{\boldsymbol{W}}_r$ and $\boldsymbol{V}_r = \tilde{\boldsymbol{V}}_r$.
So

$$\begin{aligned} H_r(s) &= \tilde{\boldsymbol{c}}\tilde{\boldsymbol{V}}_r(s\boldsymbol{I}_r - (\tilde{\boldsymbol{W}}_r{}^T\tilde{\boldsymbol{V}}_r)^{-1}\tilde{\boldsymbol{W}}_r{}^T\tilde{\boldsymbol{A}}\tilde{\boldsymbol{V}}_r)^{-1}(\tilde{\boldsymbol{W}}_r{}^T\tilde{\boldsymbol{V}}_r)^{-1}\tilde{\boldsymbol{W}}_r{}^T\tilde{\boldsymbol{b}} \\ &= (\boldsymbol{c}\boldsymbol{V}_r - d_r\boldsymbol{f}^T\boldsymbol{V}_r)(s\boldsymbol{I}_r - (\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{A}\boldsymbol{V}_r + \\ & d_r(\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{g}\boldsymbol{f}^T\boldsymbol{V}_r)^{-1}((\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{b} - d_r\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{g}) \\ &= (\boldsymbol{c}_r - d_r\boldsymbol{e}^T)(s\boldsymbol{I}_r - (\boldsymbol{A}_r + d_r(\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{e}\boldsymbol{e}^T))^{-1}(\boldsymbol{b}_r - d_r(\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{e}). \end{aligned}$$

□

For the remainder of the paper we will denote the transfer function $H_r(s)$ of Corollary 2.1 by $H_r^{d_r}(s)$ to emphasize its dependence on the d_r-term of its state-space realization, and use $H_r(s)$ only to refer to a reduced-order system that does not have any special dependence on d_r. An important result of the above characterization of $H_r^{d_r}(s)$ is that it gives us a parameterization of all reduced-order models that interpolate $H(s)$ at some fixed $2r$ points in $\mathbb{C}$ [17]. This means that for any initial selection of interpolation points, we now have a convenient means of searching for the $H_r^{d_r}(s)$ which provides the minimal error to $H(s)$ in the $\mathcal{H}_\infty$-norm. Indeed, this is the basis for the method of approximation we develop here.

A second important result of this freedom in the d_r term is that it is now possible to construct reduced-order models of order r satsifying $2r+1$ interpolation conditions. This result is not of small consequence; indeed, it is an essential step towards constructing reduced-order models that are minimal in the $\mathcal{H}_\infty$-norm. To see this, consider the following theorem due to Trefethen [23]. In the original work, his argument was in the context of functions holomorphic on the unit disc with singularities outside the disc, but our formulation is specific to rational functions and is equivalent under a conformal mapping of the right half plane to the unit disc.

Theorem 2.1.3. *If $H(s)$ is asymptotically stable, then $H_r(s)$ is an optimal order r approximation to $H(s)$ if $H(s)$-$H_r(s)$ has at least $2r+1$ zeros in the right half plane, and the image of the imaginary axis under $H(s)$-$H_r(s)$ is a perfect circle about the origin.*

Thus, control over $2r+1$ interpolation points of the reduced-order models $H_r^{d_r}(s)$ allows us, at least in principle, to satisfy sufficient conditions for $\mathcal{H}_\infty$-optimal approximations to $H(s)$. With this theorem in hand we may finally begin to attack the $\mathcal{H}_\infty$- optimal problem from an interpolation-based approach. But before we become lost in the ecstasies of this accomplishment, we remember the very sobering fact that the theorems above give us no insight whatsoever into where the optimal interpolation points are located in the right half plane. Indeed, there are very few results available that might give us insight into where these prized points lie, with only certain notable exceptions [3]. In light of this fact, we leave off our merrymaking for the moment, and instead begin experimenting with approximations.

Starting with a set of at least r interpolation points, Algorithm 2.1.1 gives us a simple but computationally expensive method for obtaining the reduced order model $\hat{H}_r^{d_r}(s)$ that both satisfies the interpolation constraints and produces the smallest error in the $\mathcal{H}_\infty$-norm.

Algorithm 2.1.1 (Finding the Optimal Interpolant $\hat{H}_r^{d_r}(s)$).

1. Choose an initial set of 2r interpolation constraints (at least r interpolation points)
2. Compute $\hat{H_r^{d_r}} = \arg\min_{d_r \in \mathbb{R}} \left\| H(s) - H_r^{d_r}(s) \right\|_{\mathcal{H}_\infty}$

Offhand, we note that the approximation to $H(s)$ that consistently yields small errors in the $\mathcal{H}_\infty$-norm is the optimal Hankel norm approximation, but since finding the interpolation points which would give the optimal Hankel norm approximation can only be done by computing the optimal Hankel norm approximation first, it would be better to choose a different approximation to work with. Instead, we take a shot in the dark, and play with the result of starting with the r interpolation points which satisfy the Meier-Luenberger conditions for optimal- $\mathcal{H}_2$ approximations. Recall that if $H_r(s)$ is an optimal $\mathcal{H}_2$ approximation to $H(s)$, then $H_r(s)$ interpolates $H(s)$ and $H_r'(s)$ interpolates $H'(s)$ at the reflection of the poles of $H_r(s)$ across the imaginary axis.

We therefore modify Algorithm 2.1.1 slightly, replacing Step 1.) with Step 1.a) which is to compute the interpolations points which satisfy the necessary conditions for $\mathcal{H}_2$ -optimality via IRKA, and then proceed with the rest of Algorithm 2.1.1 as usual.

2.1.1 Numerical Results

Here are the results of this simple method applied to two Oberwolfach benchmark examples. The first is an order 10 approximation to the canonical CD Player Model of order 120, and the second an order 4 approximation to the Spiral-Inductor PEEC model of order 1434 . We compared our method with balanced truncation, the optimal hankel norm approximation, and the $\mathcal{H}_2$ -optimal approximation.

Table 2.1: Comparison of Order 10 Approximations to Order 120 CD Player Model

Order 10 Approx. ($\sigma_{11} = 4.02 \times 10^{-2}$)			
	Rel. Er.	Abs. Error	D
$H_r^{d_r}(s)$	8.56×10^{-4}	5.86×10^{-2}	-3.57×10^{-2}
Opt. Hankel	1.1×10^{-3}	7.41×10^{-2}	-8.1×10^{-3}
Bal. Trun.	1.3×10^{-3}	9.1×10^{-2}	0
Bal. Trun. D	1.24×10^{-3}	8.5×10^{-2}	-9.6×10^{-3}
Opt. $\mathcal{H}_2$	1.4×10^{-3}	9.38×10^{-2}	0

Table 2.2: Comparison of Order 4 Approximations to Order 1434 Spiral-Inductor Model

Order 4 Approx. ($\sigma_5 = 1.218 \times 10^{-4}$)			
	Rel. Er.	Abs. Error	D
$H_r^{d_r}(s)$	1.51×10^{-4}	2.35×10^{-4}	1.0×10^{-4}
Opt. Hankel	7.93×10^{-5}	1.24×10^{-4}	1.24×10^{-4}
Bal. Trun.	1.78×10^{-4}	2.78×10^{-4}	0
Opt. $\mathcal{H}_2$	2.1×10^{-4}	3.3×10^{-4}	0

These results were obtained by computing the actual $\mathcal{H}_\infty$-norm of the error system at each function call of the optimization routine. However, it was possible to obtain comparable results by computing an approximation to the $\mathcal{H}_\infty$-norm through sampling of frequencies along the imaginary axis. The number of samples was not exhorbinantly large (500 frequencies logarithimically scaled along the imaginary axis). In the case of the CD player model, it produced an absolute error of 5.87e-2. Thus, it is clear that the $\mathcal{H}_\infty$-norm of the $\mathcal{H}_2$ -optimal approximation can be significantly improved upon, and potentially at a small computational cost. The approximation $\hat{H_r^{d_r}}$ to the full order PEEC model yielded similar results, and we found it was possible in both cases to nearly match the $\mathcal{H}_\infty$-error by the optimal d_r-term. In general, significant improvements on the $\mathcal{H}_2$ -optimal approximation $H_r(s)$, appeared to depend heavily on the sensitivity of the poles of $H_r(s)$ to the perturbation which yields the approximation $H_r^{d_r}(s)$. If the approximations $H_r^{d_r}(s)$ were stable in only a very small range

of values for d_r relative to the lower bound on the approximation error, the approximation $H_r^{d_r}(s)$ tended to yield only minor improvements. The CD player model, and PEEC model are examples of systems for which there was significant freedom in the d_r-term for the IRKA interpolation points.

Pursuing this method of approximation in its own right, we will side step the investigation of $\mathcal{H}_\infty$-optimal interpolation points for the remainder of the chapter, and instead consider whether there is a cheaper way to find the optimal d_r-term. Further experiments for a variety of hard to approximate dynamical systems makes it reasonable to choose the $\mathcal{H}_2$ -optimal interpolation points as the starting point of the d_r-term minimization. Therefore we will restrict ourselves to this initialization of Algorithm 2.1.1, and further look at the special case of finding the optimal d_r-term for the class of state-space symmetric systems.

2.2 The optimal d_r-term for state-space symmetric systems

The systems that we will focus on are called state-space symmetric (SSS). A dynamical system $H(s)=(\boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{c},\ \mathrm{d})$ is SSS provided that $\boldsymbol{A}=\boldsymbol{A}^T$, and $\boldsymbol{b} = \boldsymbol{c}^T$. The relaxation dynamics which characterize SSS systems make them particularly easy to approximate by systems of very low order, and thus our interest lies primarily in minimizing the cost of the model reduction. SSS systems regularly occur on a very large scale in partial element equivalent circuits (PEEC) that arise from a partial discretization of the Maxwell's equations governing inductance in a circuit. The spiral inductor model we approximated in Section 2.1 provides a good example of an SSS system. The results developed here are motivated by the repeated and consistent observation that if the interpolation points $\{\sigma_i\}_{i=1}^r$ satisfied the

necessary conditions for $\mathcal{H}_2$ optimality, then $\tilde{H}_r^{d_r}$ solved the following minimization problem

$$\tilde{H}_r^{d_r} = \arg\min_{d_r \in \mathbb{R}} \left\| H(s) - H_r^{d_r}(s) \right\|_{\mathcal{H}_\infty} \tag{2.5}$$

provided there exists some point $i\tilde{\omega} \neq 0$ such that $\left\| \tilde{H}_r^{d_r}(s) \right\|_{\mathcal{H}_\infty} = \tilde{H}_r^{d_r}(i\tilde{\omega}) = \tilde{H}_r^{d_r}(0)$. Moreover, the point $i\tilde{\omega}$ generally corresponded to the point at ∞. This observation suggested that finding $\tilde{H}_r^{d_r}(s)$ was equivalent in most cases to finding the the value of d_r-term for which $H_r^{d_r}(0) = d_r$. Searching for the d_r-term which satisfied this condition would therefore allow us to avoid the very costly computation of the $\mathcal{H}_\infty$-norm at each step of the minimization routine.

We first consider SSS systems having the zeros-interlacing-the-poles (ZIP) property.

Definition 1. A system

$$H(s) = K \frac{\prod_{i=1}^{n-1} (s - z_i)}{\prod_{j=1}^{n} (s - \lambda_j)} \tag{2.6}$$

is a strictly proper ZIP system provided that $\lambda_1 < z_1 < \lambda_2 < z_2 < \lambda_3 < \cdots < z_{n-1} < \lambda_n$. If $H(s)$ is minimal and $|H(\infty)| = K > 0$ then $H(s)$ is ZIP provided that it can be factored as:

$$H(s) = K + P(s)/D(s), \tag{2.7}$$

where $P(s)/D(s)$ is a strictly proper ZIP system.

ZIP systems have a number of interesting simplifying properties which we will find useful in our search to find the approximation $H_r^{d_r}(s)$ that produces the minimal error. We first prove a new result which shows every state-space symmetric system with real coefficients is either a ZIP system, or reducible to a minimal system which is ZIP. We will then exploit this pole-zero

structure to give a new proof of the important result that if $H(s)$ is ZIP, then $\|H(s)\|_{\mathcal{H}_\infty} = |H(0)|$. First, we will need a number of alternative but equivalent characterizations of ZIP systems summarized in the following theorem due to B. Srinivasan, and P. Myszkorowski [21].

Proposition 2.2.1 (**Characterization of ZIP Systems** [21])**.** *The following statements are equivalent:*

(i) $H(s)$ *is a strictly proper ZIP system*

(ii) $H(s)$ *can be written as*

$$H(s) = \sum_{i=1}^{n} \frac{b_i}{s - \lambda_i} \tag{2.8}$$

with $\lambda_i < 0,\ b_i > 0\ \lambda_i \neq \lambda_j\ \forall i \neq j$.

(iii) $H(s)$ *has a diagonal realization* $H(s)$=($\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$, 0)*, where* $\boldsymbol{A}$=$diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$*, and* $\boldsymbol{b}^T = \boldsymbol{c} = [\sqrt{b_1}, \sqrt{b_2}, \ldots, \sqrt{b_n}]$.

(iv) $H(s)$ *has a SSS principle axis balanced realization.*

Lemma 2.2.2. *If* $H(s)$ *is SSS and* $H(\infty) = 0$*, then* $H(s)$ *is minimal if and only if the poles of* $H(s)$ *are distinct. Further, every SSS system* $H(s)$ *satisfying* $H(\infty) = 0$ *has a SSS minimal realization with distinct poles, and is therefore a strictly proper ZIP system.*

Proof. Suppose to the contrary that $H(s)$=($\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$, d) is SSS and minimal and has a pole λ^* of order $k > 1$. $\boldsymbol{A}$ is symmetric means it is normal, and therefore has a unitary diagonalization $\boldsymbol{A}$=$\boldsymbol{U}^T\boldsymbol{\Lambda}\boldsymbol{U}$. Thus, $H(s) = (\boldsymbol{\Lambda}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{b}}^T)$, where $\tilde{\boldsymbol{b}}(i) = \sqrt{|\boldsymbol{U}^T\boldsymbol{b}(i)|}$, and $\boldsymbol{b}(i)$ denotes the *ith* entry of the vector $\boldsymbol{b}$. Therefore $H(s)$ can be written as:

$$H(s) = \sum_{i=1}^{n} \frac{\tilde{\boldsymbol{b}(i)}^2}{s - \lambda_i} \tag{2.9}$$

$$= \sum_{l=1}^{k} \frac{\tilde{\boldsymbol{b}}(i_l)^2}{s - \lambda^*} + \sum_{j=1}^{n-k} \frac{\tilde{\boldsymbol{b}}(i_j)^2}{s - \lambda_{i_j}} \tag{2.10}$$

$$= \frac{\sum_{i=1}^{k} \tilde{\boldsymbol{b}}(i_l)^2}{s - \lambda^*} + \sum_{j=1}^{n-k} \frac{\tilde{\boldsymbol{b}}(i_j)^2}{s - \lambda_{i_j}} \tag{2.11}$$

Thus, $H(s)$ is not minimal, which is a contradiction. Now suppose that the poles of $H(s)$ are all distinct. Then, since $\boldsymbol{A}$ is normal, $H(s)$ can be written as in (iii) of Proposition 2.2.1, and is therefore minimal. It follows that if $H(s)$ is not minimal, it may be reduced to a minimal system given by equation 2.11 with a SSS realization given by (ii) of Proposition 2.2.1, and is therefore a strictly proper ZIP system. □

Lemma 2.2.2 implies that any properties which hold for the transfer function of a ZIP system must hold for the transfer function of a SSS system. We make use of this fact in the proof of Theorem 2.2.3 given below.

Theorem 2.2.3. *If $H(s)$ is asympotically stable and SSS with $H(\infty) = 0$, then $\|H(s)\|_{\mathcal{H}_\infty} = |H(0)|$*

Proof. We may assume by Lemma 2.2.2 that $H(s)$ is ZIP in the sense that it has a minimal realization which is ZIP. Thus, it suffices to show that $|H(i\omega)|$ has a global max at $\omega = 0$. Let $f(\omega) = \log(|H(i\omega)|)$. Then

$$f(\omega) = \log(K) + \sum_{i=1}^{n-1} \log(|i\omega + z_i|) - \sum_{j=1}^{n+1} \log|(i\omega + \lambda_i)| \tag{2.12}$$

and therefore

$$f'(\omega) = \sum_{i=1}^{n-1} \frac{\omega}{z_i^2 + \omega^2} - \sum_{j=1}^{n} \frac{\omega}{\lambda_i^2 + \omega^2}. \tag{2.13}$$

It is clear from (2.13) that $f'(0) = 0$, and the ZIP property guarantees that for $\omega > 0$, $f'(\omega) < 0$, since $\frac{\omega}{\omega^2+z_i^2} < \frac{\omega}{\omega^2+\lambda_i^2}$. Multiplying through this inequality by -1, we see that when $\omega < 0$, $f'(\omega) > 0$. Thus $f(0)$ is a global maximum, hence $|H(0)|$ is the global maximum, and therefore $|H(0)| = \|H(s)\|_{\mathcal{H}_\infty}$. □

The results of [16] show that if $H(s)$ is SSS and $H_r(s)$ is computed by balanced truncation, or if $H_r(s)$ is an optimal Hankel norm approximation to $H(s)$ then $H_r(s)$ is also SSS, and therefore ZIP. Ideally then, we would like to find reduced order models by interpolation that also preserve the ZIP property. When $H(s)$ is SSS and the shifts used to construct $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$ are equal, then $\boldsymbol{V}_r$=$\boldsymbol{W}_r$. Let $\boldsymbol{Q}_r\boldsymbol{R} = \boldsymbol{V}_r$ be the QR decomposition of $\boldsymbol{V}_r$. Then we may use $\boldsymbol{Q}_r$ as our projection matrix in place of $\boldsymbol{V}_r$ by Proposition 1.4.1.

The reduced order model $H_r(s)$ constructed from the Rational Krylov method then becomes

$$H_r(s) = \boldsymbol{b}^T\boldsymbol{Q}_r(s\boldsymbol{I} - \boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T\boldsymbol{b} \tag{2.14}$$

and $H_r(s)$ is clearly SSS. The family of systems $H_r^{d_r}(s)$ parameterized by the d_r-term which interpolate $H(s)$ at the same $2r$ interpolation points as $H_r(s)$ can easily be written in terms of the projection $\boldsymbol{Q}_r\boldsymbol{Q}_r^T$. Let $\boldsymbol{f}_Q$ and $\boldsymbol{g}_Q$ be the solutions to

$$\boldsymbol{f}_Q^T\boldsymbol{Q}_r = \boldsymbol{e}^T\boldsymbol{R}^{-1}, \qquad \boldsymbol{Q}_r^T\boldsymbol{g}_{\boldsymbol{Q}} = \boldsymbol{R}^{-T}\boldsymbol{e}, \tag{2.15}$$

and let

$$\boldsymbol{e}_Q = \boldsymbol{R}^{-T}\boldsymbol{e}. \tag{2.16}$$

Then substituting these expressions into the original equation for $H_r^{d_r}(s)$ gives:

$$\begin{aligned} H_r^{d_r}(s) &= (\boldsymbol{b}^T\boldsymbol{Q}_r\boldsymbol{R} - d_r\boldsymbol{f}^T\boldsymbol{Q}_r\boldsymbol{R})(s\boldsymbol{R}^T\boldsymbol{Q}_r^T\boldsymbol{Q}_r\boldsymbol{R}- \qquad (2.17)\\ &\quad \boldsymbol{R}^T\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r\boldsymbol{R} - d_r\boldsymbol{R}^T\boldsymbol{Q}_r^T\boldsymbol{g}\boldsymbol{f}^T\boldsymbol{Q}_r\boldsymbol{R})^{-1}(\boldsymbol{R}^T\boldsymbol{Q}_r^T\boldsymbol{b} - d_r\boldsymbol{R}^T\boldsymbol{Q}_r^T\boldsymbol{g}) + d_r \\ &= (\boldsymbol{b}^T\boldsymbol{Q}_r - d_r\boldsymbol{e}_Q^T)(s\boldsymbol{I} - (\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r + d_r\boldsymbol{e}_Q\boldsymbol{e}_Q^T))^{-1}(\boldsymbol{Q}_r^T\boldsymbol{b} - d_r\boldsymbol{e}_Q) + d_r. \qquad (2.18) \end{aligned}$$

We summarize these results in the following proposition.

Proposition 2.2.4. *Given a stable, order n, SSS system with transfer function $H(s)$ having simple poles, every reduced-order system $H_r(s)$ of order r, constructed by a one sided projection $\boldsymbol{Q}_r$ having orthonormal columns, is also state-space symmetric and has the ZIP property. If $H_r(s)$ has a minimal state space realization $H_r(s) = (\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r, \boldsymbol{Q}_r^T\boldsymbol{b}, \boldsymbol{b}^T\boldsymbol{Q}_r, 0)$, then the perturbed system $H_r^{d_r}(s)$ with state space realization $H_r^{d_r}(s) = (\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r + d_r\boldsymbol{e}\boldsymbol{e}^T, \boldsymbol{Q}_r^T\boldsymbol{b} - d_r\boldsymbol{e}, \boldsymbol{b}^T\boldsymbol{Q}_r - d_r\boldsymbol{e}, d_r)$ is also state-space symmetric and ZIP, but not strictly proper.*

Proof. It is clear that both $H_r(s)$ and $H_r^{d_r}(s)$ have state-space symmetric realizations. It only remains to show that their respective poles are simple, and the proposition follows by Lemma 2.2.2. Since the poles of $H(s)$ are simple, the eigenvalues of $\boldsymbol{A}$ are distinct, and since $\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{A}_r$ is a compression of $\boldsymbol{A}$, its eigenvalues are distinct by the Cauchy interlacing theorem. Similarly since the eigenvalues of $\boldsymbol{A}_r$ are distinct, then by the Weyl Interlacing theorems the eigenvalues of $\boldsymbol{A}_r + d_r\boldsymbol{e}\boldsymbol{e}^T$ are distinct, and provided $\boldsymbol{A}_r + d_r\boldsymbol{e}\boldsymbol{e}^T < 0$ the system remains stable. $H_r^{d_r}(s)$ is therefore ZIP, but not strictly proper, since it can be written as a strictly proper ZIP system plus its D-term. □

In the SSS case, the necessary conditions for optimal $\mathcal{H}_2$ interpolation points force $\boldsymbol{V}_r$=$\boldsymbol{W}_r$

in IRKA. Thus if $H(s)$ is SSS, and $H_r(s)$ is a reduced order model computed by IRKA, then $H_r(s)$ is also SSS. It follows that $H_r(s)$ inherits all the ZIP properties of the original system. Moreover, if $H_r(s)$ is computed from IRKA, we may view it as the result of a compression of the full order state-space by taking an orthonormal basis for $\boldsymbol{V}_r$ for the underlying projection, so $H_r(s)$ is a special case of a more general method of model reduction done by compression. For model reduction done by compression we obtain the following upper bound on the $\mathcal{H}_\infty$-norm of the reduced order model.

Lemma 2.2.5. *If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is hermitian negative definite and $\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r$ is a compression of $\boldsymbol{A}$, then $\boldsymbol{A}_Q = -\boldsymbol{A}^{-1} + \boldsymbol{Q}_r(\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T \geq 0$ and if rank($\boldsymbol{Q}_r$)=r, then rank($\boldsymbol{A}_Q$) $= n - r$.*

Proof. Let -$\boldsymbol{A}$= $\boldsymbol{L}\boldsymbol{L}^T$ be the Cholesky factorization of -$\boldsymbol{A}$. Then

$$\boldsymbol{A}_Q = \boldsymbol{L}^{-T}\boldsymbol{L}^{-1} - \boldsymbol{Q}_r(\boldsymbol{Q}_r^T\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T.$$

Factoring out $\boldsymbol{L}^{-T}$ on the left and $\boldsymbol{L}$ on the right we have

$$\boldsymbol{A}_Q = \boldsymbol{L}^{-T}(\boldsymbol{I} - \boldsymbol{L}^T\boldsymbol{Q}_r(\boldsymbol{Q}_r^T\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T\boldsymbol{L})\boldsymbol{L}^{-1}.$$

By a theorem in [18],

$$(\boldsymbol{I} - \boldsymbol{L}^T\boldsymbol{Q}_r(\boldsymbol{Q}_r^T\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T\boldsymbol{L})$$

is positive semi-definite and of rank $n - r$, and the theorem follows. □

Corollary 2.2.6. *If $H(s)$ is SSS and $H_r(s)$ is a reduced order model constructed from a compression of the full order state-space, then $\|H_r(s)\|_{\mathcal{H}_\infty} \leq \|H(s)\|_{\mathcal{H}_\infty}$*

Proof. Both $H(s)$ and $H_r(s)$ are SSS, and therefore $H(0) = \|H(s)\|_{\mathcal{H}_\infty}$ and $H_r(0) =$

$\|H_r(s)\|_{\mathcal{H}_\infty}$. Thus, it suffices to show that $H(0) - H_r(0) \geq 0$.

$$H(0) - H_r(0) = -\boldsymbol{b}^T\boldsymbol{A}^{-1}\boldsymbol{b} + \boldsymbol{b}^T\boldsymbol{Q}_r(\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T\boldsymbol{b} \tag{2.19}$$

$$= \boldsymbol{b}^T(-\boldsymbol{A}^{-1} + \boldsymbol{Q}_r(\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r)^{-1}\boldsymbol{Q}_r^T)\boldsymbol{b} \tag{2.20}$$

$$\geq 0 \text{ by Lemma 2.2.5.}$$

□

With this result, and the two lemmas that follow we will be able to produce the main result of this section.

The first lemma is a curious property of IRKA applied to SSS systems.

Lemma 2.2.7. *Suppose $H_r(s)$ is computed by IRKA applied to the full order SSS system $H(s)$. Assume that in place of $\boldsymbol{V}_r$, we use $\boldsymbol{Q}_r$ as our projection basis, and let $\boldsymbol{A}_r$, $\boldsymbol{e}_Q$, $\boldsymbol{b}_r$ be as in 2.18. Then $\boldsymbol{e}_Q^T\boldsymbol{A}_r^{-1}\boldsymbol{b}_r$=0 if r is even and -2 if r is odd.*

Proof. From [4] we have that

$$\boldsymbol{A}_r = (\boldsymbol{\Sigma} - \boldsymbol{q}\boldsymbol{e}^T) \tag{2.21}$$

Where Σ=diag$(\sigma_1, \ldots, \sigma_r)$, $\boldsymbol{e}$ is a column of r ones and $\boldsymbol{q} = (\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{V}_r^T\boldsymbol{b}$, with $\boldsymbol{W}_r$ and $\boldsymbol{V}_r$ defined as in the Rational Krylov method. It follows from 2.21 that

$$\boldsymbol{I} - \boldsymbol{A}_r^{-1}\boldsymbol{\Sigma} - \boldsymbol{A}_r^{-1}\boldsymbol{b}_r\boldsymbol{e}^T = \boldsymbol{0} \tag{2.22}$$

$$\boldsymbol{e}^T + \boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{\Sigma} - \boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{b}_r\boldsymbol{e}^T = \boldsymbol{0}. \tag{2.23}$$

If $\boldsymbol{e}^T$ is a left eigenvector of $\boldsymbol{A}_r^{-1}\boldsymbol{\Sigma}$ with eigenvalue -1 for r even and 1 for r odd the Lemma

follows. Making use of the Shurman-Morrison identity we have:

$$\begin{aligned} -\boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{\Sigma} &= \boldsymbol{e}^T(\boldsymbol{\Sigma}^{-1} + \frac{1}{1-\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{q}}\boldsymbol{\Sigma}^{-1}\boldsymbol{q}\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}) - \boldsymbol{\Sigma}^{-1} \\ &= -(\boldsymbol{e}^T + \frac{\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{q}\boldsymbol{e}^T}{1-\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{q}}) \\ &= \frac{-\boldsymbol{e}^T}{1-\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{q}}. \end{aligned} \tag{2.24}$$

Thus for $\boldsymbol{e}^T$ to be a left eigenvector of $\boldsymbol{A}_r^{-1}\boldsymbol{\Sigma}$ with eigenvalue -1 for r even, 1 for r odd, $\boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{q}$ must be 0 for r even and 2 for r odd. From [4] we have that $\mathrm{q}_i = p(\sigma_i)/w'(\sigma_i)$ where p=det$(\boldsymbol{I}z - \boldsymbol{\Sigma})$ is a monic polynomial of degree r, and $w'(\sigma_i) = \Pi_{j\neq i}^r(\sigma_i - \sigma_j)$. Since p is monic and of degree r, $p(z) - z^r$ is a polynomial of degree $r-1$ and is therefore uniquely determined by the points $\sigma_1, \ldots, \sigma_r$. If $w(z) = \Pi_{i=1}^r(z-\sigma_i)$ then we have that

$$p(z) - z^r = \sum_{i=1}^r (p(\sigma_i) - \sigma_i^r)(\frac{w(z)}{w'(\sigma_i)(z-\sigma_i)}) \tag{2.25}$$

$$= \sum_{i=1}^r \frac{p(\sigma_i)w(z)}{w'(\sigma_i)(z-\sigma_i)} - \sum_{i=1}^r \frac{\sigma_i^r w(z)}{w'(\sigma_i)(z-\sigma_i)}. \tag{2.26}$$

Evaluating $p(z)$ at 0 we have:

$$\Pi_{i=1}^r\sigma_i = ((-1)^{r+1}\Pi_{i=1}^r\sigma_i)(\sum_{i=1}^r \frac{p(\sigma_i)}{w'(\sigma_i)\sigma_i} - \sum_{i=1}^r \frac{\sigma_i^{r-1}}{w'(\sigma_i)}) \tag{2.27}$$

implies that:

$$(-1)^{r+1} = \boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{q} - \sum_{i=1}^r \frac{\sigma_i^{r-1}}{w'(\sigma_i)}. \tag{2.28}$$

Finally we reduce the conclusion to demonstrating that

$$\sum_{i=1}^{r} \frac{\sigma_i^{r-1}}{w'(\sigma_i)} = 1. \tag{2.29}$$

This amounts to demonstrating that 2.29 is a rather odd way of writing $\det(\mathbf{V})/\det(\mathbf{V})$, where $\mathbf{V}$ is the Vandermonde matrix

$$\mathbf{V} = \begin{pmatrix} 1 & \sigma_1 & \sigma_1^2 & \dots & \sigma_1^{r-1} \\ 1 & \sigma_2 & \sigma_2^2 & \dots & \sigma_2^{r-1} \\ 1 & \sigma_3 & \sigma_3^2 & \dots & \sigma_3^{r-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r^2 & \sigma_r^3 & \dots & \sigma_r^{r-1} \end{pmatrix}. \tag{2.30}$$

Since all the shifts are distinct, this determinant is nonzero, so $\det(\mathbf{V})/\det(\mathbf{V})$ makes sense.

$$\sum_{i=1}^{r} \frac{\sigma_i^{r-1}}{w'(\sigma_i)} = \sum_{k=1}^{r} \frac{(-1)^{r+1}\sigma_k^{r-k}}{\prod_{j=1}^{k-1}(\sigma_k - \sigma_j)\prod_{j=k+1}^{r}(\sigma_j - \sigma_k)} \tag{2.31}$$

thus, for a fixed term k^* and $k < k^*$, the terms in the denominator of k that are absent from k^* are $\prod_{j=k+1}^{r}(\sigma_j - \sigma_k)$, $j \neq k^*$ Assumming that terms missing from the denominator of k^* are added sequentially by first adding terms for $k = 1$, then for $k = 2$, ... and so on then for $k > k^*$, terms of $\prod_{j=1}^{k-1}(\sigma_j - \sigma_k)$ will already be included for each $j < k$ and we need only include the terms $\prod_{j=k+1}^{r}(\sigma_j - \sigma_k)$ in the denominator of the k^*th summand. Thus, arranging this sum with a common denominator gives:

$$\sum_{k=1}^{r} \frac{\sigma_k^{r-1}}{w'(\sigma_k)} = \sum_{k=1}^{r} \frac{(-1)^{r-k}\sigma_k^{r-1} \prod\limits_{1\le i<j\le r, i\neq k} (\sigma_j - \sigma_i)}{\prod\limits_{1\le i<j\le r} (\sigma_j - \sigma_i)} = \frac{\sum\limits_{k=1}^{r} (-1)^{r-k}\sigma_k^{r-1} \prod\limits_{1\le i<j\le r, i\neq k} (\sigma_j - \sigma_i)}{\prod\limits_{1\le i<j\le r} (\sigma_j - \sigma_i)} \tag{2.32}$$

$$= \det(\mathbf{V})/\det(\mathbf{V}). \tag{2.33}$$

Where 2.33 follows from the well known equality $\det \mathbf{V} = \prod\limits_{1\le i<j\le r} (\sigma_j - \sigma_i)$ which occurs in the denominator, and the fact that if we consider the minors of $\mathbf{V}$ found by eliminating the rth column and kth row of $\mathbf{V}$, these minors are $r-1$ by $r-1$ Vandermonde matrices, so for a fixed k, the determinant of the kth minor is given by $\prod\limits_{1\le i<j\le r, i\neq k} (\sigma_j - \sigma_i)$. Thus we may view the sum in the numerator as computation of $\det \mathbf{V}$ by expansion by minors, and the result follows. □

This curious result is crucial to understanding the behavior of the function $f(d_r) = H_r^{d_r}(0)$. We summarize some of the key properties of $f(d_r)$ in the lemma below.

Lemma 2.2.8. *Let $H_r(s)$ be computed from IRKA. Define $f(d_r) = -(\boldsymbol{b}_r^T - d_r\boldsymbol{e}^T)(\boldsymbol{A}_r - d_r\boldsymbol{e}\boldsymbol{e}^t)^{-1}(\boldsymbol{b}_r - d_r\boldsymbol{e})$, where $f(0) = H_r(0)$. Then $f'(0) = 0$, f is convex, and for $d_r > 0$, $f(d_r) > f(-d_r)$.*

Proof. Let

$$v = \boldsymbol{b}_r^T(\boldsymbol{A}_r + d_r\boldsymbol{e}\boldsymbol{e}^T)^{-1}\boldsymbol{e} \tag{2.34}$$

and

$$u = -\boldsymbol{e}^T(\boldsymbol{A}_r + d_r\boldsymbol{e}\boldsymbol{e}^T)^{-1}\boldsymbol{e}. \tag{2.35}$$

.

Then

$$f'(d_r) = v^2 + 2v + 2d_r uv + 2d_r u + d_r^2 u^2, \tag{2.36}$$

so $f'(0) = (\boldsymbol{b}_r^T \boldsymbol{A}_r^{-1}\boldsymbol{e})^2 + 2\boldsymbol{b}_r^T \boldsymbol{A}_r^{-1}\boldsymbol{e}$. By Lemma 2.2.7, $f'(0)$ is therefore 0, regardless of the parity of the shifts.

$$f''(d_r) = 2uv^2 + 4uv + 4d_r u^2 v + 2u + 4d_r u^2 + 2d_r^2 u^3 \tag{2.37}$$

Using the fact that

$$v = \boldsymbol{b}_r^T(\boldsymbol{A}_r + d_r \boldsymbol{e}\boldsymbol{e}^T)^{-1}\boldsymbol{e} = \boldsymbol{b}_r^T \boldsymbol{A}_r^{-1}\boldsymbol{e} - \frac{d_r \boldsymbol{b}_r^T \boldsymbol{A}_r^{-1}\boldsymbol{e}\boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}}{1 + d_r \boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}},$$

due to the Shurman-Morrison identity, we see that $v = 0$ whenever r is even. Note that $u > 0$ for all d_r since$(\boldsymbol{A}_r + d_r \boldsymbol{e}\boldsymbol{e}^T)$ is negative definite. Therefore, for an even number of shifts, $f''(d_r)$ is the sum of nonnegative terms and f is convex. When r is odd,

$$v = -2 + \frac{2d_r \boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}}{1 + d_r \boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}},$$

and

$$u = -\boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e} + \frac{d_r(\boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}\boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e})}{1 + d_r \boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}}.$$

Substituting u and v into the equation for f'' and simplifying we find that

$$f''(d_r) = \frac{-2\boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e}}{(1 + d_r \boldsymbol{e}^T \boldsymbol{A}_r^{-1}\boldsymbol{e})^3} > 0$$

for all d_r in the domain of f. Hence, f is convex regardless of the parity of the shifts, and since $f'(0) = 0$, f is increasing for $d_r > 0$ and decreasing for $d_r < 0$. Now suppose $d_r > 0$. If

r is even, then after some tedious algebraic simplification we have

$$f(d_r) - f(-d_r) = \frac{d_r^3(\boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e})^2}{1 + d_r\boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e}} + \frac{d_r^3(\boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e})^2}{1 - d_r\boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e}} > 0$$

for all d_r in the domain of f, since d_r is in the domain with $\boldsymbol{A}_r + d_r\boldsymbol{e}\boldsymbol{e}^T$ negative definite is equivalent, by the Shurman-Morrison identity, to $1+d_r\boldsymbol{e}^T\boldsymbol{A}_r\boldsymbol{e} > 0$ for $d_r > 0$. If r is odd, then again after simplifying we have

$$f(d_r) - f(-d_r) = \frac{(1 + d_r)d_r^3(\boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e})^3(1 - \boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e})}{-1 - d_r\boldsymbol{e}^T\boldsymbol{A}_r^{-1}\boldsymbol{e}} > 0$$

for all d_r in the domain of f. □

In broad strokes the picture afforded by the above lemmas is that finding the optimal d_r-term involves sliding the the value of $H(0)-H_r^{d_r}(0)$ down by increasing the value of d_r which at the same time increases $|H(\infty) - H_r^{d_r}(\infty)|$. Assuming the error at these two points dominates the error along the imaginary axis for all values of d_r minimizing the error then amounts to finding the d_r-term for which the error system at zero is equal to the error system at infinity. When the error at some finite point along the imaginary axis grows larger than the value of the error at infinity, we obtain the result that to match the error at 0 with the error at this interior point is locally optimal.

Theorem 2.2.9. *Let σ_{r+1} be the $r+1$th Hankel singular value of the SSS system $H(s)$. Let $H_r(s)$be a reduced ordered model computed by IRKA. If $d_r{}^* \geq \sigma_{r+1}$ and $H_r^{d_r^*}(s)$ is asymptotically stable with $\left\|H(s) - H_r^{d_r^*}(s)\right\|_{\mathcal{H}_\infty} = d_r^*$, and $H(0) - H_r^{d_r^*}(0) = d_r^*$, then $d_r^* = \arg\min_{d_r\in\mathbb{R}} \left\|H(s) - H_r^{d_r}(s)\right\|_{\mathcal{H}_\infty}$.*

Proof. Suppose $0 \leq D < d_r^*$ and $\left\|H(s) - H_r^D\right\|_{\mathcal{H}_\infty} \leq \left\|H(s) - H_r^{d_r^*}(s)\right\|_{\mathcal{H}_\infty}$. From Lemma

2.2.8, $H_r^D(0) < H_r^{d_r^*}(0)$, so it follows that $H(0) - H_r^D(0) > H - H_r^{d_r^*}(0) = d_r^*$. But this contradicts the original assumption. If $-d_r^* \leq D \leq 0$, we have $H(0) - H_r^D(0) > H(0) - H_r^{-D}(0) > H(0) - H_r^{d_r^*}(0)$, also by Lemma 2. Hence, $d_r^* = \min_{d_r \in \mathbb{R}} \left\| H(s) - H_r^{d_r}(s) \right\|_{\mathcal{H}_\infty}$. □

Theorem 2.2.10. *Suppose $H_r(s)$ is a reduced order model of $H(s)$ computed by IRKA, and let $E_{d_r} = H(s) - H_r^{d_r}(s)$. Let $D^0 = \{d_r \geq 0 : E_{d_r}(0) = \|E_{d_r}(s)\|_{\mathcal{H}_\infty}$ and $H_r^{d_r}(s)$ is stable$\}$ and assume $D^0 \neq \emptyset$. If $d_r^* = \sup D^0$ then $d_r^* \in D^0$, and there exists some point $i\omega^*$ such that $E_{d_r^*}(0) = E_{d_r^*}(i\omega^*) = \|E_{d_r}(s)\|_{\mathcal{H}_\infty}$. Furthermore, if $-d_r^* < d_r < d_r^*$, then $\|E_{d_r}(s)\|_{\mathcal{H}_\infty} > \left\|E_{d_r^*}(s)\right\|_{\mathcal{H}_\infty}$.*

Proof. Let $H_r^{d_r}(s) = H(d_r, s) = f(d_r)g(d_r, s)h(d_r)$ where $f(d_r) = \boldsymbol{c}_r - d_r\boldsymbol{e}^T$, $g(d_r, s) = (s\boldsymbol{I} - \boldsymbol{A}_r - d_r\boldsymbol{e}\boldsymbol{e}^T)^{-1}$ and $h(d_r) = \boldsymbol{b}_r - d_r\boldsymbol{e}$. h and f are clearly continuous functions of d_r, and g is continuous as a function of d_r due to the fact that the eigenvalues of $\boldsymbol{A}_r + d_r\boldsymbol{e}\boldsymbol{e}^T$ are continuous functions of d_r. Thus $H(d_r, s)$ is a continuous function of d_r. The poles of $H_r^{d_r}(s)$ are increasing as d_r increases, but since $H_r^{d_r^*}(s)$ is stable, $H(d_r, s)$ is holomorphic in a full half-pslane for $Re(z) > \mu$ and $\mu < 0$. Thus, $H(d_r, s)$ is holomorphic along the imaginary axis. It follows that for some $\epsilon > 0$, if $d_r \in [d_r^* - \epsilon, d_r^* + \epsilon]$ then the curve $|E_{d_r}(i\omega)| = |H(i\omega) - H(d_r, i\omega)|$ is homotopic to the curve $|E_{d_r^*}(i\omega)|$. So suppose that $|E_{d_r^*}(0)| - |E_{d_r^*}(i\omega)| = s < 0$ for some $\omega \in \mathbb{R}_+$. Let $\delta < \max\{\epsilon, s/2\}$. Then there exists some $d_r \in D^0$ such that $\big|\, |E_{d_r}(i\omega)| - |E_{d_r^*}(i\omega)| \,\big| < \delta$ and $E_{d_r}(0) < E_{d_r^*}(0) + \delta$ by Lemma 2.2.8 and the fact that we can choose $d_r \in [d_r^* - \delta, d_r^* + \delta]$. But $\big|\, |E_{d_r}(i\omega)| - |E_{d_r^*}(i\omega)| \,\big| < s/2 \Rightarrow E_{d_r}(0) < |E_{d_r}(i\omega)|$, which contradicts our assumption that $d_r \in D^0$. Thus, $|E_{d_r^*}(0)| - |E_{d_r^*}(i\omega)| = s \geq 0$, $\forall \omega \in \mathbb{R}$, and $d_r^* \in D^0$.

$|E_{d_r}(s)|^2$ is a rational function, so it oscillates a fixed number of times depending on the zeros of $H_r^{d_r}(s)$. By Rouche's Theorem, the zeros of $H_r^{d_r}(s)$ are continuous in d_r, so for d_r sufficiently close to d_r^*, the local extrema of $E_{d_r}(s)$ will occur in a neighborhood of the local extrema of $E_{d_r^*}(s)$. Let $E_{d_r^*}(0) - E_{d_r^*}(i\omega_{\max}) = u > 0$ where $i\omega_{\max}$ is the next largest local

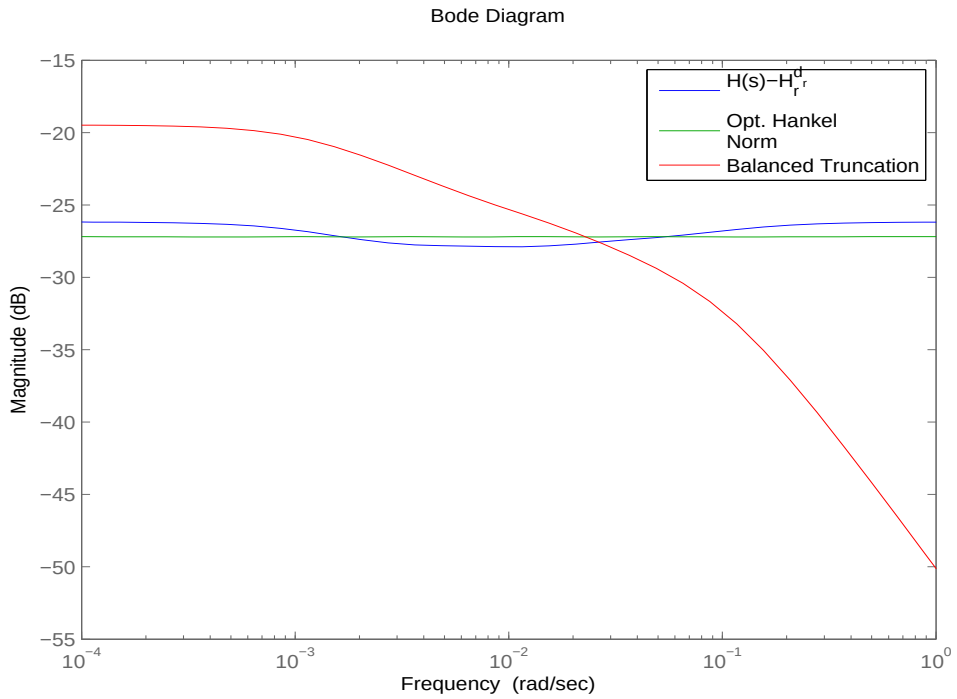
maximum compared to $E_{d_r^*}(0)$. Then choose $\delta > 0$ so that if $|d_r^* - d_r| < \delta \rightarrow ||E_{d_r}(i\omega)| - |E_{d_r^*}(i\omega)|| < u/2$. Then for any d_r in the interval, $(d_r^*, d_r + \delta), d_r \in D^0$ which contradicts the assumption that $d_r^* = \sup D^0$. Thus $u = 0$. Finally, it is clear that the $\mathcal{H}_\infty$-norm of any $E_{d_r}(s)$ with $-d_r^* < d_r < d_r^*$ is at most $E_{d_r}(0)$, and for every d_r in this interval $E_{d_r}(0) > E_{d_r^*}(0) = \|E_{d_r *}(s)\|_{\mathcal{H}_\infty}$ □

These results suggests a cheap and practical method for searching for the optimal $H_r^{d_r}(s)$ that does not require computing the $\mathcal{H}_\infty$ norm. Namely, we start by finding the d_r-term that satisfies the condition $H(0) - H_r^{d_r}(0) = d_r$ We then sample the imaginary axis one time, comparing the values of the error system at the sample points to the values at 0 and infinity. If there is some value in the interior of the interval from 0 to infinity which is larger than the value at the endpoints, we take a small collection of sample points in a neighborhood about this maximum, and decrease the d_r-term until the value of the error system at zero is equal to the max over this collection of sample points. Some examples below show the results of this method.

Example 2.2.1. Let $H(s)$=($\boldsymbol{A}$,$\boldsymbol{b}$,$\boldsymbol{c}$, 0) of order 400 be SSS where the eigenvalues of $\boldsymbol{A}$ are logarithmically spaced in the interval [-1e-1,-1e-3], and $\boldsymbol{b}$=ones(400,1). This system was chosen due to the fact that the Hankel singular values of $H(s)$decay slowly, so it more difficult to approximate. The order of approximation is $r = 4$. A system $\tilde{H}_r^{d_r}(s)$ was found which satisfied the conditions of Theorem 2.2.9, and is therefore optimal over the family of systems $H_r^{d_r}(s)$. The $\mathcal{H}_\infty$-norm of the error was $\left\|H(s) - \hat{H}_r^{d_r}(s)\right\|_{\mathcal{H}_\infty} = .0491$, and $\sigma_5 = .0436$. We show the magnitude bode plot of error function $H(s) - \tilde{H}_r^{d_r}(s)$ along with the error plots for the optimal Hankel norm approximation and the approximation computed by balanced truncation in Figure 2.1.

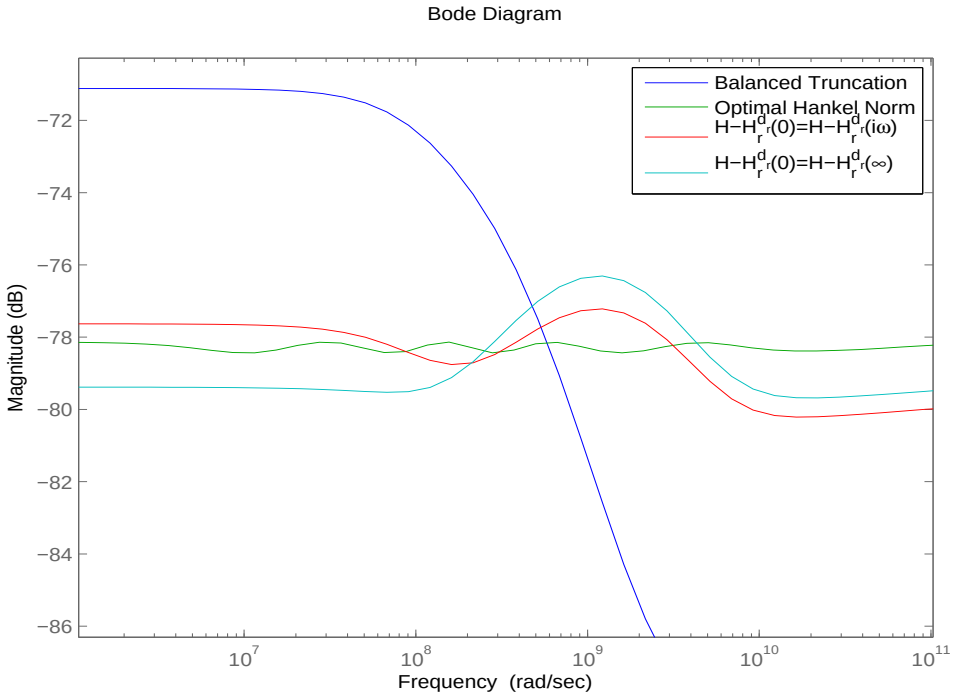
Example 2.2.2. Here the full order system to be approximated is the spiral inductor system PEEC model. This system has a minimal SSS realization of order 1434, making it moderately

Figure 2.1: Comparison of the Error Systems



large and therefore expensive to reduce. Here again, the order of approximation was $r = 4$. The reduced order model $\hat{H}_r^{d_r}(s)$ that forced $H(0) - \hat{H}_r^{d_r}(0) = H(\infty) - \hat{H}_r^{d_r}(\infty)$ did not satisfy the conditions of Theorem 2.2.9, so we instead searched for the system $\tilde{H}_r^{d_r}(s)$ that minimized the value of the error system at 0 and at some point in $(0, \infty)$. The $\mathcal{H}_\infty$-norm of the error was $\left\| H(s) - \hat{H}_r^{d_r}(s) \right\|_{\mathcal{H}_\infty} = 1.45 \times 10^{-}4$, and $\sigma_5 = 1.38 \times 10^{-4}$. The bode plot of the error systems for the approximations $\hat{H}_r^{d_r}$, and $\tilde{H}_r^{d_r}(s)$, along with the optimal Hankel norm approximation and the approximation computed by balanced truncation is shown in Figure 2.2. It is clear from the bode plot that the system $\hat{H}_r^{d_r}$ yields a smaller approximation error than the system $\tilde{H}_r^{d_r}(s)$, as is expected from theorem 2.2.10. Note that because the linear solves required in the method were expensive, only 30 points along the imaginary axis were sampled, and only 10 were used in the maximum of the other 30, yielding good results.

Figure 2.2: Comparison of the Error Systems For PEEC Model



The results presented in this section lead to a numerically efficient method for constructing reduced order models of SSS systems that are locally $\mathcal{H}_\infty$-optimal. Starting from the locally $\mathcal{H}_2$ -optimal approximation, $H_r(s)$, we found the nearby system $\tilde{H}_r^{d_r}(s)$ that minimized the $\mathcal{H}_\infty$-norm of the error over all stable reduced order models $H_r^{d_r}(s)$ satisfying the same interpolation conditions as $H_r(s)$. Surprisingly, this tweaking of the approximation $H_r(s)$ yields an approximation that in many cases is nearly optimal. The method's dependence on the dynamics of SISO SSS systems does not in principle lend itself to generalization, but it does offer a further advantage in the pursuit of general $\mathcal{H}_\infty$ optimal interpolation conditions by providing examples of interpolation zeros which yield nearly optimal results. Analyzing these interpolation points may offer insight into the general problem of $\mathcal{H}_\infty$ optimal interpolation points. To that end, we now turn to an empirical investigation of general patterns in the nearly optimal interpolation points.

Chapter 3

How Then Shall We Balance Our Charges?

3.1 Potential Theory and Rational Interpolation in the Complex Plane

The connections between rational interpolation and potential theory in the complex plane are well known and have been developed at length by Walsh, Bagby and others [25], [20]. Its relevance to the optimal $\mathcal{H}_\infty$ model reduction problem is easily motivated by considering the logarithm of the modulus of the error system $H(s)$-$H_r(s)$=$E(s)$:

$$\log|E(s)| = \log\left(\frac{\prod_{i=1}^{n+r}|(s-z_i)|}{\prod_{k=1}^{n+r}|(s-\lambda_k)|}\right) \tag{3.1}$$

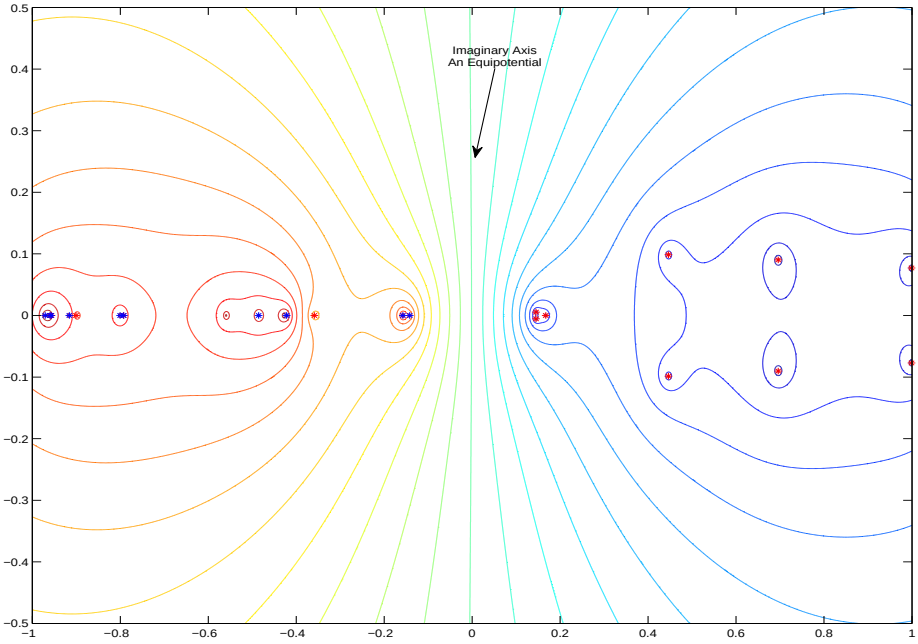
$$= \sum_{i=1}^{n+r}\log(|s-z_i|) - \sum_{k=1}^{n+r}(s-\lambda_k) \tag{3.2}$$

The singularities of the this sum occur at the zeros and poles of the $E(s)$, and 3.2 can be thought of as the sum of the potentials in the complex plane associated with point charges of amplitude -1 located at the zeros and of amplitude 1 located at the poles. Recall from Theorem 2.1.3 that for $H_r(s)$ to be optimal it is sufficient for $E(s)$ to be constant along the imaginary axis, and have at least 2r+1 zeros in the $\mathbb{C}_+$. This translates into the requirement that the sum (3.2) be constant along the imaginary axis, hence the imaginary axis is an equipotential of the the charge distribution. We can see that this occurs even for nearly optimal approximations. Figure 3.1 shows the equipotentials of the error system for an order 4 optimal Hankel norm approximation to an order 10 SSS system $H(s)$. The poles of $H(s)$ are randomly distributed through the interval (-1, 0), and the approximation is essentially optimal (numerically speaking).

In so far as the optimal $\mathcal{H}_\infty$-problem is difficult (and it is), the problem of inducing an equipotential along the imaginary axis given a fixed distribution of charges in the left half plane is equally difficult. Following a layman's derivation of this problem's connection to potential theory inspired by [8], we can simplify this problem somewhat by taking the limit as $n \to \infty$ of a rescaled problem:

$$f(z) = n^{-}1\sum_{i=1}^{n}\log|z-z_i| - n^{-1}\sum_{k=1}^{n}\log|z-\lambda_k| + C$$

Figure 3.1: Equipotentials of the Error System



In the limit as $n \to \infty$ we imagine a unit positive charge distributed in a continuous fashion through the left half plane and a unit negative charge distributed through the right half plane and and assume there is no conductivity over the imaginary axis. This situation is analogous in the physical setting to charging an electric capacitor. The aim is to minimize the max over the imaginary axis of $f(z)$. The negative charge settles to an equilibrium distribution over the set S_- where all the poles lie in the left half plane and the positive charge settles to an equilibrium over the set S_+ where all the zeros lie in the right half plane. The sets S_- and S_+ can be viewed as the positive and negatively charged plates of the capacitor. Formally, we assume that the set $S_c = S_- \cup S_+$ is a doubly connected set, and in the language of potential theory it is called a condenser [10]. In a process called *balayage* this distribution of charge is "swept" to the boundaries of the condenser so that the equipotentials of the charge distribution remain the same when viewed from the compliment of the S_c [8]. It has

been shown by [25], and [2] *et al* that it is possible to choose discrete sets of points along the boundary of S_c, such as the Bagby-Leja, or Walsh-Fekete points, which are asymptotically equivalent to the equilibrium charge distribution over the boundary. It follows that these points would be asymptotically minimal interpolation points in the $\mathcal{H}_\infty$-norm. For a given order of approximation r, we can therefore view the $2r+1$ interpolation points which are optimal, or nearly optimal, as the set of point charges which induce a nearly equivalent equipotential profile to the full distribution of charges. In the same spirit, a reasonable heuristic strategy would be to find the $r+1$ points in S_- which are "equivalent" to the total charge distribution in S_-, and then match both these $r+1$ points, and the r reduced order poles with interpolation zeros on the boundary of S_+. Thus we are effectively replacing the spectrum of $\boldsymbol{A}$ with $r+1$ equivalent charges and matching these equivalent charges in the right half plane. This strategy is plausible provided that the remaining interpolation zeros do not make any significant contribution to the total potential. In some sense then, the problem depends on interpolation zeros which we have no *a priori* knowledge of, but are ancillary to the enforcement of the $2r+1$ interpolation conditions we control. We will call these zeros the ancillary zeros of the error system. The question that naturally arises is where these ancillary zeros occur? Plausibly, if one could in some way anticipate these zeros of the error system in a manner relatively independent of the interpolation zeros, one would then be free to choose the appropriate interpolation points accordingly. With this possibility in mind, we start with the empirical observation that for a nearly optimal approximation, the $n-r-1$ ancillary zeros of the error system essentially cancel $n-r-1$ poles of the full order system, in a sense reducing the error system to an order $2r+1$ rational function. This observation is hinted at qualitatively Figures 3.2-3.5, which show the ancillary zeros and full system poles of several different benchmark models. The zeros in these figures plainly capture the spectrum of the full order system coefficient matrix.

Figure 3.2: Poles of 2-D beam model and ancillary zeros of order 20 optimal Hankel norm approximation

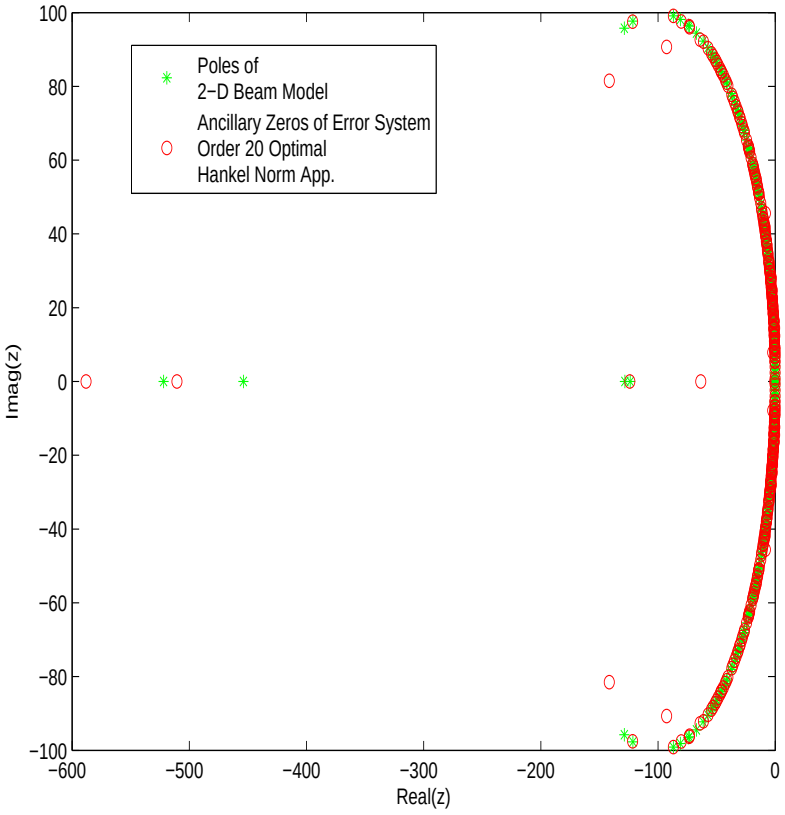


Figure 3.3: Poles of CD Player model and ancillary zeros of order 8 optimal $H_r^{d_r}(s)$ approximation

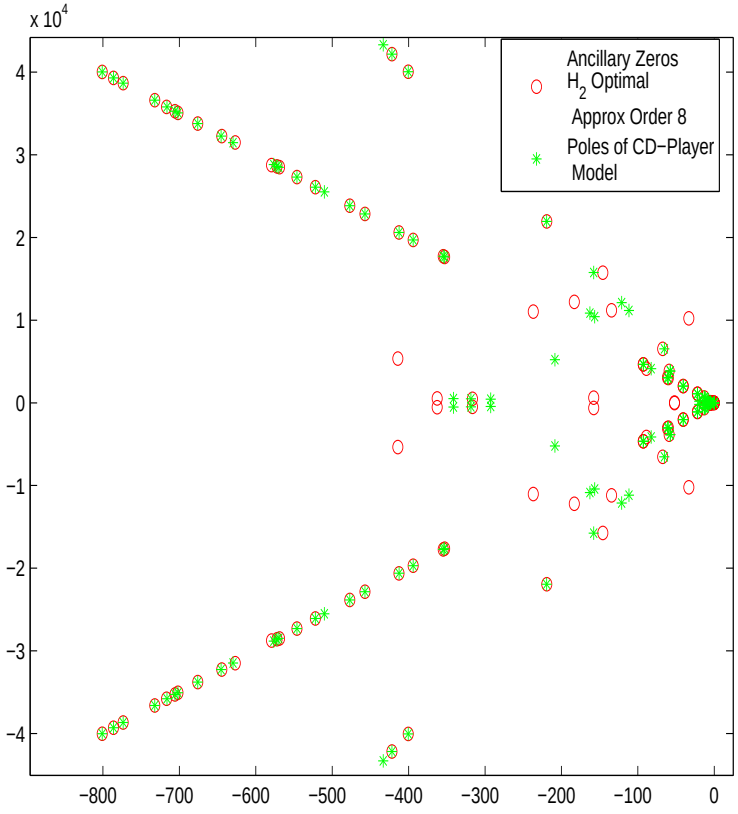
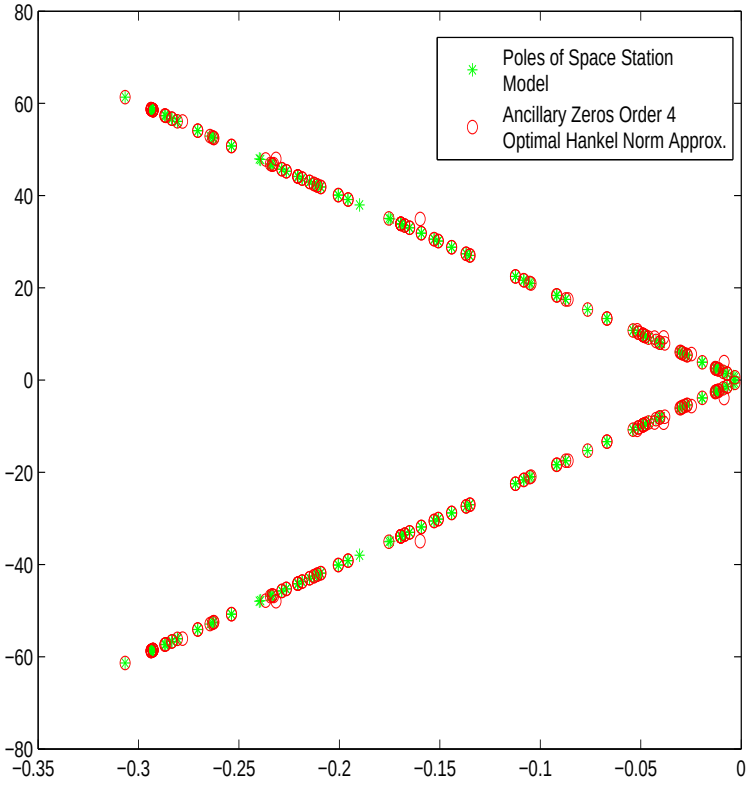


Figure 3.4: Poles of International Space-Station Model Order 270 and Ancillary Zeros of Order 4 Optimal Hankel Norm Approximation

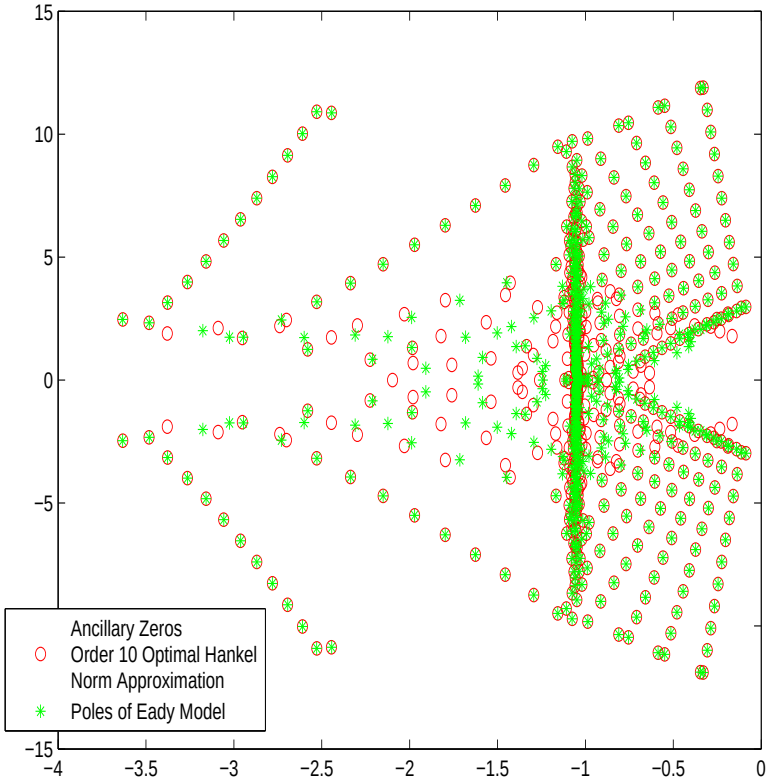


A further experiment that lends credibility to our observation involves measuring, for a given model, the distance from the ancillary zeros of a nearly optimal approximation to their nearest neighbors in the set of full-system poles.

Let $Z = \{z_i, i = 1, \ldots, n+r\}$ and $\Lambda = \{\lambda_i, i = 1, \ldots, n+r\}$. If $Z \subsetneq \mathbb{C}_-$, then set $Z = \tilde{Z} = (Z - \mathbb{C}_+) \cup -(Z - \mathbb{C}_-)$. Assume that each $z_i \in Z$ may be uniquely paired with its nearest neighbor $\lambda_k \in \Lambda$, and order the λ_k's so that the nearest neighbor of z_i is λ_i. Let

$$\phi_i = z_i - \lambda_i \tag{3.3}$$

Figure 3.5: Poles of Eady Model Order 548 and Ancillary Zeros of Order 10 Optimal Hankel Norm Approximation



Writing the expression for the error according to this ordering we obtain:

$$\log|E(i\omega)| = \log|\beta| + \log\left(\frac{\prod_{i=1}^{n+r}|(i\omega - z_i)|}{\prod_{k=1}^{n+r}|(i\omega - \lambda_k)|}\right) \tag{3.4}$$

$$= \sum_{i=1}^{n+r}\log|1 - \frac{\phi_i}{i\omega - \lambda_i}|, \tag{3.5}$$

where β is the value of the system at infinity. For any point $i\omega$, negative terms of the sum can only reduce the total error at the point, so to get a crude sense of how large the sum

might become, we need only look at the positive terms. Each term of the sum reaches a unique maximum at some point along the imaginary axis, roughly when $i\omega - \mathrm{Im}(\lambda_i) = 0$. Let $M_i = \log|1 - \frac{\phi_i}{-\mathrm{Re}(\lambda_i)}|$. Focusing on the contribution to the sum due to the ancillary zero-pole pairings, we might expect that the effect of near pole-zero cancellation on the total contribution would be on the order of

$$M = \sum_{k, M_{i_k} > 0} M_{i_k} \tag{3.6}$$

and that this contribution be substantially less than the value of $\log|\beta|$. Indeed, this is in general what we find when the approximation is nearly optimal. In the table below, we looked at the optimal Hankel norm and balanced truncation approximation to the ISS-1R model for orders 2 through 20, omitting the odd valued orders because of the slow decay of the Hankel singular values. For each order We compared the value of M for the optimal Hankel norm and balanced truncation approximations for each order. Note that the order 18 approximation seems to be an exception to the trend, however we found that in this case the ancillary zeros which appeared to be farthest from all full-system poles had a unique nearest neighbor in the set of reduced order poles. Adding these poles into the collection of error system poles and recomputing M yielded a result consistent with the other approximations. The value of M for the balanced truncation could not be similarly reduced in a significant way.

Table 3.1: Contribution of Ancillary Zeros to Error for International Space Station Model

Order	Lower Bound	$\log(\beta_{\mathrm{H}})$	$\log(M_{\mathrm{H}})$	$\log(M_{\mathrm{bal}})$
2	1.69×10^{-2}	-4.08	5.59×10^{-7}	26.54
4	5.31×10^{-3}	-5.28	1.76×10^{-6}	32.8
6	1.50×10^{-4}	-6.5	6.12×10^{-6}	32.8
8	5.93×10^{-4}	-7.43	1.20×10^{-5}	31.5
10	3.21×10^{-4}	-8.04	8.73×10^{-3}	23.3
12	2.52×10^{-4}	-8.419	3.98×10^{-2}	20.6
14	1.13×10^{-4}	-9.11	0.128	43.79
16	1.10×10^{-4}	-9.12	8.54×10^{-2}	35.3
18	1.08×10^{-4}	-9.36	11.37	37.7
20	1.00×10^{-4}	-9.206	0.14	38.8

Table 3.2: Comparison of error norms for approximations of table 3.1

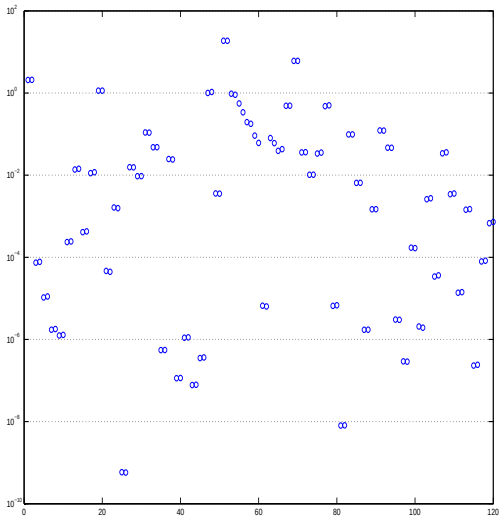
Order	Lower Bound	$\|H(s) - H_{\mathrm{H}}(s)\|_{\mathcal{H}_\infty}$	$\|H(s) - H_{\mathrm{bal}}(s)\|_{\mathcal{H}_\infty}$
2	1.69×10^{-2}	1.7×10^{-2}	3.37×10^{-2}
4	5.31×10^{-3}	5.32×10^{-3}	1.06×10^{-2}
6	1.50×10^{-4}	1.53×10^{-4}	3.00×10^{-3}
8	5.93×10^{-4}	6.12×10^{-4}	1.20×10^{-3}
10	3.21×10^{-4}	3.37×10^{-4}	6.43×10^{-4}
12	2.52×10^{-4}	2.73×10^{-4}	4.51×10^{-4}
14	1.13×10^{-4}	1.26×10^{-4}	2.27×10^{-4}
16	1.10×10^{-4}	1.61×10^{-4}	2.21×10^{-4}
18	1.08×10^{-4}	1.271×10^{-4}	2.17×10^{-4}
20	1.00×10^{-4}	1.04×10^{-4}	2.01×10^{-4}

3.2 Observations on Where the Ancillary Zeros of the Error System Do Not Occur

One further experimental observation will serve to motivate a heuristic method of approximation which we hope to develop further in the future. In general, we found that the poles of the full order system that were essentially unmatched by the ancillary zeros of an arbitrary stable order r approximation were, statistically speaking, invariant. Figure 3.2 illustrates this behavior by plotting the average distance from a full order pole to its nearest neighbor

in the set of ancillary zeros. The average was taken by fixing an order of approximation, and then randomly selecting interpolation points within a boundary that roughly approximated the reflection of the spectrum of $\boldsymbol{A}$ over the imaginary axis. From these interpolation points a reduced order model was constructed, and then its corresponding set of ancillary zeros was computed by simply removing the known interpolation points from the set of all the error system zeros. Finally, the distance from a pole of the full order system to its nearest neighbor in this set of ancillary zeros was computed. We kept track of this distance for each pole for several thousand randomly chosen sets of interpolation points and then averaged the distances.

Figure 3.6: Average Distance of CD-Player Poles to Nearest Ancillary Zeros of Order 6 Approximations



The significance of these unmatched poles derives from the fact that one can closely identify the neighborhood of good interpolation points by noting which poles, on the average, have the farthest nearest neighbor in a random set of ancillary zeros. To the extent that this observation holds, it suggests that a good starting point for choosing interpolation points which make the $\mathcal{H}_\infty$-norm of the error small are near the mirror image of the poles that

are difficult to match with the ancillary zeros of an arbitrary approximation. Of course, we do not want to have to compute several thousand random approximations just to get within a neighborhood of the "good" points, so it is noteworthy that we found that these unmatched poles often lie very close to the interpolation points computed by IRKA, which gives a plausible account of why the interpolation points found using IRKA often perform so well in the $\mathcal{H}_\infty$-norm for many of our examples. Table 3.2 demonstrates this pattern with four models having significantly different pole distributions. The averages for the unmatched poles were computed in the same manner as above and then compared with the fixed points of IRKA, initialized with 4 interpolation points.

Table 3.3: Correspondence of Unmatched Poles to IRKA Interpolation Points For Order 4 Approximations to Various Models

Model	Unmatched Poles	IRKA interpolation points
CD player	$-12.27 \pm 306.54i$, $-19.75 \pm 196.6i$	$12.3 \pm 306.6i$, $19.8 \pm 196.2i$
ISS-1R	$-0.0039 \pm 0.7751i$, $-0.0100 \pm 1.9920i$	$0.0039 \pm 0.7751i$, $0.0100 \pm 1.9920i$
2-D Beam	$-0.0051 \pm 0.1047i$,$-0.0066 \pm 0.5686i$	$0.0051 \pm .1047i$,$0.0066 \pm 0.5683i$
Random	$-0.01 \pm 789i$, $-353 \pm 3104i$	$0.01 \pm 789i$, $594 \pm 3090i$

A continued investigation of these types of patterns reveals many other interesting correlations between the IRKA points and interpolations points which effectively make the $\mathcal{H}_\infty$-norm small. For example, we also found that when the optimal model $H_r^{d_r}(s)$ was found, starting from the $\mathcal{H}_2$ optimal model, the poles of this reduced order model tended to lie very close to the poles of the optimal Hankel norm model, provided that the optimal Hankel norm model was an especially good approximation.

3.3 Interpolation Points from Penzl's Heuristic Method

If the unmatched poles of the full system are $\lambda_1, \lambda_2, \ldots, \lambda_{r+1}$, and the reduced system poles are $\tilde{\lambda_1}, \ldots, \tilde{\lambda_r}$ then, the the problem of choosing the optimal interpolation points amounts to choosing interpolation points $z_1^*, \ldots, z_{2r+1}^*$, such that

$$\{z_1^*, \ldots, z_{2r+1}^*\} = \arg \min_{z_1,\ldots,z_{2r+1} \in \mathbb{C}} \max_{z \in \mathbb{C}} \frac{\prod_{i=1}^{2r+1} |(z - z_i)|}{\prod_{i=1}^{r+1} |z - \lambda_i| \prod_{j=1}^{r} |z - \tilde{\lambda_i}|}. \tag{3.7}$$

This minimization and the previous observations we have made suggest the following rule of thumb, which we have found generally produces reduced-order models with a small $\mathcal{H}_\infty$-norm error.

RULE OF THUMB:

> Choosing approximately $r+1$ interpolation points at the mirror image of the most poorly matched poles of the full-order system, and r interpolation points at the mirror image of the reduced order poles yields good approximations in the $\mathcal{H}_\infty$-norm.

The minimization problem 3.7 is a very close relative to the the rational Zolatarev problem, for which all the zeros and poles are free in the minimization [10]. The rational Zolatarev problem has been the subject of much research, in part because the solution to this problem over the spectrum of a matrix provides a solution (in an asymptotic sense) to finding the optimal shifts in the ADI iteration [22]. Penzl developed an effective heuristic method for solving the optimal ADI parameters problem [19]. Instead of solving the rational Zolatarev problem over a condenser, he suggested minimizing (3.7) over a set of Ritz values, and harmonic Ritz values of $\boldsymbol{A}$. This method has a good theoretical basis, since the Ritz values of $\boldsymbol{A}$ have an interpretation in terms of potential theory as the set of charges that induce a nearly

equivalent potential to the potential of the original distribution of eigenvalues [15]. The points chosen by Penzl's algorithm are also the Bagby-Leja interpolation points computed over a discrete set [19]. We made use of Penzl's heuristic by solving the same minimization problem, but over a different set of discrete points. Instead of using the Ritz values of $\boldsymbol{A}$ and the harmonic Ritz values of $\boldsymbol{A}$, we used the IRKA points and the harmonic Ritz values of $\boldsymbol{A}$. We found by repeated experiment that the Ritz values of $\boldsymbol{A}$ tended to be poor interpolation points, so these were not included in the set over which we minimized. The idea was to use the Penzl algorithm to get $r+1$ interpolation points. The remaining r interpolation points would be forced at the mirror image of the reduced order poles, so they effectively made no contribution to the error. The d_r-term in this case was simply the chosen to enforce interpolation at the $r+1$st interpolation point (always real-valued) which we did not use to construct either $\boldsymbol{W}_r$ or $\boldsymbol{V}_r$. We will refer to this method as the Equivalent Charges method. We found that this approach yielded improvements over the optimal $\mathcal{H}_2$ approximation for smaller orders of approximation–between $r=2$ and $r=8$. The tables show this for three different systems. The number of harmonic Ritz values we computed was $r+1$, and when the order is small, these values coincided closely with the $\mathcal{H}_2$ -optimal points computed from IRKA. For higher orders of approximation, the similarities between the harmonic Ritz values and the $\mathcal{H}_2$ -optimal interpolation points diverged completely, which may explain why the Equivalent Charges method of choosing interpolation points performed poorly for higher orders of approximation on models with several resonances.

Table 3.4: Comparison of Method Performance For Heat Model of Order 197

Order	$\|E_{\text{bal}}\|_{\mathcal{H}_\infty}$	$\|E_{\text{Hank}}\|_{\mathcal{H}_\infty}$	$\|E_{\mathcal{H}_2}\|_{\mathcal{H}_\infty}$	$\left\|\tilde{E}_{d_r}\right\|_{\mathcal{H}_\infty}$	$\|E_{\text{Eq. Charges}}\|_{\mathcal{H}_\infty}$
2	0.462	0.267	0.462	0.257	0.280
4	4.01×10^{-2}	2.02^{-2}	3.94×10^{-2}	2.10×10^{-2}	2.98×10^{-2}
6	1.10×10^{-3}	4.95×10^{-4}	1.30×10^{-3}	5.50×10^{-4}	1.00×10^{-3}
8	1.80×10^{-4}	1.32×10^{-4}	2.12×10^{-4}	1.67×10^{-4}	1.70×10^{-4}

Figure 3.7: Magnitude Bode Plot of FOM Model and Equivalent Charges Approximation of Order 12

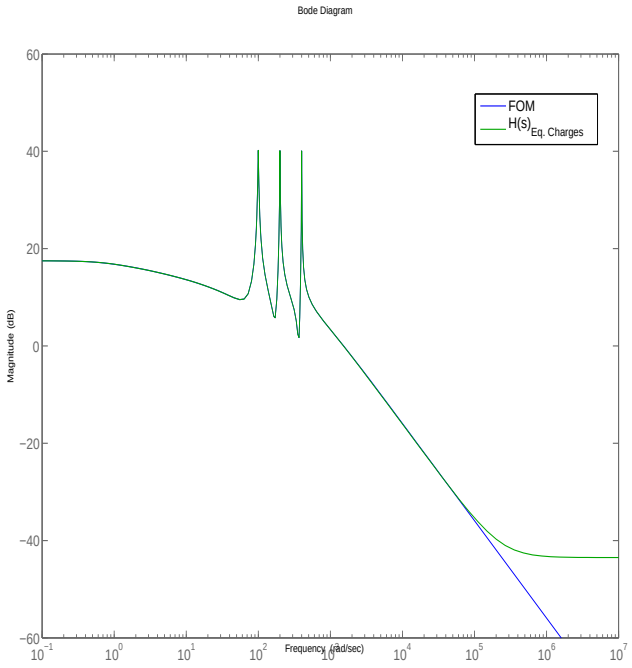


Table 3.5: Comparison of Method Performance For Eady Model Order 598

Order	$\|E_{\text{bal}}\|_{\mathcal{H}_\infty}$	$\|E_{\text{Hank}}\|_{\mathcal{H}_\infty}$	$\|E_{\mathcal{H}_2}\|_{\mathcal{H}_\infty}$	$\left\|\tilde{E}_{d_r}\right\|_{\mathcal{H}_\infty}$	$\|E_{\text{Eq. Charges}}\|_{\mathcal{H}_\infty}$
2	197	136	261	259	220
4	71.7	43.5	77.7	77.5	77.1
6	18.8	8.54	20.4	20.28	19.542
8	2.585	1.32	2.365	2.29	2.48

Table 3.6: Comparison of Method Performance For FOM Model Order 1006

Order	$\|E_{\text{bal}}\|_{\mathcal{H}_\infty}$	$\|E_{\mathcal{H}_2}\|_{\mathcal{H}_\infty}$	$\|E_{\text{match}}\|_{\mathcal{H}_\infty}$
6	7.29	7.28	4.25
8	1.00	2.16	0.395
10	0.100	0.264	0.122
12	9.00×10^{-3}	2.16×10^{-2}	9.20×10^{-3}
14	7.37×10^{-4}	1.65×10^{-3}	4.64×10^{-4}
16	5.58×10^{-5}	1.19×10^{-4}	5.23×10^{-5}

These results suggest that searching for the $r+1$ points which are good equivalent charges to the full spectrum is a promising route for an interpolation-based approach to the optimal $\mathcal{H}_\infty$ model reduction problem.

Chapter 4

Conclusions and Future Work

The $\mathcal{H}_\infty$ model reduction problem has been around for some time, and has even obtained the prominence of an "unsolved problem" [6], due to its difficulty and evident importance in applications. A great deal of work has been done on this problem from an SVD-based approached to model reduction, but even a lengthy search of the literature will reveal that little has been done on the $\mathcal{H}_\infty$model reduction problem from an interpolation-based approach. The results we presented here show that an interpolation-based approach to this problem bears much fruit. We showed it was possible to parameterize all interpolants of a given order by the D-term of the system realization, and that this parameterization can be exploited to significantly improve the $\mathcal{H}_\infty$-norm of the $\mathcal{H}_2$ -optimal approximation, as in the case of the CD-player model, or for state-space symmetric systems generally. For the special case of state-space symmetric systems, theorems 2.2.9 and 2.2.10 provide sufficient conditions for an inexpensive means of minimizing the $\mathcal{H}_\infty$ approximation error when starting from the $\mathcal{H}_2$ -optimal approximation computed from the IRKA iteration. Our results showed that this method consistently yielded approximations which well outperformed the balanced truncation method. Finally, we presented several numerical experiments which suggest that finding

the optimal interpolation points will ultimately depend on concretizing the close connections between rational interpolation and potential theory. One possible way to solidify these connections would involve computing the Bagby-Leja and Walsh-Fekete points for condensers which approximate the equilibrium charge distribution over the numerical range of the A-matrix and testing the ability of these points to yield interpolants with small approximation error. Effective methods have already been developed for computing the Walsh-Fekete and Bagby-Leja points over polygonal domains [22]. Another possible direction for further research would be to develop a characterization of the phenomena of the unmatched full-order poles. In particular, understanding where they occur and why they occur. Finally, we hope to investigate how the parameterization of interpolants by the system's D-term might be similarly exploited in the multi-input multi-ouput setting.

Bibliography

[1] A.C. Antoulas. *Approximation of large-scale dynamical systems*. Society for Industrial Mathematics, 2005.

[2] T. Bagby. On interpolation by rational functions. *Duke Math. J*, 36(1):95–104, 1969.

[3] CA Beattie and S. Gugercin. Krylov-based model reduction of second-order systems with proportional damping. In *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05*, pages 2278–2283, 2005.

[4] CA Beattie and S. Gugercin. Structured Perturbations in Rational Krylov Methods for Model Reduction. 2007.

[5] Christopher Beattie and Serkan Gugercin. Interpolatory projection methods for structure-preserving model reduction. *Systems & Control Letters*, 58(3):225 – 232, 2009.

[6] V. Blondel and A. Megretski. *Unsolved problems in mathematical systems and control theory*. Princeton University Press, 2004.

[7] A. Böttcher and B. Silbermann. *Introduction to large truncated Toeplitz matrices*. Springer Verlag, 1999.

[8] T.A. Driscoll, K.C. Toh, and L.N. Trefethen. From potential theory to matrix iterations in six steps. *SIAM Review*, pages 547–578, 1998.

[9] P. Feldman and RW Freund. Efficient linear circuit analysis by Pade approximation via the Lanczos process, IEEE Trans. *Computer-Aided Design*, 14(5):l.

[10] AA Gonchar. Zolotarev problems connected with rational functions, Math. *USSR-Sb*, 7:623–635, 1969.

[11] E.J. Grimme. *Krylov projection methods for model reduction.* PhD thesis, University of Illinois, 1997.

[12] S. Gugercin, A. C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2):609–638, 2008.

[13] D. Kavranoğlu and M. Bettayeb. Characterization of the solution to the optimal $\mathcal{H}_\infty$ model reduction problem. *Systems & Control Letters*, 20(2):99–107, 1993.

[14] G. Keith. All optimal Hankel-norm approximations of linear multivariable systems and their $\mathcal{L}_\infty$-error bounds. *Int J Control*, 39(6):1115–1193, 1984.

[15] A.B.J. Kuijlaars. Convergence analysis of Krylov subspace iterations with methods from potential theory. *SIAM REVIEW*, 48(1):3, 2006.

[16] WQ Liu, V. Sreeram, and KL Teo. Model Reduction for State-space Symmetric Systems. *Systems & Control Letters*, 34:209–215, 1998.

[17] A.J. Mayo and A.C. Antoulas. A framework for the solution of the generalized realization problem. *Linear Algebra and its Applications*, 425(2-3):634 – 662, 2007. Special Issue in honor of Paul Fuhrmann.

[18] K.S. Miller. *Some eclectic matrix theory.* Malabar, Florida: Robert E, 1987.

[19] T. Penzl and P. Sfb. A cyclic low rank Smith method for large sparse Lyapunov equations. *Siam J. Sci. Comput*, 1998.

[20] E.B. Saff and V. Totik. *Logarithmic potentials with external fields*. Springer, 1997.

[21] B. Srinivasan and P. Myszkorowski. Model reduction of systems with zeros interlacing the poles. *Systems & Control Letters*, 30(1):19–24, 1997.

[22] G. Starke. Fejer-Walsh points for rational functions and their use in the ADI iterative method. *Journal of Computational and Applied Mathematics*, 46(1-2):129–141, 1993.

[23] L.N. Trefethen. Near-circularity of the error curve in complex Chebyshev approximation. *J. Approx. Theory*, 31(4):344–367, 1981.

[24] CD Villemagne and RE Skelton. Model reduction using a projection formulation, Intl. *J. Contr*, 46:2141–2169, 1987.

[25] J.L. Walsh. *Interpolation and approximation by rational functions in the complex domain*. American mathematical society, 1969.

[26] A. Yousuff and RE Skelton. Covariance Equivalent Realizations with Application to Model Reduction of Large Scale Systems. *Control and Dynamic Systems*, 22, 1984.

[27] A. Yousuff, DA Wagie, and RE Skelton. Linear system approximation via covariance equivalent realizations. *Journal of mathematical analysis and applications*, 106(1):91–115, 1985.

[28] K. Zhou, J.C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall Upper Saddle River, NJ, 1996.