Inexact Solves in Interpolatory Model Reduction

Sarah Wyatt

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

Master of Science
in
Mathematics

Serkan Gugercin, Chair
Christopher Beattie
Eric de Sturler

May 4, 2009
Blacksburg, Virginia

Keywords: $\mathcal{H}_2$ Approximation, Rational Krylov, interpolation, model reduction

Inexact Solves in Interpolatory Model Reduction

Sarah Wyatt

Abstract

Dynamical systems are mathematical models characterized by a set of differential or difference equations. Due to the increasing demand for more accuracy, the number of equations involved may reach the order of thousands and even millions. With so many equations, it often becomes computationally cumbersome to work with these large-scale dynamical systems. Model reduction aims to replace the original system with a reduced system of significantly smaller dimension which will still describe the important dynamics of the large-scale model. Interpolation is one method used to obtain the reduced order model. This requires that the reduced order model interpolates the full order model at selected interpolation points. Reduced order models are obtained through the Krylov reduction process, which involves solving a sequence of linear systems. The Iterative Rational Krylov Algorithm (IRKA) iterates this Krylov reduction process to obtain an optimal $\mathcal{H}_2$ reduced model. Especially in the large-scale setting, these linear systems often require employing inexact solves. The aim of this thesis is to investigate the impact of inexact solves on interpolatory model reduction.

We considered preconditioning the linear systems, varying the stopping tolerances, employing GMRES and BiCG as the inexact solvers, and using different initial shift selections. For just one step of Krylov reduction, we verified theoretical properties of the interpolation error. Also, we found a linear improvement in the subspace angles between the inexact and exact subspaces provided that a good shift selection was used. For a poor shift selection, these angles often remained of the same order regardless of how accurately the linear systems were solved. These patterns were reflected in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ errors between the inexact and exact subspaces, since these errors improved linearly with a good shift selection and were

typically of the same order with a poor shift. We found that the shift selection also influenced the overall model reduction error between the full model and inexact model as these error norms were often several orders larger when a poor shift selection was used. For a given shift selection, the overall model reduction error typically remained of the same order for tolerances smaller than 1×10^{-3}, which suggests that larger tolerances for the inexact solver may be used without necessarily augmenting the model reduction error. With preconditioned linear systems as well as BiCG, we found smaller errors between the inexact and exact models while the order of the overall model reduction error remained the same. With IRKA, we observed similar patterns as with just one step of Krylov reduction. However, we also found additional benefits associated with using an initial guess in the inexact solve and by varying the tolerance of the inexact solve.

Acknowledgments

I would like to express my gratitude to Dr. Gugercin, who has been an absolutely phenomenal advisor. Beginning with my first question, Dr. Gugercin has exemplified the qualities of patience, intellectual generosity and superb communication of his vast knowledge. I would especially like to thank him for his time, understanding and care during every step of this process. Also, I would like to thank Dr. de Sturler, Dr. Beattie, Dr. Zietsman, Dr. Borggaard, and Dr. Drmac who have shown me the beauty of numerical analysis. In addition, I would like to thank my friends and family. Many thanks to Matthew for his kindness and for enriching my appreciation of the applications of dynamical systems. Thank you to Shannon, Lynn, Courtney, Amy, John, Lacey, Garrett and Kasie for their encouragement and support. A very sincere thanks to Nik, who, among so many other gifts, gave me more RAM, taught me the efficient way, challenged the veracity of my assumptions, and opened my myopic eyes. Also, a very heartfelt thank you to my parents, Leslie and Will for loving and supporting me in every way possible. I would especially like to thank Nik and my family for their compassion during and after the fall semester of 2008 and for imbuing my life with love and happiness.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Linear Dynamical Systems

Dynamical systems are mathematical models characterized by a set of differential or difference equations which capture behavior of natural and artificial processes. Today's problems often lead to an almost insatiable demand for more precision, requiring a myriad of equations to describe the system. Oftentimes, these dynamical systems may involve thousands and even millions of equations. Although these complex models may capture the overall dynamics of the system more accurately, limited computational resources, inaccuracy, and ill-conditioning often result in these large models being computationally cumbersome or even intractable to use in a practical setting. Therefore, the original system is replaced with a reduced system described by a smaller set of equations. The aim of model reduction is to obtain this reduced system, which will still delineate the important intricacies of the original system and yet be feasible to use in practice.

1.2 Notation

Table 1.1: Notation

$\mathbb{R}^{m\times n}$	set of real matrices of size m by n
$\mathbb{C}^{m\times n}$	set of complex matrices of size m by n
s	a complex number in $\mathbb{C}$
$\bar{s}$	complex conjugate of s
$\mathbf{A}^*$	complex conjugate transpose of $\mathbf{A}$ regardless of whether $\mathbf{A}$ is real or complex
$\text{span}\{\mathbf{x}_1, ..., \mathbf{x}_l\}$	span of the vectors $\mathbf{x}_1, ..., \mathbf{x}_l$
$\|\mathbf{x}\|_p$	p vector norm of $\mathbf{x}$
$\mathbf{I}$	Identity matrix of appropriate size
$\imath$	$\sqrt{-1}$
$H(s)$	full order model and transfer function
$H_r(s)$	reduced order model and transfer function using exact solves
$\widetilde{H}_r(s)$	reduced order model and transfer function using inexact solves
$\|H\|_{\mathcal{H}_\infty}$	$\mathcal{H}_\infty$ norm of $H(s)$
$\|H\|_{\mathcal{H}_2}$	$\mathcal{H}_2$ norm of $H(s)$
$\text{Ran}(\mathbf{M})$	Range of $\mathbf{M}$
$\langle \mathbf{x}, \mathbf{y} \rangle$	inner product of vectors $\mathbf{x}$ and $\mathbf{y}$
ε	inexact solver's stopping tolerance (i.e. $\frac{\|\mathbf{Ax}-\mathbf{b}\|}{\|\mathbf{b}\|} < \varepsilon$)
$\text{logspace}(a, b, n)$	n points between decades 10^a and 10^b
prec.	Preconditioned linear systems
unprec.	Unpreconditioned linear systems
$\lambda = 10$	10% perturbation from optimal shifts

1.3 Model Reduction

In this thesis, we study model reduction for single input/singe output (SISO) continuous time, time-invariant, linear dynamical systems with finite dimension. These are represented in state space as

$$H(s) \quad : \quad \begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}\, u(t) \\ y(t) &= \mathbf{c}^T \mathbf{x}(t), \end{cases} \tag{1.3.1}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$. $\mathbf{x}(t) \in \mathbb{R}^n$ is the *state*, $u(t) \in \mathbb{R}$ is the *input*, and $y(t) \in \mathbb{R}$ is the *output* of $H(s)$. Also, we assume the system is stable, namely the real parts of the eigenvalues of $\mathbf{A}$ all lie in the left half plane. Another useful description of the dynamical system is given by the transfer function. To understand this concept, we first define the Laplace transform of a function.

Definition 1.1. *Let f(t) be a piecewise continuous and bounded function defined for $t \geq 0$. Then the unilateral Laplace transform of f(t), denoted by F(s), exists for all s in the open right half plane and is defined by $F(s) := \mathcal{L}(f(t)) = \int_0^\infty f(t)e^{-st}dt$.*

From this definition, one can derive that $\mathcal{L}(\dot{f}(t)) = s \cdot \mathcal{L}(f(t)) - f(0)$. Using this property, we apply the Laplace transform to (1.3.1) with $\mathbf{x}(0) = 0$ and obtain

$$s \cdot \mathbf{X}(s) = \mathbf{A}\mathbf{X}(s) + \mathbf{b}U(s) \qquad \text{and} \qquad Y(s) = \mathbf{c}^T\mathbf{X}(s) \tag{1.3.2}$$

where $\mathbf{X}(s) = \mathcal{L}(\mathbf{x}(t))$, $\mathbf{U}(s) = \mathcal{L}(\mathbf{u}(t))$, $\mathbf{Y}(s) = \mathcal{L}(\mathbf{y}(t))$ [18]. From (1.3.2), we then have $Y(s) = H(s)U(s)$, and the transfer function of the system is defined as

$$H(s) = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}. \tag{1.3.3}$$

Because $H(s)$ describes a single input/single output system, $H(s)$ is a scalar rational function. We will follow the common convention of denoting both the underlying system and its transfer function by $H(s)$.

The dimension of the associated state space is defined to be equal to that of the dimension of $\mathbf{A}$. Since $\mathbf{A}$ is an $n \times n$ matrix, the dimension of H is n. Through model reduction, we obtain a dynamical system, denoted by H_r, defined as follows:

$$\mathbf{H}_r(s) \quad : \quad \begin{cases} \dot{\mathbf{x}}_r(t) = \mathbf{A}_r \mathbf{x}_r(t) + \mathbf{b}_r u(t) \\ y_r(t) = \mathbf{c}_r^T \mathbf{x}_r(t), \end{cases} \tag{1.3.4}$$

where $\mathbf{A}_r \in \mathbb{R}^{r \times r}$, $\mathbf{b}_r \in \mathbb{R}^r$, and $\mathbf{c}_r \in \mathbb{R}^r$ with $r \ll n$. The associated transfer function of the reduced order model is then

$$H_r(s) = \mathbf{c}_r^T (s\mathbf{I} - \mathbf{A_r})^{-1} \mathbf{b_r}. \tag{1.3.5}$$

For $H_r(s)$ to be of practical use in simulation and control problems, there are several properties we want our reduced order model to possess. We want to obtain H_r in a computationally feasible and efficient manner even for large-scale dynamical systems. Furthermore, we would like the model reduction process to preserve stability. Finally and perhaps most importantly, H_r needs to capture the input and output relationship of the original system, i.e. $y(t) \approx y_r(t)$. For more details, see [1].

1.4 Model Reduction by Projection and Interpolation

In this section, we describe a projection process that leads to a reduced order system $H_r(s)$. Assume $\mathcal{V}_r$ and $\mathcal{W}_r$ are r-dimensional subspaces of $\mathbb{R}^n$ with $\mathcal{V}_r \cap \mathcal{W}_r^{\perp} = \{0\}$. For any input $u(t)$, if $\mathbf{v}(t) \in \mathcal{V}_r$ solves $\dot{\mathbf{v}}(t) - \mathbf{A}\mathbf{v}(t) - \mathbf{b}u(t) \perp \mathcal{W}_r$ for all t, then define

$$y_r(t) := \mathbf{c}^T \mathbf{v}(t). \tag{1.4.1}$$

Assume $\mathbf{V}_r, \mathbf{W}_r \in \mathbb{R}^{n\times r}$ are matrices such that $\mathcal{V}_r = \text{Range}(\mathbf{V}_r)$ and $\mathcal{W}_r = \text{Range}(\mathbf{W}_r)$. By the assumption that $\mathcal{V}_r \cap \mathcal{W}_r^{\perp} = \{0\}$, $\mathbf{W}_r^*\mathbf{V}_r$ is invertible. From (1.4.1), we can deduce $\mathbf{v}(t) = \mathbf{V}_r\mathbf{x}_r(t)$ with $\mathbf{x}_r(t) \in \mathbb{R}^r$ for all t, implying $\dot{\mathbf{v}}(t) = \mathbf{V}_r\dot{\mathbf{x}}_r(t)$. Since $\mathcal{W}_r = \text{Range}(\mathbf{W}_r)$ and $\dot{\mathbf{v}}(t) - \mathbf{A}\mathbf{v}(t) - \mathbf{b}u(t) \perp \mathcal{W}_r$ for all t, we have

$$\mathbf{W}_r^*(\dot{\mathbf{v}}(t) - \mathbf{A}\mathbf{v}(t) - \mathbf{b}u(t)) = 0.$$

Equivalently, this gives

$$\mathbf{W}_r^*(\mathbf{V}_r\dot{\mathbf{x}}_r(t) - \mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) - \mathbf{b}u(t)) = \mathbf{W}_r^*\mathbf{V}_r\dot{\mathbf{x}}_r(t) - \mathbf{W}_r^*\mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) - \mathbf{W}_r^*\mathbf{b}u(t) = 0.$$

Solving for $\dot{\mathbf{x}}_r(t)$, we have

$$\dot{\mathbf{x}}_r(t) = (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) + (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{b}u(t)$$

which leads us to define

$$\mathbf{A}_r = (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{A}\mathbf{V}_r$$

and

$$\mathbf{b}_r = (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{b}.$$

Finally, from $y_r(t) = \mathbf{c}^T\mathbf{v}(t) = \mathbf{c}^T\mathbf{V}_r\mathbf{x}_r(t)$, we obtain

$$\mathbf{c}_r^T = \mathbf{c}^T\mathbf{V}_r.$$

In summary, we have constructed a projector, $\Gamma_r = \mathbf{V}_r\mathbf{Z}_r^*$ where $\mathbf{V}_r$, $\mathbf{Z}_r \in \mathbb{R}^{n\times r}$ with

$\mathbf{Z}_r^*\mathbf{V}_r = \mathbf{I}_r$ such that

$$\dot{\mathbf{x}}_r(t) = \mathbf{Z}_r^*\mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) + \mathbf{Z}_r^*\mathbf{b}u(t) \qquad \text{and} \qquad y_r(t) = \mathbf{c}\mathbf{V}_r\mathbf{x}_r(t)$$

where $\mathbf{A}_r = \mathbf{Z}_r^*\mathbf{A}\mathbf{V}_r, \mathbf{b}_r = \mathbf{Z}_r^*\mathbf{b}$, and $\mathbf{c}_r = \mathbf{c}\mathbf{V}_r$. Therefore, the underlying oblique projector is

$$\Gamma_r = \mathbf{V}_r(\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*.$$

In this thesis, we will construct reduced models through interpolation (also known as moment matching). The k^{th} moment at a point $\sigma_i \in \mathbb{C}^r$ is defined as the k^{th} derivative of the transfer function evaluated as σ_i. The aim of model reduction by moment matching is for the reduced model, $H_r(s)$, to interpolate $H(s)$ at selected interpolation points or shifts, which will be denoted by σ_k. Oftentimes, we wish for interpolation of the first moment as well. Therefore, we want to find $H_r(s)$ such that

$$H_r(\sigma_k) = H(\sigma_k) \qquad \text{and} \qquad H_r'(\sigma_k) = H'(\sigma_k) \qquad \text{for} \qquad \text{k} = 1, ..., \text{r}. \tag{1.4.2}$$

Naturally, we desire for the process with which we achieve the interpolation to be numerically robust. Previous research has proven that the computation of moments is extremely ill-conditioned; as a result, we wish for the interpolation to be achieved without explicitly computing the moments [10].

To connect the moment matching and projection frameworks of model reduction, we present the following results without proof. The lemma provides insight for selecting $\mathcal{V}_r$ and $\mathcal{W}_r$ while its corollary links the projection and interpolatory frameworks of model reduction. These results are important since they elucidate a way to match the moments without ex-

plicitly computing them. The proofs may be found in [15], [12], [9], [25], and [26].

Lemma 1.2. *[15] Suppose $\sigma \in \mathbb{C}$ is not an eigenvalue of $\mathbf{A}$ nor of $\mathbf{A}_r$.*

$$\textit{If } (\sigma\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \in \mathcal{V}_r, \textit{ then } H_r(\sigma) = H(\sigma).$$
$$\textit{If } (\sigma\mathbf{I} - \mathbf{A})^{-*}\mathbf{c} \in \mathcal{W}_r, \textit{ then } H_r(\sigma) = H(\sigma).$$
$$\textit{If both } (\sigma\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \in \mathcal{V}_r \textit{ and } (\sigma\mathbf{I} - \mathbf{A})^{-*}\mathbf{c} \in \mathcal{W}_r, \textit{ then } H_r(\sigma) = H(\sigma) \textit{ and } H_r'(\sigma) = H'(\sigma).$$

Corollary 1.3. *[15] Consider the system $H(s)$ defined by $\mathbf{A}, \mathbf{b}, \mathbf{c}$, a set of distinct shifts given by $\{\sigma_1, ..., \sigma_r\}$ that is closed under conjugation and subspaces spanned by the columns of $\mathbf{V}_r$ and $\mathbf{W}_r$ with*

$$Ran(\mathbf{V}_r) = span\{(\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, ..., (\sigma_r\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\} \textit{ and}$$
$$Ran(\mathbf{W}_r) = span\{(\sigma_1\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}, ..., (\sigma_r\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}\}.$$

Then $\mathbf{V}_r$ and $\mathbf{W}_r$ can be chosen to be real matrices and the reduced order system H_r defined by $\mathbf{A}_r = (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{A}\mathbf{V}_r$, $\mathbf{b}_r = (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{b}$, and $\mathbf{c}_r^T = \mathbf{c}^T\mathbf{V}_r$ is itself real and matches the first two moments of $H(s)$ at each of the interpolation points σ_k, i.e. $H(\sigma_k) = H_r(\sigma_k)$ and $H'(\sigma_k) = H_r'(\sigma_k)$ for $k = 1, ..., r$.

From these theorems, we note that interpolation-based model reduction requires the use of $(\sigma_i\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$ and $(\sigma_i\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}$. In practice, we will never explicitly compute the inverse of $(\sigma_i\mathbf{I} - \mathbf{A})$ and $(\sigma_i\mathbf{I} - \mathbf{A})^*$. Instead, we will solve the systems $(\sigma_i\mathbf{I} - \mathbf{A})\mathbf{x}_i = \mathbf{b}$ and $(\sigma_i\mathbf{I} - \mathbf{A})^*\mathbf{y}_i = \mathbf{c}$. Especially if the dimension of the dynamical system is large, inexact solves will become necessary. In chapters 3, 4, and 5, we will be investigating the impact of these inexact solves.

1.5 $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms

Throughout this thesis, we will use the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms. Define $\mathcal{L}(i\mathbb{R})$ to be the Hilbert space of matrix valued functions on $i\mathbb{R}$ such that

$$\|F\|_2^2 := \frac{1}{2\pi}\int_{-\infty}^{\infty} trace(F^*(\imath\omega)F(\imath\omega))d\omega < \infty. \tag{1.5.1}$$

The underlying inner product is defined as

$$\langle F, G\rangle := \frac{1}{2\pi}\int_{-\infty}^{\infty} trace(F^*(\imath\omega)G(\imath\omega))d\omega. \tag{1.5.2}$$

The $\mathcal{H}_2$ space is then defined to be the closed subset of $\mathcal{L}(i\mathbb{R})$ that is composed of the functions that are analytic in $Re(s) > 0$. Clearly, transfer functions associated with real stable SISO dynamical systems will be elements of this $\mathcal{H}_2$ space. If $H(s)$ belongs to the $\mathcal{H}_2$ space, then the norm is defined as

$$\|H\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} |H(\imath\omega)|^2 d\omega\right)^{1/2}. \tag{1.5.3}$$

For more details, see [1].

The $\mathcal{H}_2$ norm is one way to evaluate the performance of the reduced order model. In order for $H_r(s)$ to be of practical use, we desire $H_r(s)$ to capture the relationship between the input and output of the system, namely we want $max_{t>0}\ |y(t) - y_r(t)|$ to be uniformly small

over all inputs $u(t)$. As shown in [15], assuming $u(t)$ is such that $\int_0^\infty |u(t)|^2 dt \leq 1$, then

$$max_{t>0}|y(t) - y_r(t)| \leq \|H - H_r\|_{\mathcal{H}_2}. \tag{1.5.4}$$

Therefore, to best achieve $y(t) \approx y_r(t)$, we want $\|H - H_r\|_{\mathcal{H}_2}$ to be minimized.

Another important performance measure is the $\mathcal{H}_\infty$ norm. Define $\mathcal{L}_\infty(i\mathbb{R})$ to be the Banach space of matrix valued functions $\mathbf{F}(s)$ such that

$$\|\mathbf{F}\|_\infty := \sup_\omega \|\mathbf{F}(\imath\omega)\|_2 < \infty.$$

Then define $\mathcal{H}_\infty(i\mathbb{R})$ to be the subset of $\mathcal{L}_\infty(i\mathbb{R})$ with elements analytic in $Re(s) > 0$ such that

$$\sup_{s_R \geq 0} \sup_\omega \|\mathbf{F}(s_R + \imath\omega)\|_2 < \infty.$$

The $\mathcal{H}_\infty$ norm is then defined as

$$\|\mathbf{F}\|_{\mathcal{H}_\infty} := \sup_\omega \|\mathbf{F}(\imath\omega)\|_2.$$

One advantage of the $\mathcal{H}_\infty$ norm is its ability to capture the physical properties of the system. If we assume that $\|H - H_r\|_{\mathcal{H}_\infty} \leq \alpha$ where α is a positive scalar, then

$$\|H - H_r\|_{\mathcal{H}_\infty} = \sup_{u \neq 0} \frac{\|y - y_r\|_{\mathcal{L}_2}}{\|u\|_{\mathcal{L}_2}} \leq \alpha.$$

In this way, the $\mathcal{H}_\infty$ norm describes the input and output relationship of the error system.

1.6 Interpolation using $\mathcal{H}_2$ optimality conditions

Given a system $H(s)$, the aim of optimal $\mathcal{H}_2$ model reduction is to find a reduced order system $H_r(s)$ such that

$$H_r(s) = \arg \min_{deg(\widehat{H}_r)=r} \|H(s) - \widehat{H}_r(s)\|_{\mathcal{H}_2}. \tag{1.6.1}$$

The significance of constructing a reduced order model H_r which satisfies (1.6.1) follows from (1.5.4); by finding a reduced order system that minimizes the $\mathcal{H}_2$ error, we will ensure that the maximium difference between the outputs $y(t)$ and $y_r(t)$ is as small as possible. This is an extremely important feature for our reduced order model to possess since we desire the reduced order model to capture the relationship between the input and output of the original model. To obtain the model that satisfies (1.6.1), we have the following two results, which we present without proof. The proofs may be found in [15] and [19].

Theorem 1.4. *[15] Given the full order model $H(s)$ and a reduced order model $H_r(s)$, let λ_i and $\hat{\lambda}_j$ be the poles of $H(s)$ and $H_r(s)$, respectively. Suppose that the poles of $H_r(s)$ are distinct. Let Φ_i and $\hat{\Phi}_j$ denote the residues of the transfer functions $H(s)$ and $H_r(s)$ at their poles λ_i and $\hat{\lambda}_j$, respectively: $\Phi_i = res[H(s), \lambda_i]$, $i = 1, ..., n$ and $\hat{\Phi}_j = res[H_r(s), \hat{\lambda}_j]$, $j = 1, ..., r$. The $\mathcal{H}_2$ norm of the error system is given by*

$$\|H(s) - H_r(s)\|^2_{\mathcal{H}_2} = \sum_{i=1}^{n} \Phi_i(H(-\lambda_i) - H_r(-\lambda_i)) + \sum_{j=1}^{r} \hat{\Phi}_i(H(-\hat{\lambda}_j) - H_r(-\hat{\lambda}_j)).$$

Theorem 1.4 describes the relationship between the $\mathcal{H}_2$ error and the poles of both $H(s)$ and $H_r(s)$. To minimize $\|H(s) - H_r(s)\|_{\mathcal{H}_2}$, we want $H_r(s)$ to match $H(s)$ at both the poles of $H(s)$ and at its own poles. While Gugercin and Antoulas in [14] illustrated the benefits of choosing the interpolation points, σ_i, to be the poles of $H(s)$ associated with the larger

residuals, [15] proves that the second term of the sum is actually more important. In fact, [15] shows that the optimal selection of interpolation points, σ_i, is for $\sigma_i = -\hat{\lambda}_i$.

Therefore, from theorem 1.4, the importance of the reduced order model's and full order model's poles is observed. The next theorem also reflects the pivotal importance of the reduced order model's poles.

Theorem 1.5. *Meier-Luenberger [19] Let $H(s)$ be the full order system and $H_r(s)$ be a minimizer for $\|H(s) - H_r(s)\|_{\mathcal{H}_2}$ with the simple poles of $H_r(s)$ denoted by $\hat{\lambda}_k$. Then*

$$H(-\hat{\lambda}_k) = H_r(-\hat{\lambda}_k) \qquad \text{and} \qquad H'(-\hat{\lambda}_k) = H_r'(-\hat{\lambda}_k) \qquad \text{for} \qquad k = 1, ..., r.$$

The Meier-Luenberger conditions as stated in theorem 1.5 provide the first order necessary conditions for $\mathcal{H}_2$ optimality, namely for a reduced order model to satisfy (1.6.1), Hermite interpolation of $H(s)$ must occur at the poles of $H_r(s)$.

1.6.1 Optimal $\mathcal{H}_2$

A plethora of research has shown that finding a global minimizer for the $\mathcal{H}_2$ error is an extremely arduous process. For example, see [24], [21], [7], [19], [17], and [23]. As a result, the common method is to focus on fulfilling only the first order conditions for $\mathcal{H}_2$ optimality. Especially for large-scale dynamical systems, the expensive computations involved often make these methods not practical. The key expense is usually the solution of the Lyapunov equations; however, Gugercin et al. in [15] proposed the "Iterative Rational Krylov Algorithm" (IRKA), which circumvents solving the Lyapunov equations by using iterative rational Krylov steps.

1.6.2 IRKA

The key feature of IRKA is its ability to satisfy the Meier-Luenberger conditions without explicitly computing solutions of the expensive Lyapunov equations. IRKA iterates the Krylov reduction process and assigns $\sigma_i \leftarrow -\lambda_i(\mathbf{A}_r)$ as the new interpolation points until the iteration converges to the optimal shift selection as defined by the Meier-Luenberger conditions.

The motivation for the shift iterate is as follows. Define $\boldsymbol{\sigma} := \{\sigma_1, ..., \sigma_r\}$ where σ_i is an interpolation point and define $\boldsymbol{\lambda} := \{\hat{\lambda}_1, ..., \hat{\lambda}_r\}$ where $\hat{\lambda}_i$ is an eigenvalue of $\mathbf{A}_r$. Note that $\boldsymbol{\lambda}$ is obtained by constructing $\mathbf{A}_r = (\mathbf{W}_r^*\mathbf{V}_r)^{-1}\mathbf{W}_r^*\mathbf{A}\mathbf{V}_r$ where

$$\mathbf{V}_r = \{(\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, ..., (\sigma_r\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\} \quad \textit{and}$$

$$\mathbf{W}_r = \{(\sigma_1\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}, ..., (\sigma_r\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}\}.$$

Hence, $\boldsymbol{\lambda}$ depends on $\boldsymbol{\sigma}$. To capture this relationship, we define $\boldsymbol{\lambda}(\boldsymbol{\sigma}) : \mathbb{C}^r \mapsto \mathbb{C}^r$ as $\boldsymbol{\lambda}(\boldsymbol{\sigma}) := \{\hat{\lambda}_1, ..., \hat{\lambda}_r\}$. Since the optimal shifts σ_i are required to satisfy $\sigma_i = -\hat{\lambda}_i$, the notation above gives us an equivalent condition

$$\boldsymbol{\sigma} = -\boldsymbol{\lambda}(\boldsymbol{\sigma}). \tag{1.6.2}$$

Equivalently, we can define $\mathbf{g}(\boldsymbol{\sigma}) : \mathbb{C}^r \mapsto \mathbb{C}^r$ as $\mathbf{g}(\boldsymbol{\sigma}) := \boldsymbol{\lambda}(\boldsymbol{\sigma}) + \boldsymbol{\sigma}$, implying that (1.6.2) is satisfied when $\mathbf{g}(\boldsymbol{\sigma}) = 0$. Therefore, the $\mathcal{H}_2$ optimality condition can be treated as a root-finding problem, which can be solved with methods such as the Newton method. In this framework, the Newton method is defined by

$$\boldsymbol{\sigma}^{(k+1)} = \boldsymbol{\sigma}^{(k)} - (\mathbf{I} + \mathbf{J})^{-1}(\boldsymbol{\sigma}^{(k)} + \boldsymbol{\lambda}(\boldsymbol{\sigma}^{(k)})) \tag{1.6.3}$$

where $\mathbf{J}$ denotes the $r \times r$ Jacobian of $\boldsymbol{\lambda}(\boldsymbol{\sigma})$ with respect to $\boldsymbol{\sigma}$. Since the entries of the Jacobian become small in the neighborhood of an optimal shift selection, $\mathbf{J}$ is taken to be the zero matrix, resulting in an iteration defined by $\boldsymbol{\sigma}^{(k+1)} = -\boldsymbol{\lambda}^{(k)}(\mathbf{A}_r)$. Therefore, the assignment of $\sigma_i \leftarrow -\lambda_i(\mathbf{A}_r)$ implies that upon convergence the interpolation points will satisfy the Meier-Luenberger conditions for $\mathcal{H}_2$ optimality.

Algorithm 1.6.1. *[15]* **An Iterative Rational Krylov Algorithm (IRKA):**

1. *Make an initial shift selection* σ_i *for* $i = 1, \ldots, r$
2. $\mathbf{W} = [\ (\sigma_1\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}, \cdots, (\sigma_r\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}\ ]$
3. $\mathbf{V} = [\ (\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, \cdots, (\sigma_r\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\ ]$
4. $\mathbf{W} = \mathbf{W}(\mathbf{W}^*\mathbf{V})^{-*}$ *(to make* $\mathbf{W}^*\mathbf{V} = \mathbf{I}_r$*)*
5. *while (not converged)*
   - *(a)* $\mathbf{A}_r = \mathbf{W}^*\mathbf{A}\mathbf{V}$,
   - *(b)* $\sigma_i \longleftarrow -\lambda_i(\mathbf{A}_r)$ *for* $i = 1, \ldots, r$
   - *(c)* $\mathbf{W} = [(\sigma_1\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}, \cdots, (\sigma_r\mathbf{I} - \mathbf{A})^{-*}\mathbf{c}]$
   - *(d)* $\mathbf{V} = [\ (\sigma_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, \cdots, (\sigma_r\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}\ ]$
   - *(e)* $\mathbf{W} = \mathbf{W}(\mathbf{W}^*\mathbf{V})^{-*}$ *(to make* $\mathbf{W}^*\mathbf{V} = \mathbf{I}_r$*)*
6. $\mathbf{A}_r = \mathbf{W}^*\mathbf{A}\mathbf{V}$, $\mathbf{b}_r = \mathbf{W}^*\mathbf{b}$, $\mathbf{c}_r^T = \mathbf{c}^T\mathbf{V}$

The main cost of IRKA is the solution of the $2r$ linear systems at each iteration. Especially in a large-scale setting, inexact solves will need to be employed in solving these systems. Using inexact solves creates new concerns since exact Hermite interpolation of H will no longer hold. In fact, the main aim of this thesis is to investigate the impact of this loss of interpolation on the model reduction process.

Chapter 2

Inexact Solvers

Krylov-based model reduction methods assume that the systems $(\sigma_j \mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ and $(\sigma_j \mathbf{I} - \mathbf{A})^* \mathbf{y}_j = \mathbf{c}$ are solved directly. The need for more accuracy, however, often augments the dimension of the dynamical system to the point where direct solves become computationally infeasible or ill-conditioned. Since $\mathbf{A}$ is typically sparse, this is an ideal setting in which to employ iterative methods. In this section, we will provide a brief overview of the iterative solvers we studied, namely GMRES and BiCG.

2.1 GMRES

Since solving the linear system $\mathbf{Ax} = \mathbf{b}$ is equivalent to $\mathbf{Ax} - \mathbf{b} = 0$, a reasonable goal for an iterative solver might be to minimize $\|\mathbf{Ax} - \mathbf{b}\|_2$ at each step. In fact, the Generalized Minimum Residual method (GMRES) aims to do this over a particular subspace, known as the Krylov subspace. The Krylov subspace of dimension n, is defined as

$$\mathcal{K}_n = \text{span}\{\mathbf{b}, \mathbf{Ab}, ..., \mathbf{A}^{n-1}\mathbf{b}\}.$$

2.1.1 The Arnoldi Iteration

Using just the vectors $\mathbf{b}, \mathbf{Ab}, ..., \mathbf{A}^{n-1}\mathbf{b}$ to describe this subspace is exceptionally problematic since these vectors converge to the leading eigenvector as $n \to \infty$. As a result, GMRES uses the Arnoldi iteration to construct an orthogonal basis for this Krylov subspace. We present the pseudocode for the Arnoldi-Modified Gram-Schmidt iteration, which is a numerically robust method for constructing an orthogonal basis for the Krylov subspace, $\mathcal{K}_m = \text{span}\{\mathbf{v}_1, \mathbf{A}\mathbf{v}_1, ...\mathbf{A}^{m-1}\mathbf{v}_1\}$.

Algorithm 2.1.1. *[20]* **Arnoldi-Modified Gram-Schmidt Iteration**

1. *Choose a vector* $\mathbf{v}_1$ *with* $\|\mathbf{v}_1\| = 1$.
2. *For* $j = 1, 2...m$ *Do:*
   - *(a) Compute* $\mathbf{w}_j = \mathbf{A}\mathbf{v}_j$
   - *(b) For* $i = 1, \ ... \ , j$ *Do:*
     - *i.* $h_{ij} = \langle \mathbf{w}_j, \mathbf{v}_i \rangle$
     - *ii.* $\mathbf{w}_j = \mathbf{w}_j - h_{ij}\mathbf{v}_i$
   - *(c) EndDo*
   - *(d)* $h_{j+1,j} = \|\mathbf{w}_j\|_2$. *If* $h_{j+1,j} = 0$, *stop.*
   - *(e)* $\mathbf{v}_{j+1} = \mathbf{w}_j / h_{j+1,j}$
   - *(f) EndDo*

Clearly, the only potentially undefined step of algorithm 2.1.1 is if $h_{j+1,j} = 0$ for some j. In this case, the subspace $\mathcal{K}_j$ is invariant under $\mathbf{A}$, implying that the exact solution

of $\mathbf{Ax} = \mathbf{b}$ lies in $\mathcal{K}_j$. Therefore, Arnoldi only has the potential for lucky breakdowns. In addition to Arnoldi-Modified Gram-Schmidt, Householder Arnoldi is also a numerically robust algorithm, and the pseudocode for this may be found in [20].

GMRES uses the Arnoldi method to create an orthonormal basis for the Krylov subspace $\text{span}\{\mathbf{r}_0, \mathbf{Ar}_0, ..., \mathbf{A}^k\mathbf{r}_0\}$. The aim of GMRES is to approximate the solution by $\mathbf{x}_n \in \mathcal{K}_n$ where $\mathbf{x}_n$ minimizes $\|\mathbf{r}_n\|_2 = \|\mathbf{b} - \mathbf{Ax}_n\|_2$. The least-squares problem is solved by using Givens rotations to transform the Hessenberg matrix system into an upper triangular matrix-vector equation, which is then solved by backward substitution. While another formulation of GMRES uses the modified Gram-Schmidt orthogonalization for the Arnoldi process, we present the pseudocode for GMRES using Householder orthogonalization; this is MATLAB's implementation of GMRES, which we used for our numerical simulations.

Algorithm 2.1.2. *[20]* **GMRES with Householder Orthogonalization**

1. *Compute* $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$, $\mathbf{z} := \mathbf{r}_0$
2. *For* $j = 1, ...m, m+1$, *do:*
   (a) *Compute the Householder unit vector* $\mathbf{w}_j$ *such that*
      i. $(\mathbf{w}_j)_i = 0, i = 1, ...j-1$ *and*
      ii. $(\mathbf{P}_j\mathbf{z})_i = 0, i = j+1, ..., n$ *where* $\mathbf{P}_j = \mathbf{I} - 2\mathbf{w}_j\mathbf{w}_j{}^*$;
   (b) $\mathbf{h}_{j-1} := \mathbf{P}_j\mathbf{z}$; *If* $j = 1$, *then let* $\beta := \mathbf{e}_1{}^*\mathbf{h}_0$.
   (c) $\mathbf{v} := \mathbf{P}_1\mathbf{P}_2...\mathbf{P}_j\mathbf{e}_j$
   (d) *If* $j \leq m$ *compute* $\mathbf{z} := \mathbf{P}_j\mathbf{P}_{j-1}...\mathbf{P}_1\mathbf{Av}$,
3. *EndDo*
4. *Define* $\bar{\mathbf{H}}_m =$ *the* $(m+1) \times m$ *upper part of the matrix* $[\mathbf{h}_1, ..., \mathbf{h}_m]$.

5. *Compute* $\mathbf{y}_m = Argmin_{\mathbf{y}} \|\beta \mathbf{e}_1 - \bar{\mathbf{H}}_m \mathbf{y}\|_2$. *Let* $\mathbf{y}_m = (\eta_1, \eta_2, ...\eta_m)^T$.

6. $\mathbf{z} := 0$

7. *For* $j = m, m-1, ...1$ *Do:*

   (a) $\mathbf{z} := \mathbf{P}_j(\eta_j \mathbf{e}_j + \mathbf{z})$

8. *EndDo*

9. *Compute* $\mathbf{x}_m = \mathbf{x}_0 + \mathbf{z}$

While GMRES only has "lucky breakdowns" due to the embedded Arnoldi iteration, one disadvantage is the storage required for implementation. To circumvent these storage issues, GMRES(j) restarts the GMRES iteration every j steps by using the last iterate from the previous GMRES cycle as an initial guess in the next cycle. The main conundrum associated with GMRES(j) is selecting a successful j value. Assuming exact arithmetic, GMRES is guaranteed to converge in at most n steps; however, a similar guarantee does not exist for GMRES(j). In fact, convergence may take place only at the n^{th} step, implying GMRES(j) will fail to converge for all $j < n$. For more details about GMRES, see [20], [8], [11], [2], and [22].

2.2 BiCG

The other iterative solver we studied is BiCG. BiCG is based on the Lanczos Biorthogonalization process, which constructs biorthogonal bases for the Krylov subspaces associated with $\mathbf{A}$ and $\mathbf{A}^*$.

2.2.1 Lanczos Biorthogonalization Process

Algorithm 2.2.1. *[11]* **Lanczos Biorthogonalization Process:**

1. *Given* $\mathbf{r}_0$ *and* $\hat{\mathbf{r}}_0$ *with* $\langle \mathbf{r}_0, \hat{\mathbf{r}}_0 \rangle \neq 0$, *set* $\mathbf{v}_1 = \frac{\mathbf{r}_0}{\|\mathbf{r}_0\|}$ *and* $\mathbf{w}_1 = \frac{\hat{\mathbf{r}}_0}{\langle \hat{\mathbf{r}}_0, \mathbf{v}_1 \rangle}$.
2. *Set* $\beta_0 = \gamma_0 = 0$ *and* $\mathbf{v}_0 = \mathbf{w}_0 = \mathbf{0}$.
3. *For* $j = 1, 2...$
   (a) *Compute* $\mathbf{A}\mathbf{v}_j$ *and* $\mathbf{A}^*\mathbf{w}_j$.
   (b) *Set* $\alpha_j = \langle \mathbf{A}\mathbf{v}_j, \mathbf{w}_j \rangle$.
   (c) *Compute* $\tilde{\mathbf{v}}_{j+1} = \mathbf{A}\mathbf{v}_j - \alpha_j \mathbf{v}_j - \beta_j \mathbf{v}_{j-1}$ *and* $\tilde{\mathbf{w}}_{j+1} = \mathbf{A}^*\mathbf{w}_j - \bar{\alpha}_j \mathbf{w}_j - \gamma_{j-1}\mathbf{w}_{j-1}$
   (d) *Set* $\gamma_j = \|\tilde{\mathbf{v}}_{j+1}\|$ *and* $\mathbf{v}_{j+1} = \tilde{\mathbf{v}}_{j+1}/\gamma_j$
   (e) *Set* $\beta_j = \langle \mathbf{v}_{j+1}, \tilde{\mathbf{w}}_{j+1} \rangle$ *and* $\mathbf{w}_{j+1} = \tilde{\mathbf{w}}_{j+1}/\bar{\beta}_j$.

As the algorithm illustrates, $\mathbf{A}$ and $\mathbf{A}^*$ play a dual role. This implies that both $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{A}^*\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ can be simultaneously solved, making BiCG ideal for problems requiring the solution of two linear systems involving $\mathbf{A}$ and $\mathbf{A}^*$. However, if only the solution for one system is needed, then the operations with $\mathbf{A}^*$ are superfluous. One advantage of algorithm 2.2.1 is storage, since only six vectors of length n need to be stored along with the coefficients α, β, and γ. A potential drawback is that there is no explicit reorthogonalization; instead, the orthogonality emerges from a mathematical identity. As a result, the presence of rounding errors often results quickly in the loss of orthogonality. Another disadvantage is the procedure's susceptibility to breakdowns. If $\tilde{\mathbf{v}}_{j+1} = 0$ or $\tilde{\mathbf{w}}_{j+1} = 0$, then we have a "lucky breakdown" since the invariant subspace has been found for either $\mathbf{A}$ or $\mathbf{A}^*$, respectively. However, there are cases when $\langle \tilde{\mathbf{v}}_{j+1}, \tilde{\mathbf{w}}_{j+1} \rangle = 0$ or $\langle \tilde{\mathbf{v}}_{j+1}, \tilde{\mathbf{w}}_{j+1} \rangle \approx 0$ even though $\tilde{\mathbf{v}}_{j+1} \neq 0$

and $\tilde{\mathbf{w}}_{j+1} \neq 0$. To address this issue, look-ahead methods allow slightly longer recurrences at selected steps so that the algorithm can continue until convergence.

The BiCG algorithm is then derived from this Lanczos process; therefore, the main disadvantage associated with BiCG is its erratic convergence. We present the pseudocode for the preconditioned case below. For more details about BiCG, see [20], [11], [2], and [22].

Algorithm 2.2.2. *[2]* **Preconditioned BiCG** *Let* $\mathbf{M}$ *be a preconditioner for the system* $Ax = b$.

1. *Compute* $\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}$ *for some initial guess* $\mathbf{x}^{(0)}$.
2. *Choose* $\tilde{\mathbf{r}}^{(0)} = \tilde{\mathbf{b}} - \mathbf{A}^*\tilde{\mathbf{x}}^{(0)}$.
3. *for i =1,2...*
   - (a) *Solve* $\mathbf{M}\mathbf{z}^{(i-1)} = \mathbf{r}^{(i-1)}$
   - (b) *Solve* $\mathbf{M}^*\tilde{\mathbf{z}}^{(i-1)} = \tilde{\mathbf{r}}^{(i-1)}$
   - (c) $\rho_{i-1} = \mathbf{z}^{(i-1)^*}\tilde{\mathbf{r}}^{(i-1)}$
   - (d) *if* $\rho_{i-1} = 0$, *method fails.*
   - (e) *if* $i = 1$
     - *i.* $\mathbf{p}^{(i)} = \mathbf{z}^{(i-1)}$
     - *ii.* $\tilde{\mathbf{p}}^{(i)} = \tilde{\mathbf{z}}^{(i-1)}$
   - (f) *else*
     - *i.* $\beta_{i-1} = \frac{\rho_{i-1}}{\rho_{i-2}}$
     - *ii.* $\mathbf{p}^{(i)} = \mathbf{z}^{(i-1)} + \beta_{i-1}\mathbf{p}^{(i-1)}$
     - *iii.* $\tilde{\mathbf{p}}^{(i)} = \tilde{\mathbf{z}}^{(i-1)} + \bar{\beta}_{i-1}\tilde{\mathbf{p}}^{(i-1)}$

(g) *endif*

(h) $\mathbf{q}^{(i)} = \mathbf{A}\mathbf{p}^{(i)}$

(i) $\tilde{\mathbf{q}}^{(i)} = \mathbf{A}^*\tilde{\mathbf{p}}^{(i)}$

(j) $\alpha_i = \frac{\rho_{i-1}}{\tilde{\mathbf{p}}^{(i)*}\mathbf{q}^{(i)}}$

(k) $\mathbf{x}^{(i)} = \mathbf{x}^{(i-1)} + \alpha_i\mathbf{p}^{(i)}$

(l) $\tilde{\mathbf{x}}^{(i)} = \tilde{\mathbf{x}}^{(i-1)} + \bar{\alpha}_i\tilde{\mathbf{p}}^{(i)}$

(m) $\mathbf{r}^{(i)} = \mathbf{r}^{(i-1)} - \alpha_i\mathbf{q}^{(i)}$

(n) $\tilde{\mathbf{r}}^{(i)} = \tilde{\mathbf{r}}^{(i-1)} - \bar{\alpha}_i\tilde{\mathbf{q}}^{(i)}$

(o) *continue until convergence*

While both GMRES and BiCG are Krylov subspace methods, it is important to keep in mind the distinct differences between the two solvers. The GMRES algorithm aims to solve $\mathbf{A}\mathbf{x} = \mathbf{b}$ by choosing $\mathbf{x}_n \in \mathcal{K}_n$ such that the residual $\|\mathbf{r}_n\|_2 = \|\mathbf{b} - \mathbf{A}\mathbf{x}_n\|_2$ satisfies

$$\mathbf{r}_n \perp \text{span}\{\mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, ..., \mathbf{A}^n\mathbf{b}\}.$$

In this way, $\|\mathbf{r}_n\|_2$ is minimized. Meanwhile, BiCG does not minimize $\|\mathbf{r}_n\|_2$. Instead, the algorithm selects $\mathbf{x}_n$ and $\tilde{\mathbf{x}}_n$ such that the following Petrov-Galerkin framework is employed:

$$\mathbf{b} - \mathbf{A}\mathbf{x}_n \perp \text{span}\{\mathbf{w}_1, \mathbf{A}^*\mathbf{w}_1, ..., (\mathbf{A}^*)^{(n-1)}\mathbf{w}_1\} \text{ and}$$

$$\tilde{\mathbf{b}} - \mathbf{A}^*\tilde{\mathbf{x}}_n \perp \text{span}\{\mathbf{v}_1, \mathbf{A}\mathbf{v}_1, ..., (\mathbf{A})^{(n-1)}\mathbf{v}_1\}.$$

BiCG is implemented with three-term recurrences while GMRES uses (n+1)-term recurrences, giving BiCG an advantage in terms of storage. Whereas convergence for GMRES is

monotonic, the convergence curves of BiCG are often characterized by erratic convergence. This type of convergence behavior may potentially result in the accumulation of rounding errors, which preclude the iteration from reaching the desired accuracy. For each iteration, GMRES requires one matrix-vector multiplication using $\mathbf{A}$ whereas BiCG requires multiplications involving $\mathbf{A}$ and $\mathbf{A}^*$. As a result, GMRES may require less work in comparison to BiCG if forming the matrix-vector product is computationally intensive. In other cases, even with this additional matrix-vector product, BiCG might be cheaper depending on the number of iterations GMRES requires. As with all iterative solvers, the decision to use a particular solver will depend on the given problem. For more details, see [11], [20] and [22]. For this thesis, the problem setting is to solve $(\sigma_j \mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ and $(\sigma_j \mathbf{I} - \mathbf{A})^* \mathbf{y}_j = \mathbf{c}$. In chapter 4, we will further investigate additional features offered by GMRES and BiCG in terms of obtaining a reduced inexact model in a numerically efficient and robust manner.

Chapter 3

Inexact Solves in Model Reduction

The aim of this chapter is to analyze the impact of inexact solves when just one step of Krylov reduction is performed for a given shift selection. Of course, the first question is with what accuracy should we solve the linear systems. Moreover, we investigate the role of the stopping tolerance in the interpolation error that emerges due to the loss of Hermite interpolation. To assess the quality of the inexact model, we also study both the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ error norms. Furthermore, we consider the effect of shift selection as well as the effect of preconditioning on the inexact iteration.

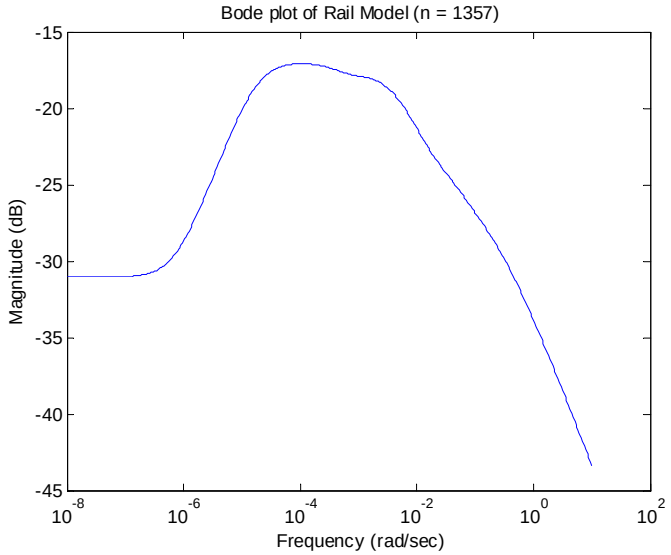
3.1 Models used in simulations

Throughout this thesis, we will be studying three models. The first two models emerge from a semi-discretized heat transfer problem for optimal cooling of steel profiles. After a finite element discretization, we obtain a descriptor system of the form

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), y(t) = \mathbf{c}^T\mathbf{x}(t)$$

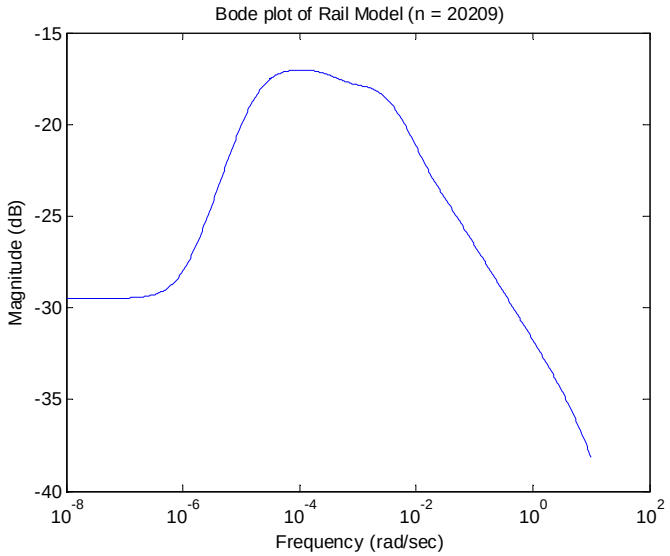
where $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times 7}, \mathbf{C} \in \mathbb{R}^{6\times n}$ and n depends on the mesh width of the discretization. We only considered the SISO system that relates the sixth input of the system to the second output. Although $\mathbf{E} \neq \mathbf{I}_n$, $\mathbf{E}$ is a positive definite matrix, and the discussion regarding $\sigma\mathbf{I}$ can be easily generalized to $\sigma\mathbf{E}$ provided that the appropriate changes are implemented. We will be considering two different discretizations. The first has a coarser mesh and results in a model of dimension 1357; we will refer to this model as simply Rail 1357. The second discretization has a finer mesh that leads to a model of dimension 20209 and will be referenced as Rail 20209. Figures 3.1 and 3.2 depict the Bode plots of Rail 1357 and Rail 20209. For more details, see [5] and [6].

Figure 3.1: Bode plot of Rail 1357



The other model we studied is the CD model. This model describes the dynamics emerging from the lens actuator and the radial arm position of a CD player. The dynamical system is of the form

Figure 3.2: Bode plot of Rail 20209



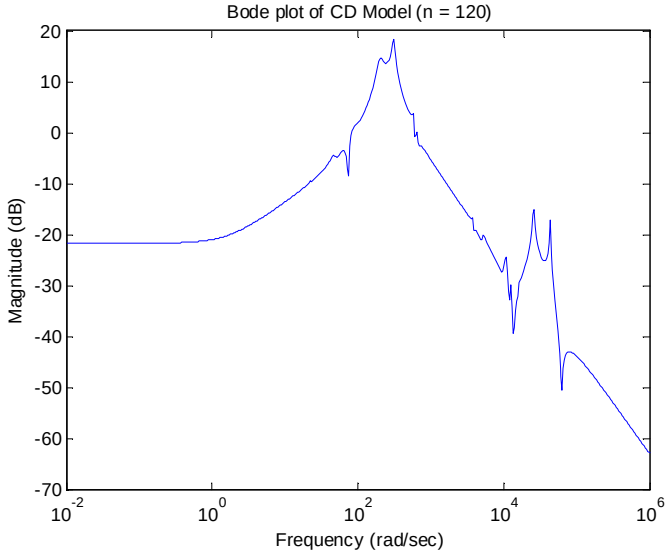
$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}\mathbf{u}(t), y(t) = \mathbf{c}^T\mathbf{x}(t)$$

where $\mathbf{A} \in \mathbb{R}^{120\times 120}$ and $\mathbf{b} \in \mathbb{R}^{120}, \mathbf{c} \in \mathbb{R}^{120}$. Figure 3.3 is the corresponding Bode plot of the CD model. For more details about this model, see [1] and [13]. It is important to note that the poles of the CD model are complex whereas the poles of the Rail models are real.

3.2 Effect of Tolerance

By solving the systems $(\sigma_j\mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ and $(\sigma_j\mathbf{I} - \mathbf{A})^*\mathbf{y}_j = \mathbf{c}$ inexactly, we will only be obtaining an approximation to the solution. As a result, a crucial question is to what accuracy should these systems be solved so that the approximation is suitable for model reduction purposes. The solver's tolerance refers to the accuracy with which the linear system is solved, namely in solving the system, $\mathbf{A}\mathbf{x} = \mathbf{b}$, the iterative solver returns a solution with relative residual $\frac{\|\mathbf{b}-\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{b}\|_2} < tolerance$. Choosing a small tolerance for the inexact solver may

Figure 3.3: Bode plot of CD Model



allow for a good approximation to the exact solution, but it also carries additional cost, which we may not be able to afford in the large-scale setting. Of course, selecting too large of a tolerance may lead to such a poor approximation of the exact solution that our reduced model is too inaccurate. In this section, we investigate the impact of the inexact solver's tolerance on model reduction. Unless otherwise stated, H, H_r, and $\widetilde{H}_r$ will denote the full order model, exact reduced order model, and the inexact reduced order model, respectively, for the remainder of this thesis.

3.2.1 Interpolation Error

Since the inexact model will no longer interpolate the full order model at the selected interpolation points, an interpolation error emerges. Beattie and Gugercin in [4] delineate connections between the interpolation error and the residuals associated with inexact solves. Following the discussion from [4], $\tilde{\mathbf{x}}_j$ will denote the inexact solution to the j^{th} linear system

$(\sigma_j \mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$. The associated residual for $\tilde{\mathbf{x}}_j$ will then be denoted by $\delta\mathbf{b}_j$, namely

$$\delta\mathbf{b}_j = (\sigma_j \mathbf{I} - \mathbf{A})\tilde{\mathbf{x}}_j - \mathbf{b}.$$

In the one-sided case, only $\widetilde{\mathbf{V}}_r$ is used, and the reduced order model is described by

$$\widetilde{\mathbf{H}}_r(s) = \widetilde{\mathbf{c}}_r^T (s\mathbf{I}_r - \widetilde{\mathbf{A}}_r)^{-1}\widetilde{\mathbf{b}}_r$$

where $\widetilde{\mathbf{A}}_r = \widetilde{\mathbf{V}}_r^* \mathbf{A} \widetilde{\mathbf{V}}_r, \widetilde{\mathbf{b}}_r = \widetilde{\mathbf{V}}_r^* \mathbf{b}$, and $\widetilde{\mathbf{c}}_r^T = \mathbf{c}^T \widetilde{\mathbf{V}}_r$. In [4], Beattie and Gugercin presented an expression that links the interpolation error at σ_j and the inexact solve's residual, namely

$$|H(\sigma_j) - H_r(\sigma_j)| = \|\delta\mathbf{b}_j\| \cdot \mathbf{c}^T (\sigma_j \mathbf{I}_n - \mathbf{A})^{-1} (\mathbf{I}_n - \mathbf{P}_r) \frac{\delta\mathbf{b}_j}{\|\delta\mathbf{b}_j\|} \tag{3.2.1}$$

where $\mathbf{P_r} = (\sigma_\mathbf{j}\mathbf{I_n} - \mathbf{A})\widetilde{\mathbf{V}}_\mathbf{r}(\sigma_\mathbf{j}\mathbf{I_r} - \widetilde{\mathbf{A}}_\mathbf{r})^{-\mathbf{1}}\widetilde{\mathbf{V}}_\mathbf{r}^*$. Similarly, for two-sided Krylov reduction, the interpolation error at σ_j can be related to the inexact solve error by the following expression:

$$|H(\sigma_j) - H_r(\sigma_j)| = \|\delta\mathbf{b}_j\| \cdot \|\delta\mathbf{c}_j\| \cdot \frac{\delta\mathbf{c}_j^T}{\|\delta\mathbf{c}_j\|} \mathbf{M}_j (\sigma_j \mathbf{I}_n - \mathbf{A}) \mathbf{M}_j \frac{\delta\mathbf{b}_j}{\|\delta\mathbf{b}_j\|} \tag{3.2.2}$$

where $\mathbf{M}_j = (\sigma_j \mathbf{I}_n - \mathbf{A})^{-1} - \widetilde{\mathbf{V}}_r (\sigma_j \mathbf{I}_r - \widetilde{\mathbf{A}}_r)^{-1} \widetilde{\mathbf{W}}_r^*$. From (3.2.2), [4] noted the quadratic behavior as an advantage of two-sided Krylov reduction since (3.2.1) demonstrates that the one-sided case is characterized by only linear behavior. We observed this behavior in our numerical results. For each model, we varied the tolerance from 1×10^{-1} to 1×10^{-10} and reduced to order six and fourteen for the Rail and CD models, respectively. Then we calculated the relative interpolation error given by $\frac{|H(\sigma_i) - \widetilde{H}_r(\sigma_i)|}{|H(\sigma_i)|}$. Throughout this data, we observed roughly linear and quadratic behavior of the relative interpolation error for the

one-sided and two-sided cases, respectively. Tables 3.1 and 3.2 report the data for the Rail 1357 model and exemplify the behavior found throughout our numerical simulations.

Table 3.1: $\frac{|H(\sigma_i)-\widetilde{H}_r(\sigma_i)|}{|H(\sigma_i)|}$; Krylov Reduction for Rail 1357; One-Sided Reduction

Shift	tolerance $= 10^{-1}$	tolerance $= 10^{-2}$	tolerance $= 10^{-3}$
σ_1	3.31×10^{-2}	2.12×10^{-3}	5.27×10^{-5}
σ_2	6.30×10^{-2}	4.06×10^{-3}	1.06×10^{-4}
σ_3	1.92×10^{-2}	8.20×10^{-4}	6.12×10^{-5}
σ_4	4.53×10^{-2}	2.16×10^{-3}	1.86×10^{-4}
σ_5	2.49×10^{-2}	2.00×10^{-3}	3.22×10^{-4}
σ_6	7.37×10^{-2}	1.46×10^{-2}	1.15×10^{-4}

Table 3.2: $\frac{|H(\sigma_i)-\widetilde{H}_r(\sigma_i)|}{|H(\sigma_i)|}$; Krylov Reduction for Rail 1357; Two-Sided Reduction

Shift	tolerance $= 10^{-1}$	tolerance $= 10^{-2}$	tolerance $= 10^{-3}$
σ_1	2.23×10^{-2}	3.95×10^{-5}	1.82×10^{-7}
σ_2	1.12×10^{-3}	4.33×10^{-6}	4.75×10^{-8}
σ_3	1.25×10^{-2}	6.42×10^{-5}	3.27×10^{-8}
σ_4	6.70×10^{-2}	1.77×10^{-3}	1.05×10^{-7}
σ_5	3.80×10^{-1}	2.08×10^{-3}	2.99×10^{-6}
σ_6	4.78	3.09×10^{-2}	6.85×10^{-6}

It is important to note that the linear systems were solved with the same accuracy; however, we see a significant improvement in the relative interpolation error with two-sided reduction.

3.2.2 Model Reduction Error

In this section, we examine the impact of the inexact solver's stopping tolerance on the model reduction errors. For a given model, we performed one step of Krylov reduction to obtain reduced order models of order six, and we considered using tolerances of $1 \times 10^{-1}, 1 \times$

$10^{-2}, ..., 1 \times 10^{-10}$. For each tolerance, we performed one step of Krylov reduction and then calculated the following error norms:

$$\|H - \widetilde{H}_r\|_{\mathcal{H}_2},\ \|H - H_r\|_{\mathcal{H}_2}, \text{ and } \|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}.$$

Similarly, we calculated $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$, $\|H - H_r\|_{\mathcal{H}_\infty}$, and $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$. By a triangle inequality argument, we have $\|H - \widetilde{H}_r\|_{\mathcal{H}_2} \leq \|H - H_r\|_{\mathcal{H}_2} + \|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$ and similarly for the $\mathcal{H}_\infty$ error. From Tables 3.3 and 3.5, we observe no improvement in $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ once $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$ is of the same or smaller order as $\|H - H_r\|_{\mathcal{H}_2}$ even though the linear systems are being solved more accurately. As Tables 3.4 and 3.6 illustrate, similar patterns occurred with the $\mathcal{H}_\infty$ norm.

Table 3.3: Krylov Reduction for CD Model; r = 6; GMRES; no prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$	$\|H - H_r\|_{\mathcal{H}_2}$	$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$
1×10^{-1}	∞	1.06×10^{1}	∞
1×10^{-2}	7.51	1.06×10^{1}	1.35×10^{1}
1×10^{-2}	∞	1.06×10^{1}	∞
1×10^{-4}	1.06×10^{1}	1.06×10^{1}	1.39
1×10^{-5}	1.05×10^{1}	1.06×10^{1}	2.45×10^{-1}
1×10^{-6}	1.06×10^{1}	1.06×10^{1}	1.61×10^{-2}
1×10^{-7}	1.06×10^{1}	1.06×10^{1}	2.60×10^{-3}
1×10^{-8}	1.06×10^{1}	1.06×10^{1}	1.66×10^{-4}
1×10^{-9}	1.06×10^{1}	1.06×10^{1}	2.10×10^{-5}
1×10^{-10}	1.06×10^{1}	1.06×10^{1}	1.77×10^{-6}

The fact that the model reduction errors $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ are of the same order for most tolerances has an important implication for interpolatory model reduction using inexact solves, namely expending the additional work to solve the systems more accurately seems to add only unnecessary expense. In the CD model, for example, the $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ norms are of the same order for tolerances 1×10^{-4} to 1×10^{-10}. Similarly, the associated $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ errors are of the same order for tolerances

Table 3.4: Krylov Reduction for CD Model; r = 6; GMRES; no prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$	$\|H - H_r\|_{\mathcal{H}_\infty}$	$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$
1×10^{-1}	∞	1.57×10^{-2}	∞
1×10^{-2}	1.89×10^{-2}	1.57×10^{-2}	5.02
1×10^{-3}	∞	1.57×10^{-2}	∞
1×10^{-4}	1.35×10^{-2}	1.57×10^{-2}	7.04×10^{-2}
1×10^{-5}	1.56×10^{-2}	1.57×10^{-2}	6.04×10^{-3}
1×10^{-6}	1.57×10^{-2}	1.57×10^{-2}	1.65×10^{-3}
1×10^{-7}	1.57×10^{-2}	1.57×10^{-2}	1.92×10^{-4}
1×10^{-8}	1.57×10^{-2}	1.57×10^{-2}	1.07×10^{-5}
1×10^{-9}	1.57×10^{-2}	1.57×10^{-2}	2.48×10^{-6}
1×10^{-10}	1.57×10^{-2}	1.57×10^{-2}	3.08×10^{-7}

1×10^{-3} to 1×10^{-10} for Rail 1357 as observed in Tables 3.5 and 3.6. Therefore, these results illustrate that we may use large tolerances in the inexact solve without augmenting the overall model reduction error.

3.3 Effect of Shift Selection

In this section, we will be examining the effect of shift selection on the overall model reduction error. The shift selections that lead to small $\|H - H_r\|_{\mathcal{H}_2}$ and $\|H - H_r\|_{\mathcal{H}_\infty}$ norms will be classified as good selections. Meanwhile, a selection is a poor shift selection if it leads to large $\|H - H_r\|_{\mathcal{H}_2}$ and $\|H - H_r\|_{\mathcal{H}_\infty}$ norms. Recall from section 1.6 that the optimal interpolation points are the mirror images of the poles of the reduced model. By definition of $\mathbf{A}_r$, this implies that the optimal interpolation points are expected to be located in the mirror spectrum of $\mathbf{A}$. As a result, choosing shifts within the convex hull of this mirror spectrum is anticipated to correspond to a good shift selection.

In this section, we present our data from the Rail 1357 and Rail 20209 models. For these

Table 3.5: Krylov Reduction for Rail 1357; r = 6; GMRES; no prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$	$\|H - H_r\|_{\mathcal{H}_2}$	$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$
1×10^{-1}	∞	6.96×10^{-6}	∞
1×10^{-2}	3.05×10^{-5}	6.96×10^{-6}	2.81×10^{-5}
1×10^{-3}	8.97×10^{-6}	6.96×10^{-6}	3.26×10^{-6}
1×10^{-4}	6.03×10^{-6}	6.96×10^{-6}	1.57×10^{-6}
1×10^{-5}	6.97×10^{-6}	6.96×10^{-6}	2.09×10^{-8}
1×10^{-6}	6.96×10^{-6}	6.96×10^{-6}	2.17×10^{-9}
1×10^{-7}	6.96×10^{-6}	6.96×10^{-6}	5.18×10^{-10}
1×10^{-8}	6.96×10^{-6}	6.96×10^{-6}	1.07×10^{-10}
1×10^{-9}	6.96×10^{-6}	6.96×10^{-6}	3.70×10^{-12}
1×10^{-10}	6.96×10^{-6}	6.96×10^{-6}	2.75×10^{-13}

models, we reduced to an order of six using $\sigma_i = \text{logspace}(-5, 0.7, 6)$ for a good shift selection and $\sigma_i = \text{logspace}(0.8, 1, 6)$ for a poor shift selection. For each model and shift selection, we considered using inexact tolerances from 1×10^{-1} to 1×10^{-10}.

3.3.1 Tolerance and Shift Selection

A relationship between the effect of tolerance and the quality of shift selection emerged when analyzing our numerical data. Let $\mathbf{V}_r$ and $\mathbf{W}_r$ be the matrices obtained when using direct solves. Also, let $\widetilde{\mathbf{V}}_r$ and $\widetilde{\mathbf{W}}_r$ be the matrices obtained when using an iterative solver. Since the Ran($\mathbf{V}_r$) and Ran($\mathbf{W}_r$) or Ran($\widetilde{\mathbf{V}}_r$) and Ran($\widetilde{\mathbf{W}}_r$) are used in constructing the reduced models, we are concerned with these subspaces. In general, if $\mathcal{X}$ and $\mathcal{Y}$ are subspaces of $\mathbb{C}$, then the angle between subspaces $\Theta(\mathcal{X}, \mathcal{Y})$ is defined as

$$\sup_{\mathbf{x}\in\mathcal{X}} \inf_{\mathbf{y}\in\mathcal{Y}} \frac{\|\mathbf{y}-\mathbf{x}\|}{\|\mathbf{x}\|} = \sin(\Theta(\mathcal{X}, \mathcal{Y})).$$

In our simulations, we performed one step of Krylov reduction using exact and inexact solves. We then calculated the angle between the subspaces formed by $\mathbf{V}_r$ and $\widetilde{\mathbf{V}}_r$ as well

Table 3.6: Krylov Reduction for Rail 1357; r = 6; GMRES; no prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$	$\|H - H_r\|_{\mathcal{H}_\infty}$	$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$
1×10^{-1}	∞	1.36×10^{-4}	∞
1×10^{-2}	6.30×10^{-4}	1.36×10^{-4}	5.86×10^{-4}
1×10^{-3}	2.10×10^{-4}	1.36×10^{-4}	1.36×10^{-4}
1×10^{-4}	1.30×10^{-4}	1.36×10^{-4}	2.36×10^{-5}
1×10^{-5}	1.36×10^{-4}	1.36×10^{-4}	2.58×10^{-6}
1×10^{-6}	1.36×10^{-4}	1.36×10^{-4}	2.17×10^{-7}
1×10^{-7}	1.36×10^{-4}	1.36×10^{-4}	5.44×10^{-9}
1×10^{-8}	1.36×10^{-4}	1.36×10^{-4}	3.60×10^{-9}
1×10^{-9}	1.36×10^{-4}	1.36×10^{-4}	1.08×10^{-10}
1×10^{-10}	1.36×10^{-4}	1.36×10^{-4}	2.78×10^{-11}

as $\mathbf{W}_r$ and $\widetilde{\mathbf{W}}_r$. Provided that a good shift selection was used, a linear improvement in the subspace angles appeared as the tolerance decreased. With a poor shift selection, the angles between the subspaces remained of the same order regardless of tolerance. Tables 3.7 and 3.8 exemplify these patterns.

Table 3.7: Krylov Reduction for Rail 1357; r = 6; $\sin(\Theta(\mathbf{V}_r, \widetilde{\mathbf{V}}_r))$; GMRES; no prec.

Tolerance	Good shift	Poor shift
1×10^{-1}	6.13×10^{-1}	9.99×10^{-1}
1×10^{-2}	1.02×10^{-1}	9.99×10^{-1}
1×10^{-3}	1.69×10^{-2}	9.96×10^{-1}
1×10^{-4}	1.24×10^{-3}	9.99×10^{-1}
1×10^{-5}	1.54×10^{-4}	9.97×10^{-1}
1×10^{-6}	1.77×10^{-5}	9.66×10^{-1}
1×10^{-7}	1.03×10^{-6}	9.95×10^{-1}
1×10^{-8}	1.09×10^{-7}	7.17×10^{-1}
1×10^{-9}	1.38×10^{-8}	1.69×10^{-1}
1×10^{-10}	1.47×10^{-9}	1.17×10^{-2}

This illustrates an important connection between the quality of shift selection and the effect of tolerance. Since a linear improvement occurs when a good shift selection is used, there is

Table 3.8: Krylov Reduction for Rail 1357; r = 6; $\sin(\Theta(\mathbf{W}_r, \widetilde{\mathbf{W}}_r))$; GMRES; no prec.

Tolerance	Good shift	Poor shift
1×10^{-1}	9.78×10^{-1}	9.99×10^{-1}
1×10^{-2}	9.96×10^{-1}	9.99×10^{-1}
1×10^{-3}	7.76×10^{-1}	9.98×10^{-1}
1×10^{-4}	1.66×10^{-2}	9.99×10^{-1}
1×10^{-5}	2.57×10^{-3}	9.97×10^{-1}
1×10^{-6}	1.56×10^{-4}	9.99×10^{-1}
1×10^{-7}	2.11×10^{-5}	9.96×10^{-1}
1×10^{-8}	3.33×10^{-6}	9.95×10^{-1}
1×10^{-9}	3.26×10^{-7}	9.83×10^{-1}
1×10^{-10}	2.24×10^{-8}	9.03×10^{-1}

a benefit in solving the linear systems more accurately as the inexact subspace will become a better approximation for the exact subspace. With a poor shift selection, however, the angles between the inexact and exact subspaces remain stagnant as the tolerance is decreased. Therefore, we can conclude that solving the systems more accurately may just introduce unnecessary costs.

These trends are then reflected in the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms. Throughout the data, $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ improved linearly as the tolerance decreased with a good shift selection; however, we do not see a similar pattern with a poor selection. Tables 3.9 and 3.10 exemplify this behavior.

Furthermore, the shift selection also impacts the overall model reduction error. In Table 3.11 and 3.12, for example, we observed that the overall model reduction $\mathcal{H}_2$ and $\mathcal{H}_\infty$ errors were two orders of magnitude larger with a poor shift selection than with a good shift selection. In general, the poor shift selection persistently yielded errors one to two orders greater than the errors obtained with a good shift selection. Another observation from Table 3.11 and 3.12 is the more erratic results obtained with a poor shift selection. The error is

Table 3.9: $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$; Krylov Reduction for Rail 1357; GMRES; no prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	∞	∞
1×10^{-2}	2.81×10^{-5}	∞
1×10^{-3}	3.26×10^{-6}	∞
1×10^{-4}	1.57×10^{-6}	2.70×10^{-4}
1×10^{-5}	2.09×10^{-8}	1.98×10^{-4}
1×10^{-6}	2.17×10^{-9}	∞
1×10^{-7}	5.18×10^{-10}	1.37×10^{-4}
1×10^{-8}	1.07×10^{-10}	1.18×10^{-4}
1×10^{-9}	3.70×10^{-12}	9.98×10^{-5}
1×10^{-10}	2.75×10^{-13}	8.55×10^{-5}

infinity for the first three tolerances and then for a tolerance of 10^{-6} when a poor shift is used. Meanwhile, for a good shift, the error often is infinity for only the larger tolerances. Then once $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty} < \infty$ for a certain tolerance, $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ remains bounded for all smaller tolerances. This pattern, as exhibited by Tables 3.11 and 3.12, held throughout all our simulations.

3.3.2 Relative Interpolation Error

From (3.2.1) and (3.2.2), [4] noted that the interpolation error depends on two factors: the quality of $\widetilde{\mathbf{V}}_r(\sigma_j \mathbf{I}_r - \widetilde{\mathbf{A}}_r)^{-1}\widetilde{\mathbf{V}}_r^*$ as an approximation to $(\sigma_j \mathbf{I}_n - \mathbf{A})^{-1}$ and the magnitude of the residual norms. As a result, a good shift selection is expected to yield a smaller relative interpolation error in comparison to the error obtained with a poor shift selection. However, our numerical results suggest that the quality of the shift selection does not greatly impact the interpolation error. From Table 3.13, we note the relative interpolation errors are roughly of the same order regardless of the shift selection. In fact, we often even observed smaller relative interpolation errors with poor shift selections as seen in Table 3.14.

Table 3.10: $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$; Krylov Reduction for Rail 1357; GMRES; no prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	∞	∞
1×10^{-2}	5.86×10^{-4}	∞
1×10^{-3}	1.36×10^{-4}	∞
1×10^{-4}	2.36×10^{-5}	2.60×10^{-3}
1×10^{-5}	2.58×10^{-6}	2.89×10^{-3}
1×10^{-6}	2.17×10^{-7}	∞
1×10^{-7}	5.44×10^{-9}	1.51×10^{-3}
1×10^{-8}	3.60×10^{-9}	1.31×10^{-3}
1×10^{-9}	1.08×10^{-10}	1.12×10^{-3}
1×10^{-10}	2.78×10^{-11}	9.75×10^{-4}

Therefore, this data suggests that the effect of shift selection impacts more drastically the overall model reduction errors $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$, $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$, $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$, and $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ rather than the interpolation error.

3.4 Preconditioning

Especially when convergence of the iterative solver is slow, preconditioning is employed. The aim of preconditioning is to find a matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{x} = \mathbf{P}^{-1}\mathbf{b}$ has faster convergence than if just $\mathbf{A}\mathbf{x} = \mathbf{b}$ is solved. In this thesis, we used an incomplete LU with a drop tolerance of 0.05. This is an LU decomposition where the diagonal entries of the upper triangular factor are preserved but all other entries are set to zero if the entry is smaller than $(drop \quad tolerance) \times \|\mathbf{A}_j\|_2$ where $\mathbf{A}_j$ is the j^{th} column of $\mathbf{A}$. For more details, see [20], [11], [2], and [22]. In our numerical simulations, we observed that preconditioning improved the convergence speed and also resulted in slightly smaller stopping residuals. An example of this is captured by Tables 3.15 and 3.16.

Table 3.11: $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$; Krylov Reduction for Rail 1357; GMRES; no prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	∞	∞
1×10^{-2}	3.05×10^{-5}	∞
1×10^{-3}	8.97×10^{-6}	∞
1×10^{-4}	6.03×10^{-6}	5.92×10^{-4}
1×10^{-5}	6.97×10^{-6}	3.03×10^{-4}
1×10^{-6}	6.96×10^{-6}	∞
1×10^{-7}	6.96×10^{-6}	5.29×10^{-4}
1×10^{-8}	6.96×10^{-6}	5.17×10^{-4}
1×10^{-9}	6.96×10^{-6}	5.07×10^{-4}
1×10^{-10}	6.96×10^{-6}	4.98×10^{-4}

With preconditioning, we observe improvements in our inexact reduced models. For example, Table 3.17 corresponds to the case presented in Tables 3.15 and 3.16 and shows the smaller error encountered due to preconditioning.

In general, we found that preconditioning often resulted in smaller $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ and $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ norms but did not impact any other performance measures. Hence, preconditioning can be employed to improve convergence speed and accuracy without any negative effects on the model reduction process.

Table 3.12: $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$; Krylov Reduction for Rail 1357; GMRES

Tolerance	Good Shift	Poor Shift
1×10^{-1}	∞	∞
1×10^{-2}	6.30×10^{-4}	∞
1×10^{-3}	2.10×10^{-4}	∞
1×10^{-4}	1.30×10^{-4}	1.57×10^{-2}
1×10^{-5}	1.36×10^{-4}	1.04×10^{-2}
1×10^{-6}	1.36×10^{-4}	∞
1×10^{-7}	1.36×10^{-4}	1.46×10^{-2}
1×10^{-8}	1.36×10^{-4}	1.44×10^{-2}
1×10^{-9}	1.36×10^{-4}	1.43×10^{-2}
1×10^{-10}	1.36×10^{-4}	1.41×10^{-2}

Table 3.13: $\frac{|H(\sigma_i)-\widetilde{H}_r(\sigma_i)|}{|H(\sigma_i)|}$; Krylov Reduction for Rail 20209; GMRES; $\varepsilon = 1 \times 10^{-6}$

Shift	Good shift	Poor shift
σ_1	4.75×10^{-17}	9.91×10^{-13}
σ_2	1.77×10^{-15}	8.39×10^{-13}
σ_3	1.87×10^{-13}	9.46×10^{-13}
σ_4	7.07×10^{-12}	7.25×10^{-13}
σ_5	1.99×10^{-11}	1.10×10^{-13}
σ_6	4.18×10^{-11}	6.17×10^{-14}

Table 3.14: $\frac{|H(\sigma_i)-\widetilde{H}_r(\sigma_i)|}{|H(\sigma_i)|}$; Krylov Reduction for Rail 20209; GMRES; $\varepsilon = 1 \times 10^{-3}$

Shift	Good shift	Poor shift
σ_1	2.70×10^{-8}	1.08×10^{-7}
σ_2	2.57×10^{-6}	8.73×10^{-8}
σ_3	1.21×10^{-4}	6.72×10^{-8}
σ_4	6.23×10^{-4}	6.23×10^{-8}
σ_5	1.70×10^{-3}	7.94×10^{-8}
σ_6	1.39×10^{-3}	1.14×10^{-7}

Table 3.15: Stopping Residuals for $\widetilde{\mathbf{V}}_r$; Krylov Reduction for Rail 1357; $\varepsilon = 1 \times 10^{-4}$; GMRES

Shift	No preconditioning	Preconditioning
σ_1	8.16×10^{-5}	3.45×10^{-5}
σ_2	8.61×10^{-5}	4.17×10^{-5}
σ_3	9.91×10^{-5}	3.58×10^{-5}
σ_4	9.72×10^{-5}	3.29×10^{-5}
σ_5	9.58×10^{-5}	4.33×10^{-5}
σ_6	9.82×10^{-5}	5.10×10^{-5}

Table 3.16: Stopping Residuals for $\widetilde{\mathbf{W}}_r$; Krylov Reduction for Rail Model 1357, $\varepsilon = 1 \times 10^{-4}$; GMRES

Shift	No preconditioning	Preconditioning
σ_1	8.01×10^{-5}	3.81×10^{-5}
σ_2	6.45×10^{-5}	6.64×10^{-6}
σ_3	8.63×10^{-5}	2.98×10^{-5}
σ_4	9.40×10^{-5}	5.26×10^{-5}
σ_5	9.88×10^{-5}	9.18×10^{-5}
σ_6	9.85×10^{-5}	9.24×10^{-5}

Table 3.17: Errors for Rail Model 1357; $\varepsilon = 1 \times 10^{-4}$

	No preconditioning	Preconditioning
$\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$	1.49×10^{-3}	8.70×10^{-5}
$\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$	1.31×10^{-4}	3.76×10^{-6}

Chapter 4

GMRES vs. BiCG

As noted in chapter 2, GMRES and BiCG offer different features. In this section, we compare the theoretical and numerical properties associated with GMRES and BiCG in the interpolatory model reduction setting.

4.1 Backward Error

Since we are solving the systems $(\sigma_j \mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{b}$ and $(\sigma_j \mathbf{I} - \mathbf{A})^* \mathbf{w} = \mathbf{c}$ inexactly, we would like to quantify the backward error. Following the discussion and notation of [3], let $\widetilde{\mathbf{v}}_j$ be an inexact solution for $(\sigma_j \mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{b}$ and $\widetilde{\mathbf{w}}_j$ be an inexact solution for $(\sigma_j \mathbf{I} - \mathbf{A})^* \mathbf{w} = \mathbf{c}$. Then we define the associated residuals of the inexact systems as:

$$\delta \mathbf{b}_j = (\sigma_j \mathbf{I} - \mathbf{A})\widetilde{\mathbf{v}}_j - \mathbf{b} \quad \text{and} \quad \delta \mathbf{c}_j = (\sigma_j \mathbf{I} - \mathbf{A})^* \widetilde{\mathbf{w}}_j - \mathbf{c}.$$

The associated "inexact bases" are then given as

$$\widetilde{\mathbf{V}}_r = [\widetilde{\mathbf{v}}_1,\, \widetilde{\mathbf{v}}_2,\, \ldots,\, \widetilde{\mathbf{v}}_r] \quad \text{and} \quad \widetilde{\mathbf{W}}_r = [\widetilde{\mathbf{w}}_1,\, \widetilde{\mathbf{w}}_2,\, \ldots,\, \widetilde{\mathbf{w}}_r]\,.$$

From numerical analysis, the standard backward error result holds for each of the linear systems solved, namely $\widetilde{\mathbf{v}}_j$ and $\widetilde{\mathbf{w}}_j$ are exact solutions of systems with perturbed right hand sides:

$$\widetilde{\mathbf{v}}_j = (\sigma_j \mathbf{I} - \mathbf{A})^{-1}(\mathbf{b} + \delta\mathbf{b}_j) \quad \text{and} \quad \widetilde{\mathbf{w}}_j = (\sigma_j \mathbf{I} - \mathbf{A})^{-*}(\mathbf{c} + \delta\mathbf{c}_j).$$

Using this result, it follows that the inexact reduced order model, $\widetilde{H}_r(s) = \widetilde{\mathbf{c}}_r^T(s\mathbf{I} - \widetilde{\mathbf{A}}_r)^{-1}\widetilde{\mathbf{b}}_r$, exactly interpolates a perturbed full order model at $s = \sigma_j$:

$$\widetilde{H}(s) = (\mathbf{c} + \delta\mathbf{c}_j)^T(s\mathbf{I} - \mathbf{A})^{-1}(\mathbf{b} + \delta\mathbf{b}_j).$$

Since $\widetilde{H}_r(s)$ will interpolate a different perturbed dynamical system, $\widetilde{H}(s)$, for $\sigma_1,\, \sigma_2,\, \ldots, \sigma_r$, this is not a desirable backward error result. Fortunately, [3] shows that if an inexact solver with the Petrov-Galerkin framework is used, then a more satisfactory backward error may be proven.

Lemma 4.1. *[3] Assume the inexact solutions are obtained using a solver with the Petrov-Galerkin framework. Define residual matrices*

$$\mathbf{R}^{\mathbf{b}} = [\delta\mathbf{b}_1,\, \delta\mathbf{b}_2,\, \ldots \delta\mathbf{b}_r] \qquad \mathbf{R}^{\mathbf{c}} = [\delta\mathbf{c}_1,\, \delta\mathbf{c}_2,\, \ldots \delta\mathbf{c}_r]$$

and backward error

$$\mathbf{E}_{2r} = \mathbf{R}_{\mathbf{b}}(\widetilde{\mathbf{W}}_r^*\widetilde{\mathbf{V}}_r)^{-1}\widetilde{\mathbf{W}}_r^* + \widetilde{\mathbf{V}}_r(\widetilde{\mathbf{W}}_r^*\widetilde{\mathbf{V}}_r)^{-1}\mathbf{R}_{\mathbf{c}}^*$$

then $\widetilde{\mathbf{H}}_r(s)$ *interpolates a perturbed dynamical system,*

$$\widetilde{\mathbf{H}}(s) = \mathbf{c}^T(s\mathbf{I} - [\mathbf{A} + \mathbf{E}_{2r}])^{-1}\mathbf{b}, \quad \text{at} \quad s = \sigma_1, \ldots, \sigma_r.$$

Since $\widetilde{\mathbf{H}}_r(s)$ interpolates $\widetilde{\mathbf{H}}$ for all of the interpolation points, this is a more attractive backward error result, suggesting that BiCG has a distinct advantage over solvers, such as GMRES, which do not have this Petrov-Galerkin framework [3].

4.2 Numerical Results

Throughout the remainder of this chapter, we present our numerical results. For the Rail 1357, Rail 20209 and CD models, we performed one step of Krylov reduction using GMRES and BiCG as the iterative solver. For each iterative method, we considered varying the inexact solve's stopping tolerance from 1×10^{-1} to 1×10^{-10}, preconditioning the linear systems, and using both a good and poor shift selection.

4.2.1 Relative Stopping Residuals

Examining the relative stopping residuals of the inexact solves, we observed another benefit associated with using BiCG as the inexact solver. In our previous discussion of iterative methods, we noted that BiCG solves the systems

$$(\sigma_j\mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b} \tag{4.2.1}$$

$$(\sigma_j\mathbf{I} - \mathbf{A})^*\mathbf{y}_j = \mathbf{c} \tag{4.2.2}$$

simultaneously whereas GMRES solves these systems independently. As a result, the stopping criteria for BiCG requires that both relative residuals associated with (4.2.1) and (4.2.2) be simultaneously less than the desired tolerance. Meanwhile, GMRES only requires that the relative residual associated with the particular linear system being solved is less than the desired tolerance. These different requirements often mean that the relative stopping residuals for GMRES are just below the desired tolerance whereas BiCG's relative stopping residuals are slightly smaller and sometimes even one to two orders smaller than those of GMRES. For example, this phenomenon is observed in Table 4.1 and 4.2.

Table 4.1: Relative Stopping Residuals for $(\sigma_j \mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$; Rail 1357; $\varepsilon = 1 \times 10^{-6}$

Shift	GMRES	BiCG
σ_1	9.12×10^{-7}	2.56×10^{-7}
σ_2	9.38×10^{-7}	6.13×10^{-7}
σ_3	8.67×10^{-7}	7.28×10^{-9}
σ_4	9.39×10^{-7}	2.29×10^{-7}
σ_5	9.91×10^{-7}	9.91×10^{-7}
σ_6	9.93×10^{-7}	6.24×10^{-7}

Table 4.2: Relative Stopping Residuals for $(\sigma_j \mathbf{I} - \mathbf{A})^* \mathbf{y}_j = \mathbf{c}$; Rail 1357; $\varepsilon = 1 \times 10^{-6}$

Shift	GMRES	BiCG
σ_1	8.12×10^{-7}	7.22×10^{-7}
σ_2	8.75×10^{-7}	9.44×10^{-7}
σ_3	7.17×10^{-7}	1.65×10^{-8}
σ_4	8.63×10^{-7}	2.09×10^{-7}
σ_5	9.76×10^{-7}	3.95×10^{-7}
σ_6	9.72×10^{-7}	4.15×10^{-8}

Moreover, this increased accuracy results in the angle between the subspaces formed by the inexact and exact Krylov matrices to be smaller in the case of BiCG. Therefore, BiCG provides increased accuracy over GMRES as exemplified by Table 4.3.

Table 4.3: Subspace Angles for Rail 1357; r = 6; $\varepsilon = 1 \times 10^{-6}$

	GMRES	BiCG
$\sin(\theta(\mathbf{V}_r, \widetilde{\mathbf{V}}_r))$	1.77×10^{-5}	6.10×10^{-6}
$\sin(\theta(\mathbf{W}_r, \widetilde{\mathbf{W}}_r))$	1.56×10^{-4}	2.57×10^{-6}

4.2.2 Relative Interpolation Errors

The impact of the increased accuracy associated with BiCG is observed in the relative interpolation error. For example, Table 4.4 depicts BiCG as giving smaller interpolation errors for all interpolation points. In general, we observed this pattern throughout our numerical simulations, suggesting another distinct advantage associated with using BiCG as the iterative solver.

Table 4.4: $\frac{|H(\sigma_i)-\widetilde{H}_r(\sigma_i)|}{|H(\sigma_i)|}$ Krylov Reduction for Rail 1357; $\varepsilon = 1 \times 10^{-6}$

Shift	GMRES	BiCG
σ_1	1.63×10^{-13}	3.38×10^{-14}
σ_2	4.23×10^{-13}	1.46×10^{-14}
σ_3	1.00×10^{-12}	8.55×10^{-16}
σ_4	3.28×10^{-14}	5.12×10^{-15}
σ_5	2.57×10^{-11}	4.37×10^{-13}
σ_6	9.47×10^{-10}	5.12×10^{-14}

4.2.3 Model Reduction Error

In our previous discussion regarding the effect of tolerance, we noted that for larger tolerances the $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ errors were larger. Decreasing the tolerance improved this overall error for the larger tolerances; however, we observed no improvement in $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ once $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$ was of the same or smaller order as $\|H - H_r\|_{\mathcal{H}_2}$ even though the linear

systems were being solved more accurately. While these patterns hold for both GMRES and BiCG, it is important to note that the increased accuracy associated with BiCG results in smaller error norms for a fixed tolerance as typified in Tables 4.5 and 4.6. Since smaller error norms were persistently associated with BiCG, this suggests yet another advantage of using BiCG in the model reduction setting.

Table 4.5: Model Reduction Error; Krylov Reduction for Rail 1357; $\varepsilon = 1 \times 10^{-6}$

	GMRES	BiCG
$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$	2.17×10^{-9}	9.38×10^{-11}
$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}$	2.17×10^{-7}	1.25×10^{-9}

Table 4.6: $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$; Krylov Reduction for Rail 1357

Tolerance	GMRES	BiCG
1×10^{-1}	7.37×10^{-3}	2.58×10^{-4}
1×10^{-2}	6.30×10^{-4}	1.30×10^{-4}
1×10^{-3}	2.10×10^{-4}	1.36×10^{-4}
1×10^{-4}	1.30×10^{-4}	1.36×10^{-4}
1×10^{-5}	1.36×10^{-4}	1.36×10^{-4}
1×10^{-6}	1.36×10^{-4}	1.36×10^{-4}
1×10^{-7}	1.36×10^{-4}	1.36×10^{-4}
1×10^{-8}	1.36×10^{-4}	1.36×10^{-4}
1×10^{-9}	1.36×10^{-4}	1.36×10^{-4}
1×10^{-10}	1.36×10^{-4}	1.36×10^{-4}

In conclusion, there exists notable advantages with using BiCG in the model reduction setting. From a theoretical perspective, the embedded Petrov-Galerkin structure of BiCG allows for a more satisfactory backward error result. Furthermore, BiCG produces smaller relative stopping residuals, which tend to result in smaller relative interpolation errors and model reduction errors when compared to GMRES. Especially since BiCG allows for the simultaneous solution of the linear systems of interest, BiCG also results in faster convergence

when compared to GMRES. Of course, the main disadvantage of BiCG, namely its potential to breakdown, remains a concern; however, our numerical results illustrate that this concern is often eliminated by simply preconditioning the linear systems. Therefore, in the one step Krylov reduction setting, we conclude that BiCG is the more appropriate inexact solver to implement.

Chapter 5

IRKA with Inexact Solves

Recalling algorithm 1.6.1, IRKA iterates the Krylov reduction process. We will be employing inexact solves in IRKA; therefore, a potential concern is that the inexact solve errors will aggregate to the point where the one step Krylov reduction results will no longer be pertinent. Fortunately, our numerical results suggest that IRKA behaves in a similar manner as just one step of Krylov reduction. The data in this section was obtained by performing the IRKA algorithm for the Rail 1357, Rail 20209 and CD models. For each of these models, we studied both a good and poor shift selection. Also, for each model and shift selection, we considered tolerances 1×10^{-1} to 1×10^{-10}, preconditioned and unpreconditioned linear systems as well as GMRES and BiCG for the inexact solver.

5.1 Tolerance

In the one step case, we observe no noticeable improvement in $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ once $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$ became of the same or smaller order as $\|H - H_r\|_{\mathcal{H}_2}$ even though the linear systems were being solved more accurately. With IRKA, we continue to see this pattern as exemplified by

Tables 5.1 and 5.2.

Table 5.1: Krylov Reduction for CD Model; r = 14; BiCG; prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$	$\|H - H_r\|_{\mathcal{H}_2}$	$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$
1×10^{-1}	5.08×10^{-1}	4.74×10^{-1}	2.73×10^{-1}
1×10^{-2}	5.08×10^{-1}	4.74×10^{-1}	2.73×10^{-1}
1×10^{-3}	5.08×10^{-1}	4.74×10^{-1}	2.73×10^{-1}
1×10^{-4}	5.08×10^{-1}	4.74×10^{-1}	2.73×10^{-1}
1×10^{-5}	4.74×10^{-1}	4.74×10^{-1}	1.62×10^{-6}
1×10^{-6}	4.74×10^{-1}	4.74×10^{-1}	1.50×10^{-6}
1×10^{-7}	4.74×10^{-1}	4.74×10^{-1}	1.98×10^{-6}
1×10^{-8}	4.74×10^{-1}	4.74×10^{-1}	1.98×10^{-6}
1×10^{-9}	4.74×10^{-1}	4.74×10^{-1}	1.54×10^{-10}
1×10^{-10}	4.74×10^{-1}	4.74×10^{-1}	4.62×10^{-11}

Table 5.2: Krylov Reduction for Rail 1357; r = 6; GMRES; prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$	$\|H - H_r\|_{\mathcal{H}_2}$	$\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$
1×10^{-1}	∞	4.30×10^{-6}	∞
1×10^{-2}	5.71×10^{-6}	4.30×10^{-6}	2.91×10^{-6}
1×10^{-3}	4.32×10^{-6}	4.30×10^{-6}	1.35×10^{-7}
1×10^{-4}	4.30×10^{-6}	4.30×10^{-6}	2.14×10^{-8}
1×10^{-5}	4.30×10^{-6}	4.30×10^{-6}	1.12×10^{-9}
1×10^{-6}	4.30×10^{-6}	4.30×10^{-6}	1.09×10^{-10}
1×10^{-7}	4.30×10^{-6}	4.30×10^{-6}	1.38×10^{-11}
1×10^{-8}	4.30×10^{-6}	4.30×10^{-6}	1.24×10^{-12}
1×10^{-9}	4.30×10^{-6}	4.30×10^{-6}	2.25×10^{-13}
1×10^{-10}	4.30×10^{-6}	4.30×10^{-6}	8.76×10^{-15}

Note that H_r is the optimal reduced order model; as a result, the small $\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}$ errors imply we are able to preserve optimality of the reduced order model. Also, the noteworthy conclusion from the stagnation of error norms is that we do not need to expend the additional work associated with solving the linear systems more accurately. In fact, for all of our models and shift selections, we found that after a tolerance of 1×10^{-4}, the overall model reduction

error remained of the same order, implying flexibility in using larger tolerances for the iterative solves.

5.2 Shift Selection

As with just one step of Krylov reduction, we continue to see linear improvement in $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ and $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ when a good shift is used. This phenomenon is typified by Tables 5.3 and 5.4. With one step of Krylov reduction, we did not observe a linear improvement with a poor shift selection, and we continue to encounter examples with IRKA of when the poor shift selection still resulted in no linear improvement as shown in Table 5.3.

However, we also observed cases, such as Table 5.4, with a roughly linear improvement in these errors when a poor initial shift selection was used in IRKA. This linear improvement for some poor shift selections is an added feature of IRKA.

Table 5.3: $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ for BiCG; Rail 1357; prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	4.71×10^{-3}	1.02×10^{-3}
1×10^{-2}	4.14×10^{-4}	4.83×10^{-2}
1×10^{-3}	3.66×10^{-6}	4.83×10^{-2}
1×10^{-4}	1.40×10^{-5}	4.83×10^{-2}
1×10^{-5}	7.11×10^{-7}	4.83×10^{-2}
1×10^{-6}	9.66×10^{-8}	4.83×10^{-2}
1×10^{-7}	9.77×10^{-9}	4.83×10^{-2}
1×10^{-8}	6.92×10^{-10}	4.83×10^{-2}
1×10^{-9}	1.45×10^{-10}	4.83×10^{-2}

Overall, we continue to observe that a good shift selection results in smaller $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ and $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ errors as shown by Tables 5.3 and 5.4. Moreover, we see that a good shift

Table 5.4: $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ for BiCG; Rail 20209; prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	2.40×10^{-2}	4.98×10^{-2}
1×10^{-2}	2.53×10^{-3}	1.78×10^{-2}
1×10^{-3}	1.46×10^{-4}	4.13×10^{-5}
1×10^{-4}	7.40×10^{-6}	9.97×10^{-6}
1×10^{-5}	7.08×10^{-7}	6.65×10^{-7}
1×10^{-6}	1.42×10^{-7}	3.37×10^{-8}
1×10^{-7}	5.69×10^{-9}	6.14×10^{-9}
1×10^{-8}	1.76×10^{-9}	2.21×10^{-10}
1×10^{-9}	5.27×10^{-11}	4.00×10^{-11}

selection results in smaller $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$errors overall, which was also an observation in the one step case. An example of this pattern is found in Table 5.5. Therefore,

Table 5.5: $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ for BiCG; Rail 1357; prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	1.43×10^{-4}	1.62×10^{-4}
1×10^{-2}	1.44×10^{-4}	1.69×10^{-2}
1×10^{-3}	1.45×10^{-4}	1.69×10^{-2}
1×10^{-4}	1.45×10^{-4}	1.70×10^{-2}
1×10^{-5}	1.45×10^{-4}	1.69×10^{-2}
1×10^{-6}	1.45×10^{-4}	1.68×10^{-2}
1×10^{-7}	1.45×10^{-4}	1.69×10^{-2}
1×10^{-8}	1.45×10^{-4}	1.69×10^{-2}
1×10^{-9}	1.45×10^{-4}	1.69×10^{-2}

our conclusion is that our observations from the one step case remain valid with IRKA. Since we are iterating the Krylov reduction process, an additional concern is the relationship between shift selection and convergence. In our numerical simulations, we observed different convergence results depending on the shift selection. These patterns are typified by Table 5.6.

Table 5.6: IRKA Iterations for BiCG; Rail 20209; no prec.

Tolerance	Good Shift	Poor Shift
1×10^{-1}	–	–
1×10^{-2}	19	36
1×10^{-3}	23	–
1×10^{-4}	24	–
1×10^{-5}	24	145
1×10^{-6}	24	59
1×10^{-7}	24	–
1×10^{-8}	24	83
1×10^{-9}	24	148

In general, IRKA may fail to converge for large tolerances; however, once it converged for a particular tolerance it would also converge for all smaller tolerances provided that a good initial shift was used. Unfortunately, erratic convergence would often be associated with a poor shift selection. This has an important implication for determining the appropriate tolerance. The predictable behavior associated with a good shift selection implies that we can solve the systems with a relatively large tolerance. However, the erratic behavior associated with a poor shift selection implies that we do not necessarily have this flexibility in choosing larger tolerances and that the iteration may actually suffer when the tolerance is decreased.

5.3 GMRES vs. BiCG

Fortunately, the conclusions from the one step Krylov reduction case regarding GMRES and BiCG remain germane with IRKA, since we continue to see the superiority of BiCG. As in the one step Krylov reduction case, we noted smaller errors between the inexact and exact models when BiCG was used. This pattern is typified by Tables 5.7 and 5.8.

Table 5.7: $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ for CD Model; r = 6; prec.

Tolerance	GMRES	BiCG
1×10^{-1}	4.32×10^{-3}	8.11×10^{-8}
1×10^{-2}	5.57×10^{-4}	7.34×10^{-8}
1×10^{-3}	1.27×10^{-5}	7.34×10^{-8}
1×10^{-4}	1.17×10^{-6}	5.43×10^{-8}
1×10^{-5}	2.48×10^{-7}	7.94×10^{-9}
1×10^{-6}	2.33×10^{-8}	3.78×10^{-11}
1×10^{-7}	2.65×10^{-10}	1.45×10^{-12}
1×10^{-8}	3.54×10^{-11}	1.44×10^{-12}
1×10^{-9}	1.43×10^{-11}	2.34×10^{-14}
1×10^{-10}	1.44×10^{-12}	2.39×10^{-14}

Also, we continue to observe that BiCG produces smaller or equivalent $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ errors. For example, in Table 5.9, we see that GMRES produced an unbounded error whereas the error remained bounded with BiCG.

Therefore, BiCG continues to offer additional advantages over GMRES, leading to the conclusion that BiCG is the iterative method of choice in the IRKA setting as well.

5.4 Using an initial guess in the iterative solver

In this section, we will let $\sigma_j^{(k-1)}$ and $\sigma_j^{(k)}$ be the j^{th} shift at iteration $k-1$ and k, respectively. Also, $\mathbf{x}_j^{(k-1)}$ and $\mathbf{y}_j^{(k-1)}$ will be the solutions of $(\sigma_j^{(k-1)}\mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ and $(\sigma_j^{(k-1)}\mathbf{I} - \mathbf{A})^*\mathbf{y}_j = \mathbf{c}$, respectively. Similarly, $\mathbf{x}_j^{(k)}$ and $\mathbf{y}_j^{(k)}$ will be the solutions of $(\sigma_j^{(k)}\mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ and $(\sigma_j^{(k)}\mathbf{I} - \mathbf{A})^*\mathbf{y}_j = \mathbf{c}$, respectively. Since we are concerned with solving $(\sigma_j^{(k)}\mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ and $(\sigma_j^{(k)}\mathbf{I} - \mathbf{A})^*\mathbf{y}_j = \mathbf{c}$, we have that if the difference between $\sigma_j^{(k)}$ and $\sigma_j^{(k-1)}$ is small, then it follows that the solution $\mathbf{x}_j^{(k)}$ or $\mathbf{y}_j^{(k)}$ should be close to $\mathbf{x}_j^{(k-1)}$ or $\mathbf{y}_j^{(k-1)}$, respectively. As a result, we anticipate less work if we use $\mathbf{x}_j^{(k-1)}$ or $\mathbf{y}_j^{(k-1)}$ as an initial guess for the iterative

Table 5.8: $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ for CD Model; r = 6; prec.

Tolerance	GMRES	BiCG
1×10^{-1}	2.96×10^{-3}	5.65×10^{-8}
1×10^{-2}	4.30×10^{-4}	5.54×10^{-8}
1×10^{-3}	1.03×10^{-5}	5.54×10^{-8}
1×10^{-4}	9.35×10^{-7}	4.55×10^{-8}
1×10^{-5}	2.72×10^{-7}	8.70×10^{-9}
1×10^{-6}	2.55×10^{-8}	3.99×10^{-11}
1×10^{-7}	2.40×10^{-10}	1.58×10^{-12}
1×10^{-8}	2.76×10^{-11}	1.57×10^{-12}
1×10^{-9}	1.57×10^{-11}	1.99×10^{-14}
1×10^{-10}	1.57×10^{-12}	7.94×10^{-15}

solve of $(\sigma_j^{(k)}\mathbf{I} - \mathbf{A})\mathbf{x}_j = \mathbf{b}$ or $(\sigma_j^{(k)}\mathbf{I} - \mathbf{A})^*\mathbf{y}_j = \mathbf{c}$, respectively. In our numerical results, we observed this benefit when using an initial guess. In Figure 5.1, for example, the norm of the relative difference between the shifts is less than 10^{-1} for most of the iteration. This small difference between successive shifts implies that the solutions for the corresponding linear systems will be similar. As a result, we expect that using an initial guess will result in fewer iterations. As Table 5.10 illustrates, the use of an initial guess does in fact decrease the number of BiCG iterations for almost all tolerances.

Unfortunately, the shifts do not converge smoothly with a poor initial shift selection. For example, the shift evolution of the Rail 20209 model when using BiCG, a poor initial guess, and a tolerance of 1×10^{-3} was characterized by erratic behavior as shown in Figure 5.2.

Due to this uneven convergence of the shifts, we might expect that $\mathbf{x}_j^{(k-1)}$ or $\mathbf{y}_j^{(k-1)}$ will not be a good initial guess for $\mathbf{x}_j^{(k)}$ or $\mathbf{y}_j^{(k)}$. Nevertheless, our numerical results suggest there still is a benefit in using the previous solution as an initial guess. In Table 5.11, for example, we observe notable advantages for most tolerances in using an initial guess even though the shift evolutions for each tolerance are erratic.

Figure 5.1: Shift Evolution of CD Model using BiCG

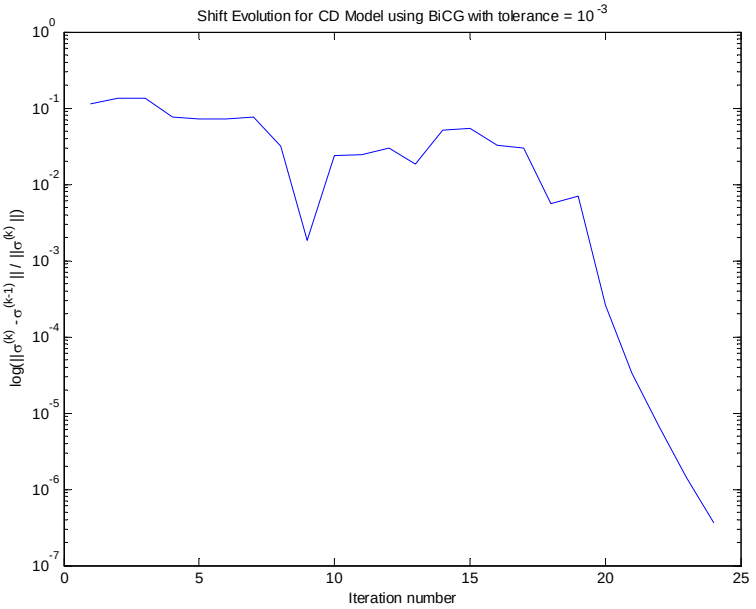


Figure 5.2: Shift Evolution of Rail 20209 using BiCG

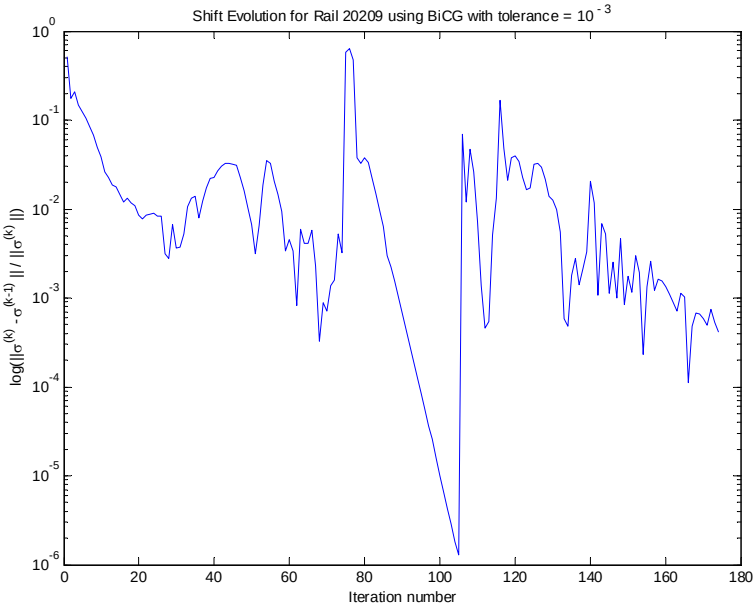


Table 5.9: $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ for Rail 1357; r = 6; prec.

Tolerance	GMRES	BiCG
1×10^{-1}	∞	6.55×10^{-6}
1×10^{-2}	5.71×10^{-6}	4.33×10^{-6}
1×10^{-3}	4.32×10^{-6}	4.31×10^{-6}
1×10^{-4}	4.31×10^{-6}	4.31×10^{-6}
1×10^{-5}	4.31×10^{-6}	4.31×10^{-6}
1×10^{-6}	4.31×10^{-6}	4.31×10^{-6}
1×10^{-7}	4.31×10^{-6}	4.31×10^{-6}
1×10^{-8}	4.31×10^{-6}	4.31×10^{-6}
1×10^{-9}	4.31×10^{-6}	4.31×10^{-6}
1×10^{-10}	4.31×10^{-6}	4.31×10^{-6}

Using an initial guess also tends to improve the accuracy with which we solve the systems. Table 5.12 reports the difference between the stopping residuals associated with using an initial guess and using just the zero vector for the first iterate, and it illustrates that using an initial guess resulted in a smaller stopping residual for most shifts. In general, we found improved accuracy when an initial guess was used.

Although the difference in stopping residuals was often small, this slight improvement in accuracy seems to impact the convergence of IRKA. In our iterations, we observed that using an initial guess tended to result in fewer IRKA iterations. Table 5.13 illustrates the number of IRKA iterations when no guess and an initial guess were used for inexact solves with tolerances 1×10^{-1} to 1×10^{-10}. Especially for the larger tolerances, the use of an initial guess in the inexact solve seems to result in substantially fewer IRKA iterations. Of course, it is important to note that sometimes using an initial guess results in additional IRKA iterations. In these cases, however, an initial guess often resulted in only a couple of additional iterations.

In general, an initial guess resulted in the IRKA iteration taking fewer iterations. Also,

Table 5.10: CD Model; r = 14; prec.; BiCG

Tolerance	Percentage Decrease of Iterations when an Initial Guess is used
1×10^{-1}	28%
1×10^{-2}	24%
1×10^{-3}	39%
1×10^{-4}	-2%
1×10^{-5}	11%
1×10^{-6}	15%
1×10^{-7}	16%
1×10^{-8}	13%
1×10^{-9}	11%
1×10^{-10}	15%

from Table 5.14, it is important to note that using an initial guess resulted in the iteration converging for shift selections and tolerances when it did not converge if no initial guess was used in the inexact solver. Therefore, the data suggests there are benefits beyond just faster convergence associated with using an initial guess in the iterative solve.

5.5 Tolerance

We also considered varying the tolerance of the inexact solve throughout the iteration. In optimization, trust-region methods begin with a larger tolerance and then decrease the tolerance as the approximation becomes closer to the solution. Using a similar idea, we considered two variations of tolerances. For the HI method, we increased the tolerance as the relative difference between the shifts became smaller. In the other method, HD, we decreased the tolerance as the relative difference between the shifts became smaller. The pseudocode for these methods may be found in the appendix.

For the Rail 1357 and CD models, we reduced to an order of six using both a good and poor

Table 5.11: Rail 20209 using BiCG, r = 6

Tolerance	Percentage Decrease of Iterations when an Initial Guess is used
1×10^{-1}	99%
1×10^{-2}	68%
1×10^{-3}	97%
1×10^{-4}	88%
1×10^{-5}	74%
1×10^{-6}	-34%
1×10^{-7}	61%
1×10^{-8}	30%
1×10^{-9}	100%
1×10^{-10}	15%

shift selection. In each case, we considered fixed tolerances 1×10^{-1}, 1×10^{-3}, and 1×10^{-5}, as well as the HI and HD methods. In general, we observed that the HI and HD methods produced $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ and $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ of the same order or one to two orders larger as when a fixed tolerance was used. An example of the behavior observed is given by Table 5.15 with the corresponding tolerances listed in Table 5.16.

In this particular case, HI and HD methods give an error on the same order as when the tolerance is fixed at 1×10^{-3}. While for the HI method we solve the systems more accurately than 1×10^{-3} for several iterations, the majority of the HD iteration does not solve the systems as accurately as 1×10^{-3}. As a result, varying the tolerance in the HD method allows us to save work by not solving the systems as accurately but still gives us error norms on the same order as if the tolerance had been fixed. In general, we did not find a distinct pattern where one method persistently produced smaller $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ and $\frac{\|H_r - \widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ norms.

One dominant trend we found is that varying the tolerance did not result in $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ being of a different order than the error associated with a fixed tolerance. Of course, this is what we expected since usually these norms are of the same order for nearly all tolerances. An example of this pattern is found in Table 5.17.

Table 5.12: Initial Guess - No Guess Stopping Residuals at the 3rd iteration CD Model; $\varepsilon = 1 \times 10^{-3}$, r = 14 (IRKA with BiCG); $\lambda = 10$, prec.

Linear System	$\mathbf{V}_r$	$\mathbf{W}_r$
σ_1	-6.86×10^{-11}	-4.55×10^{-11}
σ_2	-4.72×10^{-4}	-6.55×10^{-4}
σ_3	-5.18×10^{-4}	-7.24×10^{-4}
σ_4	-2.41×10^{-7}	-6.19×10^{-5}
σ_5	-2.34×10^{-7}	-4.98×10^{-5}
σ_6	-4.14×10^{-5}	1.50×10^{-5}
σ_7	8.77×10^{-8}	8.01×10^{-7}
σ_8	-3.61×10^{-4}	-5.87×10^{-4}
σ_9	0	0
σ_{10}	0	0
σ_{11}	0	0
σ_{12}	0	0
σ_{13}	0	0
σ_{14}	0	0

In conclusion, the IRKA iteration retains the properties of one step Krylov reduction in terms of errors and the additional benefits of BiCG. Further relationships between the shift selection and convergence behavior emerge due to the iterative nature of the algorithm. Finally, we noted the advantages of using a good initial shift selection and varying the inexact solver's tolerance throughout the iteration.

Table 5.13: Total IRKA iterations for the CD Model; r = 14; BiCG; $\lambda = 10$, prec.

Tolerance	No guess	Initial Guess
1×10^{-1}	25	14
1×10^{-2}	21	15
1×10^{-3}	25	15
1×10^{-4}	13	15
1×10^{-5}	20	23
1×10^{-6}	19	20
1×10^{-7}	20	20
1×10^{-8}	20	21
1×10^{-9}	21	21
1×10^{-10}	21	21

Table 5.14: Cases of Convergence for Tolerances 1×10^{-1} through 1×10^{-10}

Method	No guess	Initial guess
BiCG, Rail 1357, logspace$(-5, 0.7, 6)$, no prec.	7	9
GMRES, Rail, logspace$(-5, 0.7, 6)$, no prec.	7	9
BiCG, Rail 20209, logspace$(-5, 0.7, 6)$, no prec.	9	10
GMRES, Rail 20209, logspace$(-5, 0.7, 6)$, no prec.	7	10
BiCG, Rail 1357, logspace$(0.9, 1.3, 6)$, no prec.	6	4
BiCG, Rail 20209, logspace$(0.9, 1.3, 6)$, no prec.	6	10
BiCG, Rail 1357, logspace$(-5, 0.7, 6)$, prec.	7	9
GMRES, Rail, logspace$(-5, 0.7, 6)$, prec.	7	9
BiCG, Rail 20209, logspace$(-5, 0.7, 6)$, prec.	9	10
GMRES, Rail 20209, logspace$(-5, 0.7, 6)$, prec.	7	10
BiCG, Rail 1357, logspace$(0.9, 1.3, 6)$, prec.	6	4
BiCG, Rail 20209, logspace$(0.9, 1.3, 6)$, prec.	6	10
BiCG, CD Model, r = 6, $\lambda = 10$, prec.	10	10
BiCG, CD Model, r = 14, $\lambda = 10$, prec.	10	10

Table 5.15: $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ for Rail 1357 (BiCG); prec.

Tolerance	$\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$
1×10^{-1}	4.70×10^{-3}
1×10^{-3}	6.34×10^{-5}
1×10^{-5}	4.73×10^{-7}
HI	4.1×10^{-5}
HD	7.58×10^{-5}

Table 5.16: Tolerances for Rail 1357; r = 6 (BiCG); prec.

Tolerances for HI	Tolerances for HD
1×10^{-5}	1×10^{-1}
1×10^{-5}	1×10^{-1}
1×10^{-5}	1×10^{-1}
1×10^{-4}	1×10^{-2}
1×10^{-3}	1×10^{-2}
1×10^{-3}	1×10^{-3}
–	1×10^{-3}
–	1×10^{-3}

Table 5.17: $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ for Rail 1357; r = 6 (BiCG); prec.

Tolerance	$\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$
1×10^{-1}	6.55×10^{-6}
1×10^{-3}	4.31×10^{-6}
1×10^{-5}	4.31×10^{-6}
HI	4.31×10^{-6}
HD	4.31×10^{-6}

Chapter 6

Conclusions

Interpolatory model reduction finds a reduced order model that interpolates the full order model and a certain number of its derivatives at selected interpolation points. In the process, linear systems are solved. Especially in the large-scale setting, inexact solvers become necessary for these linear systems. We investigated inexact solves in the interpolatory model reduction setting by studying the Rail 1357, Rail 20209 and CD models. We considered using GMRES and BiCG, preconditioning the linear systems, varying the stopping tolerance from 1×10^{-1} to 1×10^{-10}, and using different shift selections.

For just one step of Krylov reduction, we verified the quadratic and linear behaviors of the interpolation error for the one-sided and two-sided cases. We also noted that $\|H - \widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H - \widetilde{H}_r\|_{\mathcal{H}_\infty}$ remained of the same order for most tolerances selected, implying flexibility in choosing larger stopping tolerances for the inexact solve without negatively impacting the overall model reduction error. For just one-step of Krylov reduction, we investigated the relationship between the shift selection and tolerance. We found a linear improvement in the $\sin(\Theta(\mathbf{V}_r, \widetilde{\mathbf{V}}_r))$ and $\sin(\Theta(\mathbf{W}_r, \widetilde{\mathbf{W}}_r))$ provided that a good shift selection was used. For a poor shift selection, we observed these subspace angles typically remained of the same order

regardless of how accurately the linear systems were solved. This pattern was then reflected in the overall error norms; $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ and $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ improved linearly with a good shift selection and usually were of the same order regardless of tolerance with a poor shift selection. $\|H-\widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H-\widetilde{H}_r\|_{\mathcal{H}_\infty}$ were often several orders larger when a poor shift selection was used. Also, our numerical results suggest that the tolerance of the inexact solve rather than the shift selection impacts the interpolation error. For both preconditioned and BiCG cases, we observed smaller $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_2}}{\|H_r\|_{\mathcal{H}_2}}$ and $\frac{\|H_r-\widetilde{H}_r\|_{\mathcal{H}_\infty}}{\|H_r\|_{\mathcal{H}_\infty}}$ norms but no significant improvement in $\|H-\widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H-\widetilde{H}_r\|_{\mathcal{H}_\infty}$.

With IRKA, we considered the same combinations of inexact solvers as we did for one step of Krylov reduction. Fortunately, the conclusions for just one step of Krylov reduction also hold with this iteration of the Krylov reduction process. Moreover, we found additional benefits for accuracy and IRKA convergence when the solutions from the previous iteration's linear systems were used as an initial guess for the current iteration's linear systems. Finally, we considered varying the tolerance and found no significant impact on the $\|H-\widetilde{H}_r\|_{\mathcal{H}_2}$ and $\|H-\widetilde{H}_r\|_{\mathcal{H}_\infty}$ norms.

Chapter 7

Appendix

Algorithm 7.0.1. HI: Increasing tolerance as $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|}$ decreases

if $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1 \times 10^{-3}$

$tol = 1 \times 10^{-1}$;

elseif $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1 \times 10^{-2}$

$tol = 1 \times 10^{-2}$;

elseif $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1 \times 10^{-1}$

$tol = 1 \times 10^{-3}$;

elseif $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1$

$tol = 1 \times 10^{-4}$;

else

$tol = 1 \times 10^{-5}$;

end

Algorithm 7.0.2. HD: Decreasing tolerance as $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|}$ decreases

if $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1 \times 10^{-3}$

$tol = 1 \times 10^{-5}$*;*

elseif $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1 \times 10^{-2}$

$tol = 1 \times 10^{-4}$*;*

elseif $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1 \times 10^{-1}$

$tol = 1 \times 10^{-3}$*;*

elseif $\frac{\|\boldsymbol{\sigma}_i - \boldsymbol{\sigma}_{i-1}\|}{\|\boldsymbol{\sigma}_i\|} < 1$

$tol = 1 \times 10^{-2}$*;*

else

$tol = 1 \times 10^{-1}$*;*

end

Chapter 8

Bibliography

Bibliography

[1] A.C. Antoulas, *Lectures on the approximation of linear dynamical systems*, Advances in Design and Control, SIAM, Philadelphia, 2004.

[2] Richard Barrett, Michael Berry, Tony F. Chan, James Demmel, et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, Philadelphia, 1994.

[3] Chris Beattie, *Structured Perturbations in Rational Krylov Methods for Model Reduction*, Workshop on Structured Perturbations and Distance Problems in Matrix Computations, Bedlewo, Poland, March 2007.

[4] C.A. Beattie and S. Gugercin, *Inexact Solves in Krylov-based Model Reduction*, in Proceedings of the 45th IEEE Conference on Decision and Control, 2006.

[5] P. Benner, *Solving Large-Scale Control Problems*, IEEE Control Systems Magazine, Vol. 24, No. 1, pp. 44-59, 2004.

[6] P. Benner and J. Saak, *Efficient numerical solution of the LQR-problem for the heat equation*, Proc. Appl. Math. Mech., 4, pp. 648-649, 2004.

[7] A.E. Bryson and A. Carrier, *Second-order algorithm for optimal model order reduction*, J. Guidance Contr. Dynam., pp. 887-892, 1990.

[8] E. de Sturler and H. A. van der Vorst, *Reducing the effect of global communication in GMRES(m) and CG on parallel distributed memory computers*, Applied Numerical Mathematics (IMACS), Vol. 18, 441-459, 1995.

[9] C. De Villemagne and R. Skelton, *Model reduction using a projection formulation*, Internat. J. Control, 40, pp. 2141-2169, 1987.

[10] P. Feldman and R.W. Freud, *Efficient linear circuit analysis by Padé approximation via a Lanczos method*, IEEE Trans. Computer-Aided Design, 14, pg. 639-649, 1995.

[11] Anne Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.

[12] E.J. Grimme, *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, University of Illinois, Urbana-Champaign, Urbana, IL, 1997.

[13] S. Gugercin, *Projection Methods for Model Reduction of Large-Scale Dynamical Systems*, Ph.D. thesis, Rice University, Houston, TX, 2002.

[14] S. Gugercin and A.C. Antoulas, *An $\mathcal{H}_2$ error expression for the Lanczos procedure*, in Proceedings of the 42nd IEEE Conference on Decision and Control, 2003, pp. 1869-1872.

[15] S. Gugercin, A.C. Antoulas and C.A. Beattie, *$\mathcal{H}_2$ Model Reduction for Large-Scale Linear Dynamical Systems*, SIAM Journal on Matrix Analysis and Applications, Vol.30, Issue: 2, pp. 609-638, 2008.

[16] N. Higham, *Perturbation theory and backward error for $AX - XB = C^*$*, BIT, Vol. 22, pp. 124-136, 1993.

[17] D.C. Hyland and D.S. Bernstein, *The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton and Moore*, IEE. Trans. Automa. Contr., Vol. 30, No. 12, pp. 1201-1211, 1985.

[18] Werner A. Kohler and Lee W. Johnson, *Elementary Differential Equations with Boundary Value Problems*, Addison-Wesley; 1st edition, 2004.

[19] L. Meier and D.G. Luenberger, *Approximation of Linear Constant Systems*, IEE. Trans. Automat. Contr., Vol. 12, pp. 585-588, 1967.

[20] Yousef Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2004.

[21] J.T. Spanos, M.H. Millman, and D.L. Mingori, *A new algorithm for L_2 optimal model reduction*, Automatics, pp. 897-909, 1992.

[22] Lloyd N. Trefethen and David Bau, III, *Numerical Linear Algebra*, SIAM, 1997.

[23] D.A. Wilson, *Optimum solution of model reduction problem*, in Proc. Inst. Elec. Eng., pp. 1161-1165, 1970.

[24] W-Y. Yan and J. Lam, *An approximate approach to $\mathcal{H}_2$ optimal model reduction*, IEEE Transactions on Automatic Control, AC-44, pp. 1341-1358, 1999.

[25] A. Yousouff and R.E. Skelton, *Covariance equivalent realizations with applications to model reduction of large-scale systems*, in Control and Dynamic Systems, Vol.22, C.T. Leondes, ed., Academic Press, New York, pp. 273-348, 1985.

[26] A. Yousouff, D.A. Wagie, and R.E. Skelton, *Linear system approximation via covariance equivalent realizations*, J. Math. Anal. Appl., 196, pp. 91-115, 1985.