SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Model reduction for piezo-mechanical systems using balanced truncation

av

Mohammad Monir Uddin

2011 - No 5

Model reduction for piezo-mechanical systems using balanced truncation

Mohammad Monir Uddin

Självständigt arbete i tillämpad matematik 30 högskolepoäng, avancerad nivå

Handledare: Peter Benner

2011

ABSTRACT

In today's scientific and technological world, physical and artificial processes are often described by mathematical models which can be used for simulation, optimization or control. As the mathematical models get more detailed and different coupling effects are required to include, usually the dimension of these models become very large. Such large-scale systems lead to large memory requirements and computational complexity. To handle these large models efficiently in simulation, control or optimization model order reduction (MOR) is essential. The fundamental idea of model order reduction is to approximate a large-scale model by a reduced model of lower state space dimension that has the same (to the largest possible extent) input-output behavior as the original system. Recently, the system-theoretic method "Balanced Truncation (BT)", which was believed to be applicable only to moderately sized problems, has been adapted to really large-scale problems. Moreover, it also has been extended to so-called descriptor systems, i.e., systems whose dynamics obey differential-algebraic equations. In this thesis, a BT algorithm is developed for MOR of index-1 descriptor systems based on several papers from the literature. It is then applied to the setting of a piezo-mechanical system. The algorithm is verified by real-world data describing micro-mechanical piezo-actuators. Several numerical experiments are used to illustrate the efficiency of the algorithm.

ACKNOWLEDGEMENT

With my great pleasure at first I would like to express my cordial thanks to my supervisor Professor Peter Benner for his remarkable contribution and guidance in this thesis. I am very thankful to him for introducing me in this highly fascinating and applicable research area and for the financial support to complete the thesis.

I would like to express my hearty gratitude to my teacher and advisor Professor Boris Shapiro for his valuable suggestions, essential remarks and kind gesture that he has shown to me during the course of study and completing this thesis work.

I would also like to thank Dr. Jens Saak for his important suggestions, essential discussions and cordial guidance to finish this thesis successfully.

I am thankful to all in the research group of MIIT (Mathematics in Industry and Technology) at TU Chemnitz, Germany, and the research group of CSC (Computational Methods in Systems and Control Theory) in Max Planck Institute for Dynamics of Complex and Technical Systems, Magdeburg, Germany, for their nice cooperation during my thesis work. Specially, I am grateful to Dr. René Schneider for his time to discuss some problems related to my thesis and helping me in administrative issues at the beginning of my work in Germany. I am very thankful to my dear friend and colleague Mohammad-Sahadet Hossain for his inspirations, suggestions on different issues during my stay in Chemnitz. Specially i would like to remember those interesting moments with him when we discussed among many interesting topics, the problems regarding this thesis every single day during the coffee brake.

I would like to thank Dr. Joost Rommes and Dr. B. Kranz for their valuable suggestions. I am specially grateful to them, since Dr. Joost Rommes provided me data of the power system models to check the algorithm initially, on the other hand Dr. B. Kranz provided some interesting pictures along with important information regarding the piezo-mechanical systems.

I acknowledge my gratefulness to Professor Yishao Zhou who taught several courses that helped me a lot in this thesis work. I would also like to thank Professor Andrzej Szulkin for his important suggestions. I am thankful to all my teachers in the Department of Mathematics at Stockholm University and the Department of Scientific Computing at KTH for their generous help during my course work in the Master program.

Finally, I am extremely thankful and indebted to my beloved mother and all of my family members for their inspirations and support in every step of my life.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

CHAPTER

ONE

INTRODUCTION AND OVERVIEW

Mathematical modeling and simulation have become an important part of today's scientific and technological world. For various purposes such as control, analysis, optimization and design of physical systems, simulation of mathematical models play an important role. In order to simulate accurately, when a physical model is converted into a mathematical model, in many cases its dimension becomes extremely large. Such large systems arise in different disciplines including chemical engineering, mechanical and micro-electro-mechanical systems, computational biology, aerospace engineering etc. In order to solve engineering problems dealing with these large-scale systems, enormously expensive computational efforts are needed. Sometimes the simulations, dynamic analysis and design of higher order systems are even impossible due to restrictions caused by computer memory or the applied numerical algorithms. To circumvent the size problems, it is a good idea to reduce the dimension of the mathematical models. The way how a higher dimensional model is reduced to a lower dimensional one is called *model order reduction* (MOR). See [2, 8, 63, 48, 54, 3, 52, 38, 72] to learn about motivations, applications, restrictions and techniques of MOR.

The general idea of model order reduction is to approximate a large-scale model by a reduced model of lower state space dimension, while their input-output behavior is preserved to the largest possible extent. It is a numerical process which possesses the following goals:

- The size of the reduced order model should be very small.
- The reduction process is automatic (the algorithm does not require knowledge about the nature of the underlying systems).

- System properties such as stability and passivity are preserved.
- The algorithm must be efficient.
- The error between original and reduced systems (more precisely their transfer function matrices), measured by some suitable norm, must be as small as possible.

Depending on the characteristics of mathematical models, various techniques as for example optimal Hankel norm approximation [24], singular perturbation approximation [40], moment matching approximation [29, 21], balanced truncation (BT) [43, 64] are commonly used for MOR. Among those techniques BT is one of the well accepted methods for large sparse linear time invariant (LTI) systems. One of the great advantages of this projection based method is that it does preserve the stability of the original systems, i.e., if the original model is stable, then the reduced order model is also stable. Moreover it has a global error bound and by choosing a priori error tolerances one can adapt the dimension of the reduced models. The key idea of this method is to delete the unnecessary states, which can be detected if the system is in balanced coordinates. The basic balancing technique is closely related to the solution of Lyapunov equations. Today, it is well known that among several Lyapunov solvers *alternating direction implicit* (ADI) methods are one of those attractive for very large sparse systems. Again, ADI methods require *ADI shifts parameters*, which is one more computational task.

Piezo-mechanical systems (as in Figure 1.1) considered in this thesis are highly complex systems, where an adaptive spindle support (Figure 1.1(a)) is used for additional fine positioning movements during machining operations (see [16, 44] for more details). In order to analyze (the mechanical design and the performance of) this spindle support regarding control aspects, a mathematical model is formed (see [34]) by using a finite-element (FE)-model (see Figure 1.2). The mathematical model of this adaptive spindle support is a very large multi-input multi-output second order LTI continuous-time system. It is an index-1 descriptor system (see Section 3.3) which is highly sparse. In order to be able to simulate this model efficiently, MOR is unavoidable. To apply the BT technique for MOR of such a system, at first it is converted into first order form which is also an index-1 descriptor system. In this representation the size of the system becomes larger. Under these circumstances, if the sparsity pattern of the original system can be preserved, it will not only consume less memory in storing but also the computation will be extremely fast. From the title and above discussion the motivation of this thesis is now clear. The main objective is to

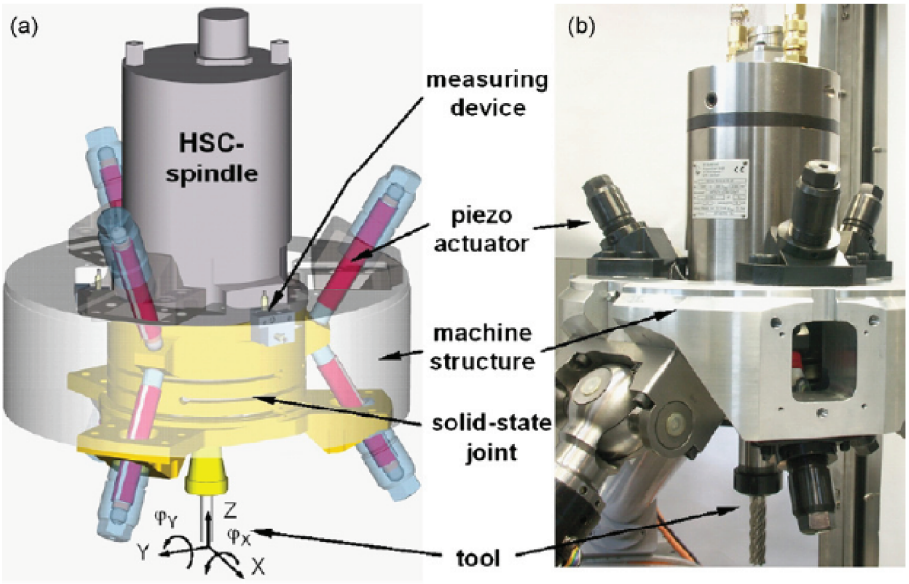


Figure 1.1: The CAD model of the adaptive spindle support (a) and real component (b) mounted in a parallel-kinematic machine (PKM) (image source: [16], with courtesy to Fraunhofer IWU, Dresden, Germany)

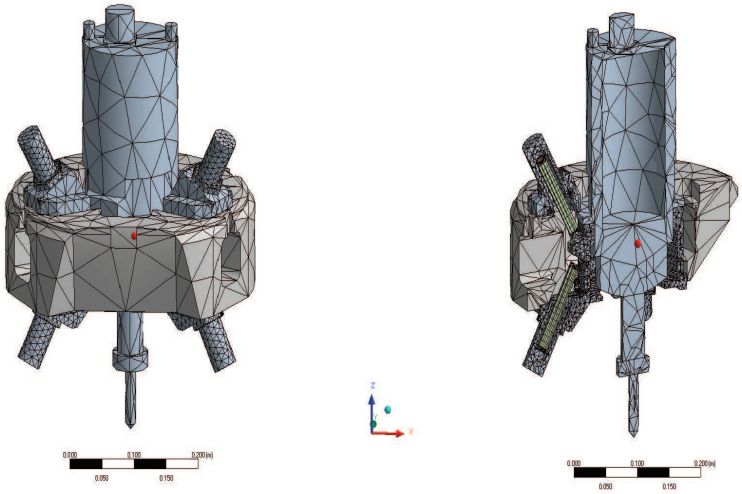


Figure 1.2: Adaptive spindle support and finite element model (image source: [16], with courtesy to Fraunhofer IWU, Dresden, Germany)

develop a BT algorithm for large sparse index-1 descriptor systems with a block-structure as imposed by the piezo-mechanical model to derive small standard state-space systems that can be used within MATLAB/Simulink for control design. The whole MOR process suggested in this thesis is represented by a flowchart given in Figure 1.3.

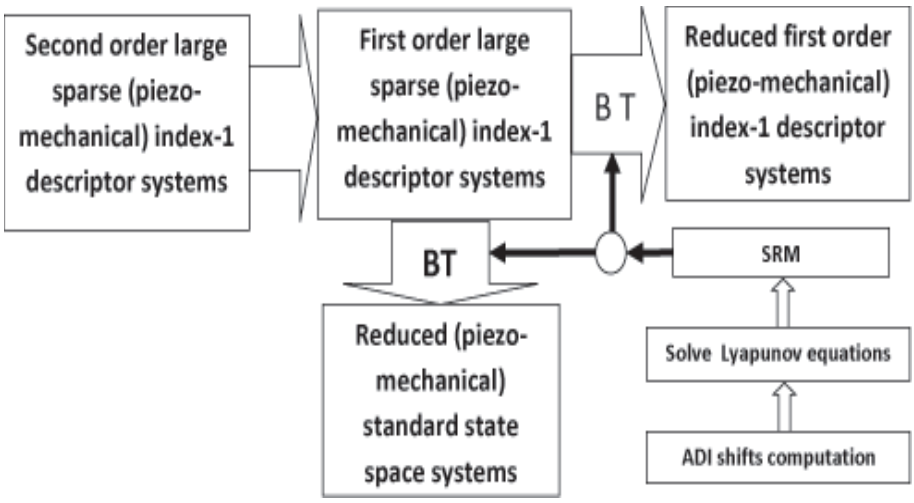


Figure 1.3: Different steps of MOR for piezo-mechanical systems using BT

This thesis consists of 6 chapters including this introductory one. Chapter 2 contains some notations and a review of some fundamental concepts, and results from linear algebra and system theory. The concepts of this chapter are used throughout the thesis. Chapter 3 then introduces the model problems. This chapter concentrates on converting a second order model into a first order model by preserving system characteristics like stability, sparsity etc in order to adapt it for the proposed algorithm. Several techniques in reducing first order representations are discussed here elaborately. The benefits of preserving the sparsity pattern are also shown by graphical illustration.

In Chapter 4 an algorithm, namely the generalized sparse low rank Cholesky factor (GSLRCF-)ADI methods, is introduced to solve the generalized Lyapunov equations for the large sparse index-1 descriptor systems. The related issues such as computation of ADI shift parameters, stopping criteria of the algorithm are included here. The efficiency of the algorithm is illustrated by numerical results at the end of the chapter.

Then Chapter 5 discusses the BT algorithms for MOR of large sparse index-

1 descriptor systems. In this chapter two algorithms are introduced separately. One derives a reduced order model in a descriptor form and another one can be used to find a reduced standard state space model. The algorithms are checked by interesting application data coming from micromechanical piezo-actuators. The numerical experiments are discussed elaborately at the end of this chapter.

Finally, Chapter 6 contains the conclusions of the thesis.

CHAPTER

TWO

PRELIMINARY CONCEPT

In this chapter, we establish some notation and review some fundamental concepts and results that will be used in the later chapters of this thesis. The first section presents basic concepts from system and control theory, emphasizing system representations from different points of view, their classifications, controllability and observability, and transfer functions. The second section contains some important definitions, propositions, and theorems form Linear Algebra and Matrix Theory. For the sake of conciseness, profound discussion of a topic, and proofs of the theorems are omitted, since details are available in the references listed at the end of the thesis.

2.1 System identification and their representations

Generally speaking, systems are devices that take an input, and yield an output. Usually, the word system is associated with dynamics, i.e., with the way a system evolves with time which can be modeled in various ways, such as as continuous and discrete are most common. To describe a system, we focus on the black box approach, given by the diagram in Figure 2.1.

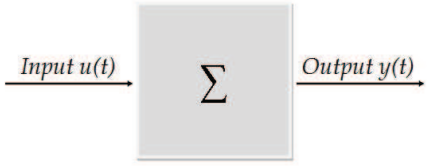


Figure 2.1: Black box

Here $\mathbf{\Sigma}$ denotes the system, $u(t)$ the input or control signal, and $y(t)$ the output signal. The way the output signal depends on the input signals is called the *input-output* relation [22]. Such a description of a system is called an *external representation*. On the other hand an *internal representation* of a system is nothing but a model and can be represented as difference (discrete case) or differential equations (continuous case). For example, a finite dimensional LTI continuous-time system is

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{2.1}$$

where $t \in \mathbb{R}$ is the time variable, $x(t) \in \mathbb{R}^n$ is the state of the system, $u(t) \in \mathbb{R}^p$ is the system's input or control, $y(t) \in \mathbb{R}^m$ is called system output, and $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{m\times n}$, $D \in \mathbb{R}^{m\times p}$ are matrices. Likewise, a finite dimensional LTI discrete-time system is

$$\begin{aligned} Ex_{\gamma+1}(t) &= Ax_\gamma(t) + Bu_\gamma(t), \qquad x_\gamma(t_0) = x_0, \\ y_\gamma(t) &= Cx_\gamma(t) + Du_\gamma(t), \end{aligned} \tag{2.2}$$

where $t \in \mathbb{Z}$. E, A, B, C and D can be defined as above.
Since the matrices E, A, B, C and D are time-invariant and linear maps in the linear spaces, system (2.1) and (2.2) are known as *linear time-invariant* systems. The *dimension* of the system, that is identified by the dimension of the state is n. If $p = m = 1$, the system is said to be a *single-input single-output* (SISO) system. A *multi-input multi-output* (MIMO) system has $m, p > 1$. For the sake of convenience we will write a system $\mathbf{\Sigma}$ compactly as (A, B, C, D, E) or as (A, B, C) if E is the identity matrix and D is absent to represent a system. Neither the time-varying system, where A, B, C, D, E are depending on time nor the discrete-time system will appear in later discussion, because they are not involved in this thesis. If E is an identity matrix (i.e., $E = I$), system in (2.1) is called a *standard state space* system. Otherwise the system in (2.1) is in generalized form and known as *generalized state space system*. Again, if E is invertible then the system in (2.1) can be written in the following standard state space form:

$$\begin{aligned} \dot{x}(t) &= \bar{A}x(t) + \bar{B}u(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{2.3}$$

where $\bar{A} = E^{-1}A$ and $\bar{B} = E^{-1}B$.

2.1.1 Descriptor systems

A descriptor system is a special form of a generalized state space model. Systems of the form in (2.1) with singular matrix E ($\det(E) = 0$) are often referred to as *descriptor systems*. In some references, they are also known as *singular systems* or *differential-algebraic equation (DAE) systems* (see [60, 23]). Such a system appears in different disciplines including power systems, piezo- mechanical systems, electrical circuits, chemical engineering, multi body systems and so on (see [41, 19, 33, 34, 8, 60]). The behavior of such a system is different from a standard state-space system. The eigen-pencil $\mathcal{P} = \lambda E - A$ (see Definition 6) is an important tool for analyzing a descriptor system. Let us consider the following descriptor system

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned} \tag{2.4}$$

where $E, A, B, C, x(t), u(t)$, and $y(t)$ are defined as in (2.1). The solution characteristics of the singular system in (2.4) are determined by the corresponding matrix pencil $\mathcal{P}$ [35]. More precisely, the system has a solution if $\mathcal{P}$ is regular (see Definition 7). The sub-space generated by solutions $x(t)$ is separated into two sub-spaces where one is formed by the eigen-spaces associated with the finite eigenvalues and another one corresponds to the infinite eigenvalues of the system, respectively. Since E is singular, $\operatorname{rank}(E) = r < n$ and $\det(\mathcal{P})$ is a (nonzero) polynomial of degree k ($0 \leq k \leq r$), then there exist non-singular matrices $P, Q \in \mathbb{R}^{n\times n}$ such that pre-multiplying (2.4) by P, and applying

$$Qx = \begin{pmatrix} \bar{z}_1(t) \\ \bar{z}_2(t) \end{pmatrix} \tag{2.5}$$

the following standard canonical form is obtained (see [35, 23]):

$$\dot{\bar{z}}_1(t) = A_1\bar{z}_1(t) + B_1u(t), \tag{2.6a}$$
$$N\dot{\bar{z}}_2(t) = \bar{z}_2(t) + B_2u(t), \tag{2.6b}$$
$$y(t) = C_1\bar{z}_1(t) + C_2\bar{z}_2(t), \tag{2.6c}$$

where $\bar{z}_1 \in \mathbb{R}^k$ and $\bar{z}_2 \in \mathbb{R}^{n-k}$ and A_1, B_1, B_2, C_1 and C_2 are constant matrices of appropriate dimensions and N is an $(n-k)\times(n-k)$ matrix of nilpotency ν, i.e., all eigenvalues of N are zero and $N^\nu = 0$ while $N^i \neq 0$ for $i \leq \nu - 1$ [35]. In the case when $k = r$ the matrix N is simply the zero matrix and the system in (2.4) is called *index-1 system* and the second sub-system in (2.6) is purely algebraic [35]. On the other hand, if $k < r$ and N is in Jordan canonical form, then the index of the system $\nu > 1$ is the size of the largest Jordan

block [35]. We refer to [4, 36] for more details about the index of a system. For any initial condition $\bar{z}_1(0)$ and any input $u(t)$, (2.6*a*) has the following unique solution

$$\bar{z}_1(t) = e^{A_1 t}\bar{z}_1(0) + \int_0^t e^{A_1(t-\tau)}B_1 u(\tau)\, d\tau, \qquad t \geq 0, \tag{2.7}$$

while the solution of (2.6*b*) including the initial condition $\bar{z}_2(0)$ is uniquely determined by the forcing inputs $u(t)$ [35]:

$$\bar{z}_2(t) = -\sum_{i=0}^{\nu-1} N^i B_2 u^{(i)}(t), \qquad t \geq 0, \tag{2.8}$$

where $u^{(i)}(t)$ denotes the ith derivatives of the input. Therefore, unlike ordinary differential equation (ODE) systems, DAE systems do not have smooth solutions for arbitrary initial conditions [35]. Only initial conditions $x(0)$ that are consistent, i.e., $x(0)$, for which $\bar{z}_2(0)$ satisfies (2.8) at $t = 0$, yield smooth solutions [35]. Furthermore, unlike ODE system, if the index ν for the DAE system in (2.4) exceeds one, then the solution of the DAE system may depend on the derivatives of the forcing input $u(t)$, which must be accordingly smooth [35].

2.1.2 Input-output relation in a system

Using the *Laplace transformation*[1] we can convert an equation from the 'time domain' into the 'complex domain'. Applying the Laplace transformation to the system in (2.1), the following relation can be found

$$Y(s) = C(sE - A)^{-1}BU(s) + C(sE - A)^{-1}EX(0) + DU(s), \tag{2.9}$$

where $X(s)$, $Y(s)$ and $U(s)$ are Laplace transformations of $x(t)$, $y(t)$ and $u(t)$, respectively. The rational matrix valued function $G(s) = C(sE - A)^{-1}B + D$ in (2.9) is called *transfer function* (transfer function matrix called by some authors) (TF) of the system in (2.1). If the initial condition $X(0) = 0$ is given, (2.9) gives the following relationship:

$$G(s) = \frac{Y(s)}{U(s)}. \tag{2.10}$$

From (2.10), it is clear that the transfer function is nothing but the ratio of the output of a system to the input of a system in the Laplace domain. The

[1] The Laplace transformation of a function $f(t)$, for all $t \geq 0$, is denoted by $F(s)$ and defined by $(\mathcal{L}[f(t)] =) \int_0^\infty f(t)e^{-st}\, dt$, where s is the complex number ($s = a + ib$, with real number a, b).

TF is a powerful tools for simulation, stability analysis, and control design. System characteristics like stability, order of a system, their classification and frequency responses, etc. can fully be described by its TF . We would like to refer to [14, 11] for properties with examples of TF.

Definition 1 ([42]):
Let G(s) be a transfer function of the system in (2.4), defined as (2.9). $G(s)$ is called *proper* if $\lim_{s\to\infty} G(s) < \infty$ and *strictly proper* if $\lim_{s\to\infty} G(s) = 0$. Otherwise $G(s)$ is called *improper*. ◊

For a descriptor system in (2.4), the TF can be divided into two parts (see [42]), $G(s) = G_{sp}(s) + P(s)$, where $G_{sp}(s)$ is strictly proper and $P(s)$ is the polynomial part of $G(s)$, respectively. One should remember that in model order reduction we only reduce the order of the $G_{sp}(s)$ part.

NOTE: *Suppose that $G(s) = G_{sp}(s) + P(s)$ is the transfer function of the full order model (2.4) and $G_r(S) = G_{r_{sp}}(s) + \hat{P}(s)$ be the transfer function of the reduced order model. Then $\|G(s) - G_r(s)\| = \|G_{sp}(s) - G_{r_{sp}}(s)\|$ ($\|.\|$ can be measured by some suitable norm) is small for a good approximation of the strictly proper part.*

Impulse response: For a comparison, we consider the transfer function in (2.10) in the time domain;

$$h(t) = \frac{y(t)}{u(t)}; \quad t \in \mathbb{R}. \tag{2.11}$$

This relationship between input and output is known as *impulse response*, $h(t)$. However, we can not use the impulse response in order to find the system output from the system input like the transfer function. If input and impulse response of a system are given, the system output can be calculated by the following convolution[2] operation:

$$y(t) = h(t) * u(t). \tag{2.12}$$

On the other hand, the *frequency response* of a system is as same as the transfer function except that it is the input-output relation of the system in the complex *Fourier domain* ($\{s = i\omega$ where, $\omega \in \mathbb{R}\}$), not in the Laplace domain. Frequency response of a system can be obtained from the TF in (2.10), by using $s = j\omega$ where $\omega \in \mathbb{R}$.

[2]The convolution of two functions, $x(t)$ and $y(t)$ is $(x * y)(t) = \int_{-\infty}^{\infty} x(\tau)y(t-\tau)\,d\tau$

2.1.3 Controllability and observability

Controllability:

For convenience, we consider the following standard state space system without any output:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \tag{2.13}$$

where $x(t), u(t) \in \mathbb{R}^n$ are vectors, and $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ are matrices. The solution to (2.13) is given by [2, 11, 17]

$$x(t) = e^{At}x_0 + \int_0^{t_f} e^{A(t-\tau)}Bu(\tau)d\tau \qquad \text{for} \quad t \geq 0. \tag{2.14}$$

The integrand in the above equation is a vector, including the integral of $e^{A\tau}Bu(t)$. So each scalar entry of $x(t)$ is just equal to the corresponding scalar entry of a vector in the right hand side. Now an interesting question comes to mind, do we have control over the values of $x(t)$, to reach a desired state, through choice of $u(t)$? If the answer is affirmative, then the system is controllable.

Definition 2 ([71]):
A system in (2.13) or the pair (A, B) is said to be *controllable* over the interval $[0, t_f]$, if for any $x_0 \in \mathbb{R}$. there exists an input $u(t)$, for $t \in [0, t_f]$, such that the state of (2.14) is $x(t_f) = 0$. ◊

To be more clear let us study three related issues with respect to the system in (2.13):

- ***Set of reachable states*** [56] : For a fixed $t \geq 0$, let $\mathcal{R}_t$ denote the set of all states that are reachable at time t by some input function $u(t)$. Namely $\mathcal{R}_t$ is the set

$$\mathcal{R}_t = \{\xi \in \mathbb{R}^n; \exists\ u(t) \text{ such that } x(t; u) = \xi\}. \tag{2.15}$$

  It turns out that $\mathcal{R}_t$ is a subspace of $\mathbb{R}^n$.

- ***Controllable subspace*** [2, 11, 56]: If a state space equation as in (2.13) is given, the *controllability matrix* $\mathcal{C}$, is defined as follows:

$$\mathcal{C} = [B\ AB\ \cdots\ A^{n-1}B]. \tag{2.16}$$

  Note that, the rank of this matrix is *full* (equal to the dimension of the system) if only if the system is fully controllable. *Controllability subspace*, denoted by $\mathcal{C}_s$ is the image of the controllability matrix, i.e.,

$$\mathcal{C}_s = \text{Im}[B\ AB\ \cdots\ A^{n-1}B]. \tag{2.17}$$

Thus we see that like $\mathcal{R}_t$, $\mathcal{C}_s$ is also a subspace of $\mathbb{R}^n$.

- ***Controllability Gramian:*** Another object associated with the state equation is the controllability Gramian. For each $t > 0$ the time dependent *controllability Gramian* is defined to be the $n \times n$ matrix

$$\mathcal{W}_c = \int_0^\infty e^{A\tau} B B^T e^{A^T \tau} \, d\tau. \tag{2.18}$$

We would like to refer to [57] for more details.

Now we are ready to state some important results regarding controllability.

Theorem 1 ([17, 11, 56]):
The following are equivalent for the matrix pair (A, B) of the system (2.13):

1. (A, B) is controllable.
2. The controllability matrix has full rank i.e., $\text{rank}(\mathcal{C}) = n$.
3. The controllability Gramian is positive definite i.e., $\mathcal{W}_c > 0$ for some $t > 0$.
4. $\text{rank}[sI - A, B] = n$ for all $s \in \mathbb{C}$.
5. The pair $(\tilde{A}, \tilde{B})$ is controllable, where $\tilde{A} = TAT^{-1}$ and $\tilde{B} = TB$ for any nonsingular $T \in \mathbb{R}^{n \times n}$. ◊

A proof is available in [17, 11, 56].

Observability:

To illustrate observability, let us consider the following system which has an output but no input

$$\begin{aligned} \dot{x}(t) &= Ax(t) \qquad \text{with} \quad x(t_0) = x_0, \\ y(t) &= Cx(t). \end{aligned} \tag{2.19}$$

The solution of this equation is clearly,

$$y(t) = Ce^{At} x_0 \qquad \text{for } t \geq 0. \tag{2.20}$$

Again the same question arises: Is it possible to find x_0 from the above equation? If the answer is yes, we say that the system (2.19) is observable.

Definition 3 ([12]):
A system as in (2.19) with initial state $x(t_0) = x_0$ is observable if the value of the initial state can be determined from the system output $y(t)$, that has been observed through the time interval $t_0 < t < t_f$. If the initial state can not be so determined, then the system is unobservable. ◊

NOTE: *The notations of controllability and observability can be thought as dual of each other and any result that we obtain for controllability has a counterpart in terms of observability.*
The following concepts are related to observability [17, 2]:

1. Observability matrix $\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$.

2. The unobservable subspace is $\mathcal{O}_s^{un} = \ker(\mathcal{O})$.

3. Observable subspace $\mathcal{O}_s = \mathrm{Im}(\mathcal{O})$.

4. Observability Gramian $\mathcal{W}_o = \int_0^\infty e^{A^T\tau} C^T C e^{A\tau}\, d\tau$.

We conclude this subsection with the following theorem:

Theorem 2 ([17, 71]):
The following are equivalent:

1. (A, C) is controllable pair.

2. The rank of the observability matrix $\mathcal{O}$ is full; i.e., $\mathrm{rank}(\mathcal{O}) = n$.

3. The observability Gramian is positive definite: $\mathcal{W}_o > 0$.

4. $\mathrm{rank}\begin{bmatrix} A - sI \\ C \end{bmatrix} = n$ for all $s \in \mathbb{C}$.

5. There exists a similarity transformation T, such that $TAT^{-1} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{12} & \tilde{A}_{22} \end{bmatrix}$ and $CT^{-1} = \begin{bmatrix} \tilde{C}_1 & 0 \end{bmatrix}$, where $(\tilde{C}_1, \tilde{A}_{11})$ is observable. ◊

A proof of this theorem is available in [17, 71].

2.1.4 Stability

The stability of a system can be described in many different ways, such as exponential stability [12], asymptotic stability, Lyapunov stability, BIBO (bounded-input bounded-output) stability. Here, we include the concept of stability of a system briefly from [11, 55, 60].

Definition 4:

A system as in (2.1) is called asymptotically stable if $\lim_{t\to\infty} x(t) = 0$ (equilibrium point of the system) for all trajectories $x(t)$ of $E\dot{x}(t) = Ax(t)$.◊

A time-invariant continuous time linear system (2.1) is stable if $\mathrm{Re}(\lambda_i(A, E)) < 0$ (E is nonsingular); $\forall\, i$ i.e., all eigenvalues of the pencil $(\lambda E - A)$ have negative real parts, because in that case, Definition 4 is satisfied [17]. Another popular test, to measure stability of a system is the *Lyapunov stability condition* which is stated in the following theorem.

Theorem 3:

Consider a system in (2.1) with nonsingular matrix E. If the system is stable then the continuous time generalized Lyapunov equation

$$AXE^T + EXA^T = -Q \tag{2.21}$$

has a unique Hermitian positive semi-definite solution X, for every Hermitian positive definite matrix Q. ◊

2.1.5 Realization

In Subsection 2.1.2, we have shown that for a given system as in (2.1) the transfer function of order $m \times p$ has the following form

$$G(s) = C(sE - A)^{-1}B + D, \qquad s \in \mathbb{C}. \tag{2.22}$$

A set of matrices E, A, B, C, D which satisfies (2.22) is called realization of $G(s)$ [2, 22]. We will denote such a realization by (E, A, B, C, D) or $\left[\begin{array}{c|c} sE - A & B \\ \hline C & D \end{array}\right]$. However, this realization for $G(s)$ is not unique. Among different relations of a particular transfer function $G(s)$, we are interested in the so called minimal one as defined in the following:

Definition 5 ([42]):

A realization of the transfer function $G(s)$ is called minimal if the matrices E and A have smallest possible dimension. ◊

Theorem 4 ([42]):
Suppose that (E, A, B, C, D) is a realization of $G(s)$ in (2.22) of the system in (2.1), then (E, A, B, C, D) is minimal if and only if the system is controllable and observable. ◊

2.1.6 Hankel operator and Hankel singular values

Recall that, $\mathcal{L}_2(I)$ be the space of square integrable functions, defined on the interval I. The Hankel operator $\mathcal{H}_{\mathbf{\Sigma}}$, of the system $\mathbf{\Sigma}$ which maps past inputs (u_-) into future outputs (y_+) is defined as follows:

$$\mathcal{H}_{\mathbf{\Sigma}} : \mathcal{L}_2(\mathbb{R}_-) \to \mathcal{L}_2(\mathbb{R}_+), \tag{2.23}$$

$$u_- \to y_+ \quad \text{where} \quad y_+(t) = \int_{-\infty}^{0} h_{\mathbf{\Sigma}}(t-\tau)u_-(\tau)\,d\tau, \quad t \geq 0.$$

In control theory, Hankel singular values are considered as a measure of energy for each state in a system. They are the basis of balanced model reduction, in which low energy states are discarded while high energy states are preserved.
The Hankel singular values of a stable system $\mathbf{\Sigma}$ are denoted by:

$$\sigma_1(\mathbf{\Sigma}) > \cdots > \sigma_{\mathbf{q}}(\mathbf{\Sigma}) \quad \text{with multiplicities } \mathbf{r_i}, \quad \mathbf{i = 1, 2, \cdots, q}, \quad \sum_{\mathbf{i=1}}^{\mathbf{q}} \mathbf{r_i} = \mathbf{n}, \tag{2.24}$$

the singular values of $\mathcal{H}_{\mathbf{\Sigma}}$ defined by (2.23) which are equal to the square roots of the eigenvalues for the product of the two Gramians [1], i.e.,

$$\sigma_i(\mathcal{H}_{\mathbf{\Sigma}}) = \sqrt{\lambda_i(\mathcal{W}_c\mathcal{W}_o)}.$$

2.2 Background in linear algebra and matrix theory

2.2.1 Generalized eigenvalue problem

Let us consider an ordered pair (A, E), where $A, E \in \mathbb{C}^{n\times n}$ and a nonzero vector $x \in \mathbb{C}^n$. If there exist $\alpha, \beta \in \mathbb{C}$ not both zero, such that

$$\alpha Ax = \beta Ex, \tag{2.25}$$

if $\alpha \neq 0$, then

$$Ax = \lambda Ex, \tag{2.26}$$

the scalar $\lambda = \frac{\beta}{\alpha}$ is called *eigenvalue or characteristic value* of the pair (A, E) and x is an *eigenvector* corresponding to the eigenvalue λ. These are the *finite* eigenvalues. On the other hand if (2.25) holds with $\alpha = 0$ and $\beta \neq 0$, then the eigenvalues of (A, E) are called *infinite*. The eigenvalue problem associated with the pair (A, E) is known as *generalized eigenvalue* problem (see [49, 5, 15, 69]). In the case $E = I$ (identity matrix) we obtain the standard eigenvalue problem.

Theorem 5 ([69]):
Let $A, E \in \mathbb{C}^{n\times n}$ and $\lambda \in \mathbb{C}$ be nonzero.

- λ is an eigenvalue of (A, E) iff $\frac{1}{\lambda}$ is an eigenvalue of (E, A).
- ∞ is an eigenvalue of (A, E) iff E is a singular matrix.
- ∞ is an eigenvalue of (A, E) iff 0 is an eigenvalue of (E, A).
- If E is nonsingular, the eigenvalues of (A, E) are exactly the eigenvalues of AE^{-1} and $E^{-1}A$. ◊

Definition 6 ([25]):
Let, $A, E \in \mathbb{C}^{n\times n}$ be matrices, the expression $A - \lambda E$ with indeterminant λ is called *eigen pencil* or *matrix pencil*. we denote this by $\mathcal{P}$. ◊

The terms matrix pencil or matrix pair can be used interchangeably, i.e., if any nonzero vector x is an eigenvector of the pencil $\mathcal{P}$, it is also called an eigenvector of the pair (A, E). Note that any $\lambda \in \mathbb{C}$ is an eigenvalue of (A, E) iff $A - \lambda E$ is singular [69] i.e.,

$$\det(\lambda E - A) = 0. \tag{2.27}$$

Equation (2.27) is known as *characteristic equation* of the pair (A, E), where the function $\Delta(\lambda) = \lambda E - A$ is the *characteristic polynomial* of degree n or less.

Definition 7 ([69]):
A matrix pair (A, E), where $A, E \in \mathbb{C}^{n\times n}$ is called *singular* if $\det(\lambda E - A) = 0$. Otherwise it is called *regular* pair. ◊

To compute eigenvalues of large sparse matrices, the Krylov-based Arnoldi process is one of the most powerful tools.

Definition 8 ([15]):
Let $A \in \mathbb{R}^{n\times n}$ and $v \in \mathbb{R}^n$ with $v \neq 0$ then,

$$K_m(A, v) = [v, Av, \cdots A^{m-1}v] \tag{2.28}$$

is called the mth *Krylov matrix* associated with A and v. The corresponding subspace

$$\mathcal{K}_m(A, v) = \text{span}(v, Av, \cdots, A^{m-1}v) \tag{2.29}$$

is called the mth *Krylov subspace* associated with A and v. ◊

Theorem 6 ([15, 65]):

Let the columns of $V_{m+1} = [v_1, v_2, \cdots, v_{m+1}] \in \mathbb{R}^{n\times(m+1)}$ form an orthogonal basis for $\mathcal{K}_{m+1}(A, v_1)$, then there exists an $(m+1) \times m$ unreduced upper Hessenberg matrix [15]

$$\widehat{H}_m = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ & \ddots & \ddots & \vdots \\ & & & h_{m,m} \\ & & & h_{m+1,m} \end{bmatrix} \tag{2.30}$$

such that

$$AV_m = V_{m+1}\widehat{H}_m. \tag{2.31}$$

Conversely, a matrix V_{m+1}, with orthonormal columns satisfies a relation of the form in (2.31) only, if the columns of V_{m+1} form a basis for $\mathcal{K}_{m+1}(A, v_1)$. ◊

Definition 9 ([15, 65]):

Let the column of $V_{m+1} = [V_m, v_{m+1}] \in \mathbb{R}^{n\times m+1}$ form an orthogonal basis. If there exists a Hessenberg matrix $\widehat{H}_m \in \mathbb{R}^{m+1\times m}$ of the form in (2.30) so that

$$AV_m = V_{m+1}\widehat{H}_m, \tag{2.32}$$

then (2.32) is called an (unreduced) Arnoldi decomposition of order m.◊

By a suitable partition of $\widehat{H}_m$, one can rewrite (2.32) as

$$AV_m = \begin{bmatrix} V_m & v_{m+1} \end{bmatrix} \begin{bmatrix} H_m \\ h_{m+1,m}e_m^T \end{bmatrix} = V_m H_m + h_{m+1,m}v_{m+1}e_m^T. \tag{2.33}$$

By the the orthogonality property of v_{m+1}, (2.33) yields (see [65] or [49] for details)

$$H_m = V_m^T A V_m. \tag{2.34}$$

Note that H_m is a projection of A onto the Krylov subspaces $\mathcal{K}_m(A, v)$. As a result, one can imagine that its eigenvalues are related to those of A and these eigenvalues are known as *Ritz values* of A.

Definition 10 ([15]):
Let $A \in \mathbb{R}^{n\times n}$ and let the columns of $V_m \in \mathbb{R}^{n\times m}$ be orthonormal. The $m \times m$ matrix $H_m = V_m^T A V_m$ is called *Rayleigh quotient*, an eigenvalue λ of H_m is called a *Ritz value*, and if v is an eigenvector of H_m associate with λ, then $V_m v$ is called a *Ritz vector* belonging to λ. ◊

2.2.2 Projection matrix

A projector or projection matrix P is a square matrix that satisfies

$$P = P^2. \tag{2.35}$$

Such a matrix is also known as *idempotent* matrix. If P is a projector, $I - P$ is also a projector because

$$(I - P)^2 = 1 - 2P + P^2 = 1 - 2P + P = 1 - P.$$

$I - P$ is called *complementary* projector to P [65].

Definition 11 ([65]):
Let S_1 and S_2 be two orthogonal subspaces of $\mathbb{C}^m$ such that $S_1 \cap S_2 = \{0\}$ and $S_1 + S_2 = \mathbb{C}^m$, where $S_1 + S_2$ denote the span of S_1 and S_2, i.e., the set of vectors $s_1 + s_2$ with $s_1 \in S_1$ and $s_2 \in S_2$. If a projector projects onto a subspace S_1 along a subspace S_2 then it is called *orthogonal projector*. ◊

Theorem 7 ([65]):
A projector P is orthogonal if and only if $P = P^*$ where P^* is the Hermitian. ◊

2.2.3 Singular value decomposition (SVD) and eigenvalue decomposition

In this subsection we introduce two important decompositions in linear spaces, namely the singular value decomposition and eigenvalue decomposition briefly.

Definition 12:
If the columns of $V \in \mathbb{C}^{n\times n}$ contain linearly independent eigenvectors of $A \in \mathbb{C}^{n\times n}$, the *eigenvalue decomposition* of A is

$$A = V \Lambda V^{-1}, \tag{2.36}$$

where $\Lambda \in \mathbb{C}^{n\times n}$ is an $n \times n$ diagonal matrix whose entries are the eigenvalues of A. ◊

Definition 13:
Let $m, n \in \mathbb{R}$ be arbitrary. Given $A \in \mathbb{C}^{m\times n}$, the *singular value decomposition* of A is a factorization

$$A = U\Sigma V^*, \tag{2.37}$$

where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary, and $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal. In addition, diagonal elements σ_j $(j = 1, 2, \cdots, k)$ of Σ are non-negative and in decreasing order; i.e., $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$, where $k = \min(m, n)$. ◊

This SVD is known as *full singular value decomposition* by some authors. On the other hand, if $U_1 = (u_1, \cdots, u_n) \in \mathbb{C}^{m\times n}$, $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$ then $A = U_1\Sigma_1 V^*$. This factorization is known as *thin* SVD or *economic* SVD of A (see [25] for details).
The following statements are true for a matrix A [65]:

1. The singular values $\sigma_1, \sigma_2, \cdots, \sigma_n$ of A are the square roots of the eigenvalues of the symmetric positive semi-definite matrix A^TA.
2. The right singular vectors are the eigenvectors of the matrix A^TA, and the left singular vectors are the eigenvectors of the matrix AA^T.
3. The rank of A is r, the number of nonzero singular values.
4. If $A = A^*$, then the singular values of A are the absolute values of the eigenvalues of A.
5. For $A \in \mathbb{C}^{n\times n}$, $\det(A) = \prod_{j=1}^{n} \sigma_j$.
6. A is the sum of rank-one matrices.
7. $\| A \|_2 = \sigma_1$ and $\| A \|_F = (\sigma_1^2 + \cdots + \sigma_r^2)^{\frac{1}{2}}$.
8. The condition number of A is $\text{cond}(A) = \frac{\sigma_{max}}{\sigma_{min}}$, where $A \in \mathbb{C}^{n\times n}$ is non-singular and σ_{max} and σ_{min} are the maximum and minimum singular value, respectively of A.

2.2.4 Some important definitions

Vector norms and matrix norms [65]: Let X be a vector space. A real valued function $\| . \|: X \to \mathbb{R}$ is said to be a norm on X if it satisfies the following properties:

1. $\| x \| \geq 0$ and $\| x \| = 0$ iff $x = 0$,
2. $\| x + y \| \leq \| x \| + \| y \|$,
3. $\| \alpha x \| = |\alpha| \, \| x \|$,

for any $x \in X$, $y \in X$ and $\alpha \in \mathbb{R}$. Let $x \in \mathbb{C}^n$. Then we define the vector $p-norm$ of x as

$$\| x \|_p = (\sum_{i=1}^{n} \| x_i \|^p)^{\frac{1}{p}}, \qquad \text{for } 1 \leq p < \infty.$$

In particular, when $p = 1, 2, \infty$ we have

$$\| x \|_1 = \sum_{i=1}^{n} |x_i|,$$
$$\| x \|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2},$$
$$\| x \|_\infty = \max_{1 \leq i \leq n} | x_i | .$$

Let $A = [a_{ij}] \in \mathbb{C}^{m \times n}$; then the matrix norm induced by a vector $p-$norm is defined as

$$\| A \|_p = \sup_{x \neq 0} \frac{\| Ax \|_p}{\| x \|_p}.$$

The matrix norms induced by vector p-norms are sometimes called *induced p-norms*. In particular the induced matrix 1-norm and 2-norm can be computed by

$$\| A \|_1 = \max_{1 \leq j \leq n} \| a_j \|_1; \quad a_j \text{ is the } jth \text{ column of } A,$$
$$\| A \|_2 = \sqrt{\lambda_{max}(A^* A)}.$$

The most important matrix norm which is not induced by a vector norm is the *Frobenius norm* defined by

$$\| A \|_F = (\sum_{i=1}^{m} \sum_{j=1}^{n} | a_{ij} |^2)^{1/2}. \tag{2.38}$$

$\mathcal{H}_\infty$ Space and $\mathcal{H}_\infty$ norm [71]: $\mathcal{H}_\infty$ is a (closed) subspace of $\mathcal{L}_\infty(j\mathbb{R})$ (Banach space of matrix-valued or scalar-valued functions on $j\mathbb{R}$) with functions $G(s)$ that are analytic and bounded in $Re(s) > 0$ (open right half plane). The corresponding norm can be defined as

$$\| G \|_\infty = \sup_{Re(s)>0} \sigma[G(s)] = \sup_{\omega \in \mathbb{R}} \sigma[G(j\omega)]. \tag{2.39}$$

Invariant subspaces: For a given transformation A, a subspace $S \subset \mathbb{C}^n$ is called *invariant*, or *A-invariant*, if $Ax \in S$ for every $x \in S$. In

other words, that S is invariant for A means that S contains its image under A i.e., $AS \subset S$. For example, 0, $\mathbb{C}^n$, ker(A), img(A) are all A-invariant subspaces.

Sparse matrices: A matrix with special structure (see Section 3.4 in Chapter 3), that has relatively few nonzero entries is called *sparse matrix*. A sparse matrix is a matrix that allows special techniques to take advantage of the large number of zero elements and their locations. Usually, standard discretizations of partial differential equations lead to large and sparse matrices. More details about sparse matrices, their properties, their representations, and the data structures used to store them, can be found in [50].

CHAPTER

THREE

INTRODUCTION TO MODEL PROBLEMS

We now turn our attention to the study of some special model problems, namely *power system* and *piezo-mechanical* models. How to reduce a higher order system into state space representations, by preserving the system properties (specially the sparsity) is one of the vital issues to discuss in this chapter. In the first subsection we introduce general second order models, whereas the second and third subsections, present power system models and piezo-mechanical models, respectively. We will give some graphical representations of system matrices in the later subsection. Reference papers are [41, 19, 33, 10, 20, 34, 52, 53, 61].

3.1 Second order models

We consider LTI continuous-time second-order systems, as they arise in many practical applications, including electrical circuits, mechanical systems, large structures and microsystems technology, etc [61]:

$$\begin{aligned} M\ddot{z}(t) + D\dot{z}(t) + Kz(t) &= B_1 u(t), \\ y(t) &= C_1\dot{z}(t) + C_2 z(t), \end{aligned} \tag{3.1}$$

where M, D, and K are matrices of dimension n, $B_1 \in \mathbb{R}^{n\times p}$, $C_1, C_2 \in \mathbb{R}^{m\times n}$, $z(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^p$ is the input vector and $y(t) \in \mathbb{R}^m$ is the output vector. In mechanical engineering the matrices M, D and K have special names: mass, damping and stiffness matrix, respectively. In case $D = 0$, the system is called *undamped*. Equivalently, the model in (3.1) can

be written as

$$\begin{aligned} \underbrace{\begin{bmatrix} \mathcal{F} & 0 \\ 0 & M \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} \dot{z}(t) \\ \ddot{z}(t) \end{bmatrix}}_{\dot{x}(t)} &= \underbrace{\begin{bmatrix} 0 & \mathcal{F} \\ -K & -D \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} 0 \\ B_1 \end{bmatrix}}_{B} u(t), \\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{C} \begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}, \end{aligned} \tag{3.2}$$

where E,$A \in \mathbb{R}^{2n\times 2n}$, $B \in \mathbb{R}^{2n\times p}$, and $C \in \mathbb{R}^{m\times 2n}$. This is the first order representation of the system in (3.1). The choice of $\mathcal{F} \in \mathbb{R}^{n\times n}$ is optional. For simplicity one may choose $\mathcal{F} = I$. Most of the cases, M, K and D are symmetric and even positive definite and K is nonsingular. Therefore it is recommended to write (3.1) as the following first order form:

$$\begin{aligned} \underbrace{\begin{bmatrix} -K & 0 \\ 0 & M \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} \dot{z}(t) \\ \ddot{z}(t) \end{bmatrix}}_{\dot{x}(t)} &= \underbrace{\begin{bmatrix} 0 & -K \\ -K & -D \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} 0 \\ B_1 \end{bmatrix}}_{B} u(t), \\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{C} \begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}. \end{aligned} \tag{3.3}$$

In this case, E is symmetric since M and K are symmetric. A is also symmetric as K and D are symmetric. A second order system in (3.1) can also be transferred into a state space form as (see [52])

$$\begin{aligned} \underbrace{\begin{bmatrix} 0 & \mathcal{F} \\ M & D \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} \ddot{z}(t) \\ \dot{z}(t) \end{bmatrix}}_{\dot{x}(t)} &= \underbrace{\begin{bmatrix} \mathcal{F} & 0 \\ 0 & -K \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \dot{z}(t) \\ z(t) \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} 0 \\ B_1 \end{bmatrix}}_{B} u(t), \\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{C} \begin{bmatrix} \dot{z}(t) \\ z(t) \end{bmatrix}, \end{aligned} \tag{3.4}$$

which is equivalent to (3.2). This representation might be helpful to have numerical stability of the system if M is badly conditioned. In this case the common choice of $\mathcal{F}$ can be I and that changes the system in (3.4) into [52],

$$\begin{aligned} \underbrace{\begin{bmatrix} 0 & I \\ M & D \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} \ddot{z}(t) \\ \dot{z}(t) \end{bmatrix}}_{\dot{x}(t)} &= \underbrace{\begin{bmatrix} I & 0 \\ 0 & -K \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \dot{z}(t) \\ z(t) \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} 0 \\ B_1 \end{bmatrix}}_{B} u(t), \\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{C} \begin{bmatrix} \dot{z}(t) \\ z(t) \end{bmatrix}. \end{aligned} \tag{3.5}$$

Also, the choice of $\mathcal{F} = M$ in (3.4) would again preserve the symmetry and might be an interesting choice. Note that, in the case of an undamped model, one can put 0 (zero matrix) instead of D, in each of the above first order representations.

Note: The input-output relation of the system in (3.1) can directly be described by the following transfer function (or transfer function matrix) (see [61, 51])

$$G(s) = (C_1 + sC_2)(s^2M + sD + K)B_1, \qquad \text{where } s \in \mathbb{C}. \tag{3.6}$$

3.2 Power system models

Linearizing around the equilibrium point, a stable time invariant power system model can be represented by a set of differential and algebraic equations [41, 19, 33, 10, 20]:

$$\begin{array}{rcl} \dot{x}(t) &=& J_a x_1(t) + J_b x_2(t) + B_1 u(t), \\ 0 &=& J_c x_1(t) + J_d x_2(t) + B_2 u(t), \\ y(t) &=& C_1 x_1(t) + C_2 x_2(t) + D_s u(t), \end{array} \tag{3.7}$$

where $x_1(t) \in \mathbb{R}^{n_1}$ is the state vector, $x_2(t) \in \mathbb{R}^{n_2}$ is the vector of algebraic variables, $u(t) \in \mathbb{R}^p$ is the input vector, and $y(t) \in \mathbb{R}^m$ is the vector of output variables. J_a, J_b, J_c, J_d are sub-matrices of the system's Jacobian matrix [41, 19]. J_d is always invertible and for large systems J_a, J_b, J_c, J_d are highly sparse [10, 41, 19]. However, this is the descriptor form of power system models. Eliminating the algebraic variables from (3.7), we can reduce the system into standard state space form

$$\begin{array}{rcl} \dot{x}(t) &=& (J_a - J_b J_d^{-1} J_c)x_1(t) + (B_1 - J_b J_d^{-1} B_2)u(t), \\ y(t) &=& (C_1 - C_2 J_d^{-1} J_c)x_1(t) + (D_s - C_2 J_d^{-1} B_2)u(t). \end{array} \tag{3.8}$$

Now, consider $J_a - J_b J_d^{-1} J_c = J$, $B_1 - J_b J_d^{-1} B_2 = B$, $C_1 - C_2 J_d^{-1} J_c = C$ and $D_s - C_2 J_d^{-1} B_2 = D_a$ to write the set of equations in (3.8) as a simpler form

$$\begin{array}{rcl} \dot{x}(t) &=& J x_1(t) + B u(t), \\ y(t) &=& C x_1(t) + D_a u(t), \end{array} \tag{3.9}$$

where $J \in \mathbb{R}^{n_1 \times n_1}$, $B \in \mathbb{R}^{n_1 \times p}$, $C \in \mathbb{R}^{m \times n_1}$ and $D_a \in \mathbb{R}^{m \times p}$ are matrices.

The descriptor form in (3.7) can also be written in a compact form as

$$\begin{array}{rcl} E\dot{z}(t) &=& Az(t) + Fu(t), \\ y(t) &=& Lz(t) + D_s u(t), \end{array} \tag{3.10}$$

where $E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)}$, $A = \begin{bmatrix} J_a & J_b \\ J_c & J_d \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)}$, $F = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times p}$, $L = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \in \mathbb{R}^{m\times(n_1+n_2)}$ and $z(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \in \mathbb{R}^{n_1+n_2}$.

3.3 Piezo-mechanical models

The piezo mechanical model considered in this thesis is assumed to be given in the form [34]

$$\begin{aligned} M\ddot{\xi}(t) + D\dot{\xi}(t) + K\xi(t) &= Qu(t), \\ y(t) &= G\xi(t), \end{aligned} \tag{3.11}$$

where $M \in \mathbb{R}^{n\times n}$, $D \in \mathbb{R}^{n\times n}$, $K \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times p}$ and $G \in \mathbb{R}^{m\times n}$ are known as mass, damping, stiffness, input and output matrices, respectively and all are highly sparse. M is singular. Reordering M, system in (3.11) can be reformulated as

$$\begin{aligned} \begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{z}(t) \\ \ddot{\varphi}(t) \end{bmatrix} + \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{z}(t) \\ \dot{\varphi}(t) \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} \\ K_{12}^T & K_{22} \end{bmatrix} \begin{bmatrix} z(t) \\ \varphi(t) \end{bmatrix} &= \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} u(t), \\ y(t) &= \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} z(t) \\ \varphi(t) \end{bmatrix}, \end{aligned} \tag{3.12}$$

where $z(t) \in \mathbb{R}^{n_1}$, $\varphi(t) \in \mathbb{R}^{n_2}$, $n_1 + n_2 = n$, M_1 is the mechanical mass matrix, D_1 is the mechanical damping matrix, and K is the stiffness matrix (including mechanical (K_{11}), electrical (K_{22}) and coupling (K_{12}) terms). The general force quantities (mechanical forces and electrical charges) are chosen as the input quantities u, and the corresponding general displacements (mechanical displacements and electrical potential) as the output quantities y. The total mass matrix contains zeros at the locations of electrical potential. More precisely, the electrical potential of piezo-mechanical systems (the electrical degree of freedom) is not associated with an inertia. The equation of motion of a mechanical system can be found in [45] (eq. 1). This equation results from a finite element discretization of the balance equations. For piezo-mechanical systems these are the mechanical balance of momentum (with inertia term) and the electro-static balance. From this, electrical potential without inertia term is obtained. Thus, for the whole system (mechanical and electrical dof) the mass matrix has rank deficiency. And so called index-1 system is arisen. This is the basic difference between the piezo-mechanical and a general mechanical system. In the general mechanical system the piezo part is absent. Therefore, the mass matrix has full rank in the general mechanical system. See [45] for more details. Equation

(3.12) is equivalent to

$$\begin{aligned} M_1\ddot{z}(t) + D_1\dot{z}(t) + K_{11}z(t) + K_{12}\varphi(t) &= Q_1u(t),\\ K_{12}^Tz(t) + K_{22}\varphi(t) &= Q_2u(t),\\ y(t) &= G_1z(t) + G_2\varphi(t). \end{aligned} \tag{3.13}$$

Let us consider $z(t) = x_1(t)$, $\dot{z}(t) = x_2(t) \Rightarrow \dot{z}(t) = \dot{x}_1(t)$, and $\ddot{z}(t) = \dot{x}_2(t)$. Applying these to (3.13), we obtain the following first order form of the piezo-mechanical model (3.11)

$$\begin{aligned} \begin{bmatrix} I & 0 & 0\\ 0 & M_1 & 0\\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(t)\\ \dot{x}_2(t)\\ \dot{\varphi}(t) \end{bmatrix} &= \begin{bmatrix} 0 & I & 0\\ -K_{11} & -D_1 & -K_{12}\\ -K_{12}^T & 0 & -K_{22} \end{bmatrix} \begin{bmatrix} x_1(t)\\ x_2(t)\\ \varphi(t) \end{bmatrix} + \begin{bmatrix} 0\\ Q_1\\ Q_2 \end{bmatrix} u(t),\\ y(t) &= \begin{bmatrix} G_1 & 0 & G_2 \end{bmatrix} \begin{bmatrix} x_1(t)\\ x_2(t)\\ \varphi(t) \end{bmatrix}. \end{aligned} \tag{3.14}$$

Let us consider $\begin{bmatrix} I & 0\\ 0 & M_1 \end{bmatrix} = E_1$, $\begin{bmatrix} 0 & I\\ -K_{11} & -D_1 \end{bmatrix} = J_a$, $\begin{bmatrix} 0\\ -K_{12} \end{bmatrix} = J_b$, $\begin{bmatrix} -K_{12}^T & 0 \end{bmatrix} = J_c$, $\begin{bmatrix} -K_{22} \end{bmatrix} = J_d$, $\begin{bmatrix} 0\\ Q_1 \end{bmatrix} = B_1$, $Q_2 = B_2$, $\begin{bmatrix} G_1 & 0 \end{bmatrix} = C_1$, $G_2 = C_2$ and $\begin{bmatrix} x_1(t)\\ x_2(t) \end{bmatrix} = x(t)$, and using in (3.14),

$$\begin{aligned} \underbrace{\begin{bmatrix} E_1 & 0\\ 0 & 0 \end{bmatrix}}_{E} \begin{bmatrix} \dot{x}(t)\\ \dot{\varphi}(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} J_a & J_b\\ J_c & J_d \end{bmatrix}}_{A} \begin{bmatrix} x(t)\\ \varphi(t) \end{bmatrix} + \underbrace{\begin{bmatrix} B_1\\ B_2 \end{bmatrix}}_{F} u(t),\\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{L} \begin{bmatrix} x(t)\\ \varphi(t) \end{bmatrix}, \end{aligned} \tag{3.15}$$

where $E \in \mathbb{R}^{(2n_1+n_2)\times(2n_1+n_2)}$, $A \in \mathbb{R}^{(2n_1+n_2)\times(2n_1+n_2)}$, $B \in \mathbb{R}^{(2n_1+n_2)\times p}$, and $C \in \mathbb{R}^{m\times(2n_1+n_2)}$. J_d is always invertible. System in (3.15) is an index-1 descriptor system (see Chapter 2). Note that matrices E, A, F, and L are extremely sparse. Since J_d is invertible one can easily put system in (3.15) as the following generalized state space form

$$\begin{aligned} E_1\dot{x}(t) &= Jx(t) + Bu(t),\\ y(t) &= Cx(t) + D_au(t), \end{aligned} \tag{3.16}$$

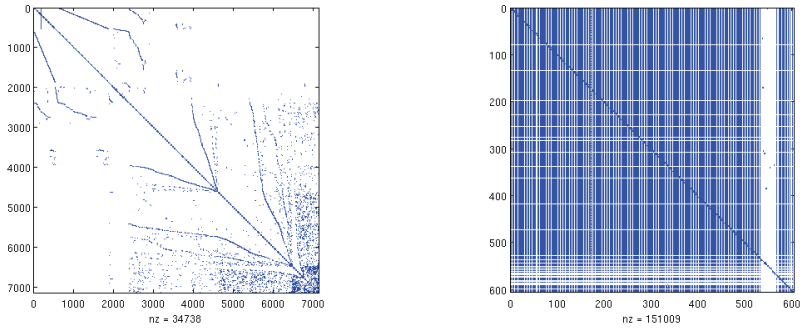
where $J = (J_a - J_bJ_d^{-1}J_c) \in \mathbb{R}^{2n_1\times 2n_1}$, $B = (B_1 - J_bJ_d^{-1}B_2) \in \mathbb{R}^{2n_1\times p}$, $C = (C_1 - C_2J_d^{-1}J_c) \in \mathbb{R}^{m\times 2n_1}$ and $D_a = (-C_2J_d^{-1}B_2) \in \mathbb{R}^{m\times p}$, or the following standard state space form

$$\begin{aligned} \dot{x}(t) &= \overline{J}x(t) + \overline{B}u(t),\\ y(t) &= Cx(t) + D_au(t), \end{aligned} \tag{3.17}$$

where $\overline{J} = E_1^{-1}J$ and $\overline{B} = E_1^{-1}B$. However, neither the representation in (3.16) nor in (3.17) is our desired form, sice in both cases system matrices J and $\overline{J}$ are dense. Although, the dimension of the system in (3.15) is larger, the system matrices are highly sparse in this representation. We will discuss this in the coming section.

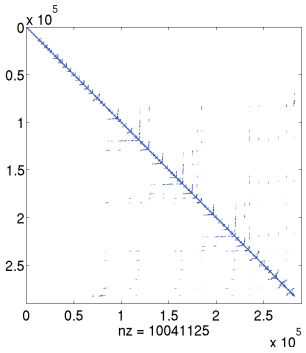
3.4 Graphical representation of the system matrices

The aim of this section is to show the sparsity pattern and the comparison of standard state space form and it's descriptor form of the models that are introduced in the above Sections. Figure 3.1 shows a power system model from [20]. Figure 3.1a shows the sparsity pattern of the unreduced Jacobian matrix A when the model is put as a descriptor form in (3.10) where the dimension of this matrix is 7135. When this model is put into standard state space form (3.9), the dimension of J (Figure 3.1b) is only 606, but the number of non-zero elements is larger than that of the descriptor form. Figure 3.2 presents the sparsity pattern of the piezo-mechanical model considered from [34]. Figure 3.2a, Figure 3.2b and Figure 3.2c are showing the sparsity pattern of the mass, damping and stiffness matrices respectively in (3.11). Figure 3.2d and Figure 3.2e are showing the sparsity pattern of the matrices E_1 and $A = \begin{bmatrix} J_a & J_b \\ J_c & J_d \end{bmatrix}$, respectively when the model is represented in descriptor form (3.15). Figure 3.2f shows the sparsity pattern of the matrix $J = (J_a - J_b J_d^{-1} J_c)$ of the state space form of the piezo-mechanical model in (3.16). From Figure 3.2d and Figure 3.2e one can easily identify that dealing with the representation of the piezo-mechanical model in (3.15) is convenient in the MOR algorithms.

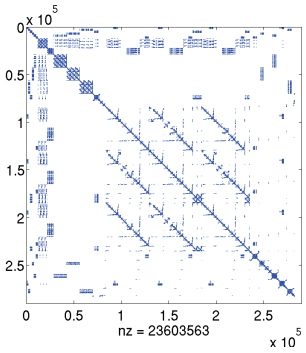


(a) System matrix A in (3.10) (b) System matrix J in (3.9)

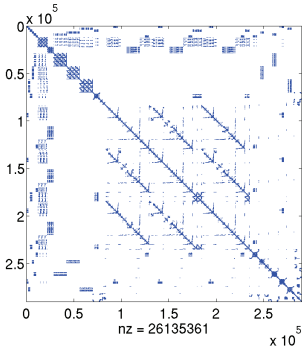
Figure 3.1: Sparsity patterns of the power system model (data source: [20])



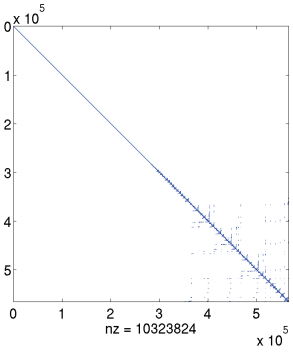
(a) Mass matrix M



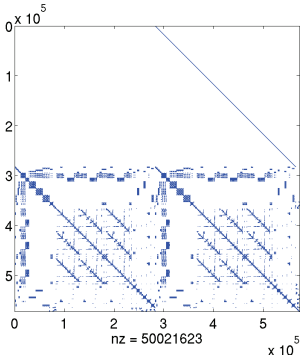
(b) Damping matrix D



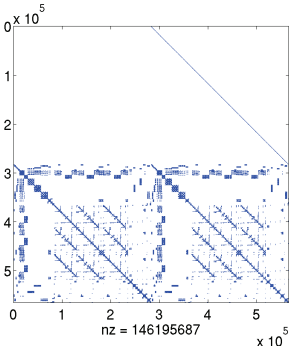
(c) Stiffness matrix K



(d) Matrix E_1



(e) Matrix A



(f) Matrix J

Figure 3.2: Sparsity patterns of the piezo-mechanical model (data source: [34])

CHAPTER

FOUR

GENERALIZED SPARSE LRCF-ADI METHODS FOR SOLVING LYAPUNOV EQUATIONS OF LARGE SPARSE SYSTEMS

It is well known that the most expensive part in approximating a large system via balanced truncation is the solution of two Lyapunov equations for the controllability Gramian and observability Gramian. Many books, journals and research papers introduce numerous techniques or methods to solve Lyaponuv equations exactly as well as approximately. Among these methods, *Bartels-Stewart methods* [6], *Hammarling methods* [32], *Krylov subspace methods* [39], *ADI methods* [67], *low rank iterative methods* [62], and later *low rank Cholesky factor ADI methods* [47, 39, 7] that have been developed over the last decades are remarkable. But in real life, problems are not confined with a particular size or structured indeed. As the methods become more efficient, the size of the feasible and the desired problems grow. As a result research is still going on in developing more and more efficient techniques. This chapter studies one such method; *sparse low rank Cholesky factor ADI* introduced in [20] and this idea is extended to the generalized case.

The first section gives a brief overview of past work on ADI methods. *GSLRCF-ADI* (generalized sparse low rank Cholesky factor ADI) algorithm is discussed in the second section. Finally the third section contains some numerical results.

4.1 A brief review of ADI methods

For convenience, let us recall the following LTI continuous-time system

$$\begin{aligned} E_1\dot{x}(t) = &\quad Jx(t) + Bu(t), \qquad x(t_0) = x_0, \\ y(t) = &\quad Cx(t) + D_a u(t), \end{aligned} \tag{4.1}$$

where E_1 is a nonsingular matrix, $J = (J_a - J_b J_d^{-1} J_c)$, $B = B_1 - J_b J_d^{-1} B_2$, $C = C_1 - C_2 J_d^{-1} J_c$ and $D_a = -C_2 J_d^{-1} B_2$. The sub-matrices J_a, J_b, J_c, J_d, B_1, B_2, C_1, C_2 are defined in the earlier chapter. $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^p$, and $y(t) \in \mathbb{R}^m$ are state, input and output vectors, respectively. When $E_1 = I$ (identity matrix), clearly (4.1) is a standard state space system. However one can write (4.1) as a descriptor form:

$$\begin{aligned} \underbrace{\begin{bmatrix} E_1 & 0 \\ 0 & 0 \end{bmatrix}}_{E} \begin{bmatrix} \dot{x}(t) \\ \dot{\varphi}(t) \end{bmatrix} = &\quad \underbrace{\begin{bmatrix} J_a & J_b \\ J_c & J_d \end{bmatrix}}_{A} \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix} + \underbrace{\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}}_{F} u(t), \\ y(t) = &\quad \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{L} \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix}. \end{aligned} \tag{4.2}$$

If the system in (4.1) is stable, the controllability Gramian $\mathcal{W}_c \in \mathbb{R}^{n\times n}$ and the observability Gramian $\mathcal{W}_o \in \mathbb{R}^{n\times n}$ are the unique solutions of the following pair of continuous-time algebraic Lyapunov equations (CALE) [31]:

- standard case (when $E_1 = I$):

$$J\mathcal{W}_c + \mathcal{W}_c J^T = -BB^T \quad \text{[controllability CALE]}, \tag{4.3a}$$

$$J^T\mathcal{W}_o + \mathcal{W}_o J = -C^T C \quad \text{[observability CALE]}, \tag{4.3b}$$

- generalized case:

$$J\mathcal{W}_c E_1^T + E_1\mathcal{W}_c J^T = -BB^T \quad \text{[controllability CALE]}, \tag{4.4a}$$

$$J^T\mathcal{W}_o E_1 + E_1^T\mathcal{W}_o J = -C^T C \quad \text{[observability CALE]}. \tag{4.4b}$$

Note that controllability Lyapunov equations and observability Lyapunov equations are dual of each other. Therefore, we emphasize on the controllability Lyapunov equations through out the chapter.

The ADI iteration was first introduced by the authors in [46] to solve parabolic and elliptic difference equations. In general formulation, consider a linear system of the form

$$(H + V)y = b, \tag{4.5}$$

where b is an $n-$dimensional vector. The matrices H and V are commuting, positive definite and symmetric. H may be represented as the discretization in $x-$direction, whereas V may be represented as the discretization in $y-$direction. Usually, the ADI method is proceeding in two stages. At first a half-step is to take in the direction H implicitly and in the direction V explicitly, then a half-step is to take in the direction V implicitly and in the direction H explicitly. Now applying this idea to the general problem (4.5) the equations for the ADI method can be written as:

$$\begin{aligned} (H + \mu_j I)y_{j-\frac{1}{2}} &= (\mu_j I - V)y_{j-1} + b, \\ (V + \mu_j I)y_j &= (\mu_j I - H)y_{j-\frac{1}{2}} + b, \end{aligned} \tag{4.6}$$

where $\mu_1, \mu_2, \cdots$ are called ADI shift parameters. Author in [67] shows that, one can considers (4.3a) as a ADI model problem, and the (approximate) solution of $\mathcal{W}_c$ can be obtained by following iterations [31]:

$$(J + \mu_i I)\mathcal{W}_c^{i-\frac{1}{2}} = -BB^T - \mathcal{W}_c^i(J^T - \mu_i I), \tag{4.7a}$$

$$(J + \mu_i I)(\mathcal{W}_c^i)^* = -BB^T - (\mathcal{W}_c^{i-\frac{1}{2}})^*(J^T - \mu_i I), \tag{4.7b}$$

where $\mathcal{W}_c^0 = 0$, and ADI shift parameters $P = \{\mu_1, \mu_2, \cdots\} \subseteq \mathbb{C}^-$ (open left half complex plane). Combining (4.7a) and (4.7b), one can find the following single equation:

$$\begin{aligned} \mathcal{W}_c^i =& (J - \mu_i^* I)(J + \mu_i I)^{-1}\mathcal{W}_c^{i-1}[(J - \mu_i^* I)(J + \mu_i I)^{-1}]^* \\ & - 2\rho_i(J + \mu_i I)^{-1}BB^T(J + \mu_i I)^{-*}, \end{aligned} \tag{4.8}$$

where $\rho_i = \text{real}(\mu_i)$. The spectral radius of the matrix $\prod_{i=1}^{l} \frac{(J-\mu_i^* I)}{(J+\mu_i I)}$, denoted by ρ_{sr}, determines the rate of ADI convergence, where l is the number of shifts used [31]. Note that since the matrix J is asymptotically stable, $\rho_{sr} < 1$ [31]. Smaller ρ_{sr} guarantees faster convergence [20]. The minimization of ρ_{rs} with respect to shift parameters μ_i is called the ADI *mini-max* problem [47]:

$$\{\mu_1, \cdots, \mu_l\} = \arg \min_{\mu_1, \mu_2, \cdots, \mu_l} \max_{\lambda \in \Lambda(J)} \prod_{i=1}^{l} \frac{|\, \mu_i - \lambda \,|}{|\, \mu_i + \lambda \,|}, \tag{4.9}$$

which gives the idea to generate the optimal and suboptimal ADI parameters. Here, $\Lambda(J)$ denotes the spectrum of J. One can use single shift

($\mu_i = \mu; i = 1, 2, \cdots l$) throughout the ADI iteration. Then the ADI iteration reduces to the *Smith method* which can be summarized as follows [31]

$$\mathcal{W}_{cs}^{k} = -2\mu \sum_{j=0}^{k-1} J_{\mu}^{j} B_{\mu} B_{\mu}^{T} (J_{\mu}^{j})^{T}, \tag{4.10}$$

where

$$J_{\mu} = (J - \mu I)(J + \mu I)^{-1},$$
$$B_{\mu} = (J + \mu I)^{-1} B.$$

However, later Penzl showed in [47], although Smith method is convenient in using single shift, the rate of convergence is not up to mark. And he proposed that instead of a single shift, if l number of shifts are used, then the ADI convergence rate is faster. He also added that if one uses more shifts than l the rate of convergence is not improved further (see [47] for more details). Penzl introduces a set of l suboptimal shift parameters which can be chosen by a *heuristic procedure* that will be discussed later.

LRCF-ADI iterations: If we assume that (BB^T) in (4.3a) has lower rank (r) compared to the order of the system (n) ($r \leq p << n$), then the numerical rank of $\mathcal{W}_c^i$ in (4.8) is usually small. Therefore one can compute a low rank factor Z_i, instead of $\mathcal{W}_c^i$ from (4.8), using the outer product

$$\mathcal{W}_c^i = Z_i Z_i^T. \tag{4.11}$$

This version of ADI iteration is known as LR-ADI in [47], CF-ADI in [39], LR-smith in [47], and LRCF-ADI in [7]. The key idea of LRCF-ADI is to use (4.11) in (4.8), and calculate the Cholesky factor Z_i, equating the terms from both sides, which gives

$$Z_i = [(J - \mu_i I)(J + \mu_i I)^{-1} Z_{i-1} \quad \sqrt{-2\rho_i}(J + \mu_i I)^{-1} B]. \tag{4.12}$$

Therefore, the ADI algorithm can be reformulated in terms of the Cholesky factor Z_i as

$$\begin{aligned} Z_1 &= \sqrt{-2\rho_1}(J + \mu_1 I)^{-1} B, && Z_1 \in \mathbb{R}^{n \times p}, \\ Z_i &= [(J - \mu_i I)(J + \mu_i I)^{-1} Z_{i-1} && \sqrt{-2\rho_i}(J + \mu_i I)^{-1} B], \quad Z_i \in \mathbb{R}^{n \times ip}, \end{aligned} \tag{4.13}$$

where $\rho_i = \text{real}(\mu_i) < 0$, with $i = 2, 3, \cdots, i_{\max}$ and $i_{\max}$ is the maximum number of iterations. In this formulation, the number of columns in each iteration is increased by p (the number of columns in B). However, by using a clever reordering the above LRCF-ADI method can be reformulated as

follows (see [31, 39] for details computation):

$$Z_1 = V_1 = \sqrt{-2\rho_1}(J + \mu_1 I)^{-1}B, \tag{4.14}$$

$$V_i = \sqrt{\frac{\rho_i}{\rho_{i-1}}}[V_{i-1} - \gamma_i(J + \mu_i I)^{-1}V_{i-1}], \tag{4.15}$$

$$Z_i = [Z_{i-1} \quad V_i], \quad i = 2, 3, \cdots, i_{max}, \tag{4.16}$$

where $\gamma_i = \mu_i + \mu_{i-1}^*$ and shift parameters are used in cyclic way (see [31] for further details). The whole procedure is summarized in Algorithm 4.1.

Algorithm 4.1 [7] LRCF-ADI (for solving CALE (4.3a))

Input: $J, B, P = \{\mu_1, \mu_2, \cdots, \mu_{i_{max}}\}$.
Output: $Z = Z_{i_{max}}$, such that $ZZ^T \approx \mathcal{W}_c$.
1. $V_1 = \sqrt{-2\mathrm{Re}(\mu_1)}(J + \mu_1 I)^{-1}B$.
2. $Z_1 = V_1$.
for $i = 2 : i_{max}$ **do**
3. $V_i = \sqrt{\frac{\mathrm{Re}(\mu_i)}{\mathrm{Re}(\mu_{i-1})}}(V_{i-1} - (\mu_i + \mu_{i-1}^*)(J + \mu_i I)^{-1}V_{i-1})$.
4. $Z_i = [Z_{i-1} \quad V_i]$.
end for

Sparse low rank Cholesky factor ADI (SLRCF-ADI) methods

The most expensive operations in LRCF-ADI algorithm outlined above are the solutions of linear system with the operators $(J + \mu_i I)$, where $i = 1, 2, \cdots, i_{max}$ in the Steps 1& 3. If the matrix J in the system (4.1) is not sparse or has a poor sparsity pattern, LRCF-ADI is too expensive because of solving a linear system to have the expression $(J + \mu_1 I)^{-1}B$ and $(J + \mu_i I)^{-1}V_{i-1}$ in the Step 1 and Step 3, respectively of the Algorithm 4.1. Under this circumstances, the SLRCF-ADI method introduced in [20], solves the same linear system of equations by operating on the sparse descriptor form of the system. Therefore the computational procedure is more efficient. One can assume that the basic difference between LRCF-ADI and SLRCF-ADI is in solving a linear system for the expression $(J + \mu_1 I)^{-1}B$ and $(J + \mu_i I)^{-1}V_{i-1}$; $\{i = 2, 3, \cdots, i_{max}\}$, respectively in the Algorithm 4.1. The procedure for the SLRCF-ADI is as follows.

The expression $X_1 = (J + \mu_1 I)^{-1}B$ can be assumed as a solution of the linear system

$$(J + \mu_1 I)X_1 = B. \tag{4.17}$$

Since, $J = (J_a - J_b J_d^{-1} J_c)$ and $B = (B_1 - J_b J_d^{-1} B_2)$, (4.17) can be reformed to

$$(J_a - J_b J_d^{-1} J_c + \mu_1 I)X_1 = B_1 - J_b J_d^{-1} B_2. \tag{4.18}$$

With $X_2 = -J_d^{-1}J_cX_1 + J_d^{-1}B_2$, (4.18) implies

$$\begin{bmatrix} J_a + \mu_1 I & J_b \\ J_c & J_d \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}. \tag{4.19}$$

System of linear equations in (4.19) can easily be solved, for X_1 by using any convenient linear system solver. To compute $(J + \mu_i I)^{-1}V_{i-1}$, in Step 3 solve the following system of linear equations:

$$\begin{bmatrix} J_a + \mu_i I & J_b \\ J_c & J_d \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} V_{i-1} \\ 0 \end{bmatrix}, \tag{4.20}$$

for X_1. The complete procedure is summarized in the Algorithm 4.2. The

Algorithm 4.2 [20] Sparse LRCF-ADI (for solving CALE (4.3a))

Input: J_a, J_b, J_c, J_d, B_1 and B_2 and $P = \{\mu_1, \mu_2, \cdots, \mu_{i_{max}}\}$.
Output: $Z = Z_{i_{max}}$, such that $ZZ^T \approx \mathcal{W}_c$.

1. Solve (4.19) for X_1.
2. Compute $V_1 = \sqrt{-2\mathrm{Re}(\mu_1)}X_1$.
3. Set $Z_1 = V_1$.

for $i = 2 : i_{max}$ **do**

5. Solve (4.20) for X_1.
6. Determine $V_i = \sqrt{\frac{\mathrm{Re}(\mu_i)}{\mathrm{Re}(\mu_{i-1})}}(V_{i-1} - (\mu_i + \mu_{i-1}^*)X_1)$.
4. $Z_i = [Z_{i-1} \quad V_i]$.

end for

low rank approximate solution Z of the observability Lyapunov equation (4.3b), for observability Gramian ($ZZ^T \approx W_o$), use the same algorithm as above. In that case, for X_1 in Step 1, solve (4.19) by replacing J_a, J_b, J_c, J_d B_1, and B_2 by J_a^T, J_c^T, J_b^T, J_d^T, C_1^T and C_2^T, respectively. And solve (4.20), for X_1 in Step 5 by replacing J_a by J_a^T, J_b by J_c^T, J_c by J_b^T and J_d by J_d^T. See [20] for further details.

4.2 Generalized sparse low rank Cholesky factor ADI (GSLRCF-ADI) methods

This section includes GSLRCF-ADI methods for solving GCALE (4.4) of the system (4.1) by exploiting the sparsity of the descriptor representation in (4.2). An algorithm namely GLRCF-ADI for solving GCALE is stated in [51] (see Chapter 5 for details). But unfortunately this technique is not efficient when index-1 system in (4.2) converts into a state space form in (4.1). Since, in this formulation J is not sparse, as a result computational cost is extreme

high. Moreover, in some cases this formulation is not possible at all for a large-scale system. Under this circumstances we can use the technique stated in SLRCF-ADI methods outlined above.

One can relate LRCF-ADI and GLRCF-ADI algorithm; when controllability GCALE in (4.4a) is solved by GLRCF-ADI, compute $(J + \mu_1 E_1)^{-1}B$ instead of $(J + \mu_1 I)^{-1}B$ in the Step 1 of the Algorithm 4.1. And in the Step 3 of the same algorithm, $(J + \mu_i I)V_{i-1}$ is changed by $(J + \mu_i E_1)(E_1 V_{i-1})$. Expression $(J + \mu_1 E_1)^{-1}B$ is nothing but the solution of the linear system

$$\begin{bmatrix} J_a + \mu_1 E_1 & J_b \\ J_c & J_d \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \tag{4.21}$$

for X_1. Likewise, X_1, the solution of the linear system

$$\begin{bmatrix} J_a + \mu_i E_1 & J_b \\ J_c & J_d \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} E_1 V_{i-1} \\ 0 \end{bmatrix}, \tag{4.22}$$

where $i = 2, 3, \cdots, i_{max}$, is the expression $(J + \mu_i E_1)(E_1 V_{i-1})$. Now, computing X_1 from (4.21) and (4.22), respectively and plug in the Step 2 and Step 6, respectively in the Algorithm 4.2, the low rank approximation of controllability Gramian of GCALE can be obtained. This is known as GSLRCF-ADI in this paper.
The complete procedure of GSLRCF-ADI method is shown in Algorithm-4.3 .

Algorithm 4.3 Generalized sparse LRCF-ADI (GSLRCF-ADI) method for solving GCALE (4.4a)

Input: E_1, J_a, J_b, J_c, J_d, B_1 and B_2 and $P = \{\mu_1, \mu_2, \cdots, \mu_{i_{max}}\}$.
Output: $Z = Z_{i_{max}}$ such that $ZZ^T \approx \mathcal{W}_c$.

1. Solve (4.21) for X_1.
2. Compute $V_1 = \sqrt{-2\mathrm{Re}(\mu_1)}X_1$.
3. Set $Z_1 = V_1$.

for $i = 2 : i_{max}$ **do**

5. Solve (4.22) for X_1.
6. Determine $V_i = \sqrt{\frac{\mathrm{Re}(\mu_i)}{\mathrm{Re}(\mu_{i-1})}}(V_{i-1} - (\mu_i + \mu_{i-1}^*)X_1)$.
4. Set $Z_i = [Z_{i-1} \quad V_i]$.

end for

When solving (4.4b) for the low rank approximation Z of the observability Gramian $\mathcal{W}_o$, such that $ZZ^T \approx \mathcal{W}_o$, with the Algorithm 4.3, inputs are E_1, J_a, J_b, J_c, J_d, C_1 and C_2. Replace E_1, J_a, J_b, J_c, J_d, B_1 and B_2 in (4.21) by E_1^T, J_a^T, J_c^T, J_b^T, J_d^T, C_1^T and C_2^T respectively and solve it for X_1. Then this X_1 put in step 2 of the Algorithm 4.3. Similarly for X_1 in Step 6, solve (4.22) replacing E_1 by E_1^T, J_a by J_a^T, J_b by J_c^T, J_c by J_b^T and J_d by J_d^T.

4.2.1 ADI shift parameter selection

The approximate solutions of Lyapunov equations in (4.4) by GSLRCF-ADI methods, mentioned above converge to the exact solutions if a system is asymptotically stable [62]. However, proper selection of ADI parameters is one of the crucial task for fast convergence. It is mentioned earlier that ADI optimal parameters μ_i where $i = \{1, 2, \cdots, l\}$ are the solutions of the discrete rational mini-max problem (4.9), where Λ is the full spectrum of $\bar{J} = (E_1^{-1}J)$ [67]. Since the entire spectrum of $\bar{J}$ may not be easily available, instead of solving (4.9), one can solve the following problem to compute ADI optimal shifts [39].

$$\min_{\mu_1,\mu_2,\cdots,\mu_l} \max_{\lambda\in\mathcal{R}} \prod_{i=1}^{l} \frac{\mid \mu_i - \lambda \mid}{\mid \mu_i + \lambda \mid}, \tag{4.23}$$

where $\mathcal{R}$ is the region in the open left half-plane. The parameters $\{\mu_i\}$ for $i = \{1, 2, \cdots, l\}$ are called optimal if they satisfy mini-max problem (4.23). But in practice, we compute near-optimal parameters, which are known as ADI *suboptimal parameters* or ADI *suboptimal shifts* [47]. Several techniques are proposed in different papers to compute ADI shift parameters, see [66, 59, 47, 9] for example. Many papers investigated the performance of these techniques (see [39, 62, 8, 7, 58, 51, 7] and the references therein).

Here, we consider the technique proposed in [47], called *heuristic procedure*. The key ideas of heuristic procedure are as follows:

- Determine $\mathcal{R}$ based on the approximation of the dominant spectrum (in magnitude) of $\bar{J}$. Therefore, compute a set $\mathcal{R}_+$ containing K_+ largest (magnitude) and a set $(1/\mathcal{R}_-)$ ($\mathcal{R}_-$ is the set, containing K_- largest Ritz-values of $\text{inv}(\bar{J})$) containing K_- smallest (magnitude) stable Ritz-values of $\bar{J}$ via the Arnoldi process. Then apply $\mathcal{R} = \mathcal{R}_+ \cup (1/\mathcal{R}_-)$.
- Choose the suboptimal ADI parameters $P = \{\mu_1, \cdots, \mu_l\}$ among the elements of $\mathcal{R}$ since the function

$$S_p(\lambda) = \prod_{i=1}^{l} \frac{\mid \mu_i - \lambda \mid}{\mid \mu_i + \lambda \mid}, \tag{4.24}$$

  becomes small over $\Lambda(\bar{J})$ if there is one of the shifts μ_i in the neighborhood of each eigenvalue [47].

The complete heuristic procedure is shown in the Algorithm 4.4.

Algorithm 4.4 [47] Heuristic procedure to compute suboptimal ADI shifts for solving (4.4)

Input: E_1, J_a, J_b, J_c, J_d, l, K_+, and K_-.
Output: $P = \{\mu_1, \mu_2, \cdots, \mu_l\}$.

1. Consider $\bar{J} = E_1^{-1}(J_a - J_b J_d^{-1} J_c)$
2. Choose a vector $v \in \mathbb{R}^n$, which may be random or a particular choice.
3. Perform K_+ steps of the Arnoldi process w.r.t. $(\bar{J}, v)$ and compute the set of Ritz-values $\mathcal{R}_+$.
4. Perform K_- steps of the Arnoldi process w.r.t. $(\text{inv}(\bar{J}), v)$ and compute the set of Ritz-values $\mathcal{R}_-$.
5. $\mathcal{R} = \{\mu_1, \mu_2, \cdots, \mu_{K_+ + K_-}\} = \mathcal{R}_+ \cup (1/\mathcal{R}_-)$.
6. IF $\mathcal{R} \not\subset \mathbb{C}^-$, STOP.
7. Detect i with $\max_{\lambda \in \mathcal{R}} S_{\{\mu_i\}}(\lambda) = \min_{\mu \in \mathcal{R}} \max_{\lambda \in \mathcal{R}} S_{\{\mu\}}(\lambda)$ and initialize

$$P = \begin{cases} \{\mu_i\} & : \text{when } \mu_i \text{ real}, \\ \{\mu_i, \mu_i^*\} & : \text{when } \mu_i \text{ complex}. \end{cases}$$

while card$(P) < l$ **do**

8. Detect i with $S_p(\mu_i) = \max_{\lambda \in \mathcal{R}} S_p(t)$ and set

$$P = \begin{cases} P \cup \{\mu_i\} & : \text{when } \mu_i \text{ real}, \\ P \cup \{\mu_i, \mu_i^*\} & : \text{when } \mu_i \text{ complex}. \end{cases}$$

end while

4.2.2 Stopping criteria

There are several criteria in the literatures to stop ADI iterative methods (see [51] and references therein for details). However we consider *residual based criteria* [51] to stop GSLRCF-ADI algorithm. Let us consider the residual of the GCALE (4.4a) be

$$\mathfrak{R} = J Z_i Z_i^T E_1^T + E_1 Z_i Z_i^T J^T + B B^T, \quad i = 1, \cdots, i_{max} \tag{4.25}$$

where, $J = J_a - J_b J_d^{-1} J_c$ and $B = B_1 - J_b J_d^{-1} B_2$. Then we use the expression

$$\frac{\|\mathfrak{R}\|_2}{\|BB^T\|_2} \leq tol, \tag{4.26}$$

to stop the GSLRCF-ADI iterations by choosing appropriate tolerance based on the problem . Note that, we use power iteration to compute the approximate 2-norm of $\mathfrak{R}$, which reduces the computational costs extremely (see [51] for further details). A similar procedure is followed, for solving observability GCALE (4.4a).

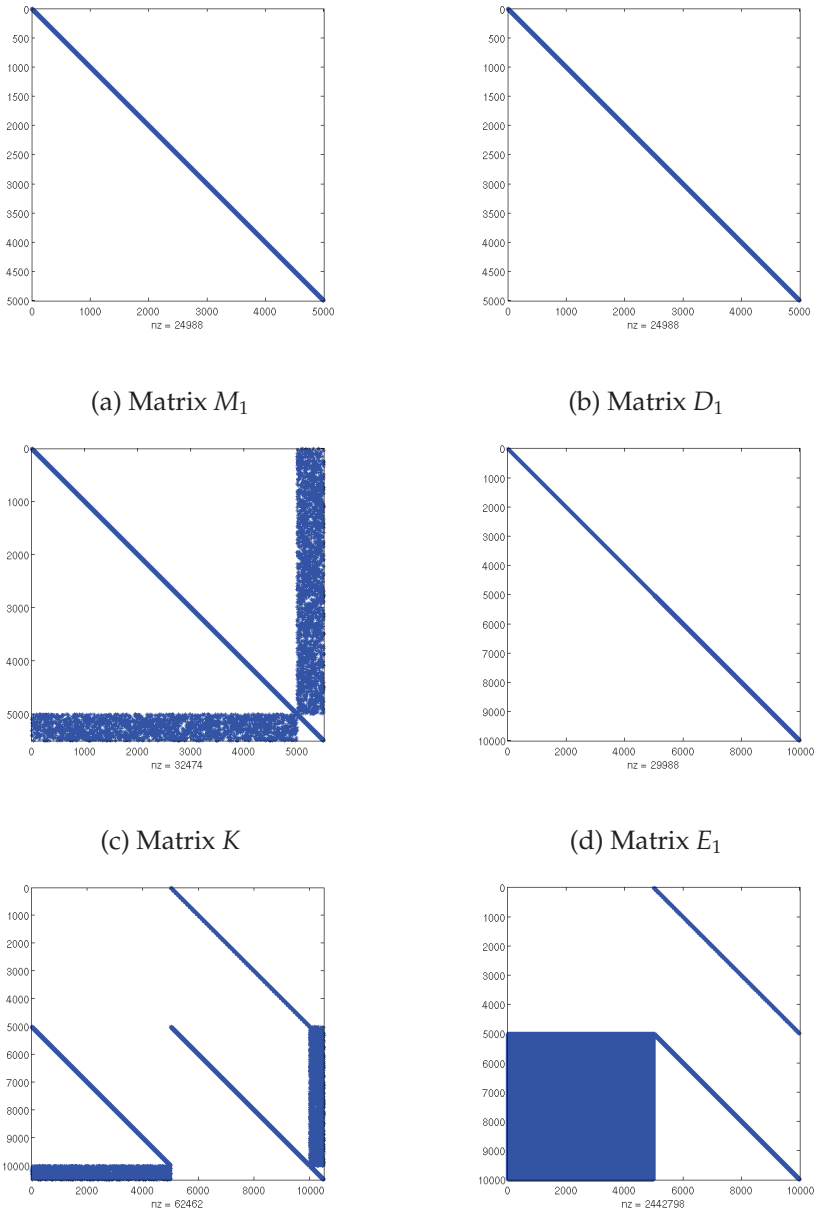
4.3 Numerical examples

To verify the accuracy and efficiency of our algorithm (GSLRCF-ADI), some results of numerical experiments are discussed in this section. A set of artificial data as given below is considered for the piezo-mechanical model (3.12). Where possible, results are compared with the results obtained by an exact solver (using MATLAB function LYAP) and GLRCF-ADI algorithm (Algorithm 5.1 in [51]).
Artificial data:

```
n = number of state variables,
l = number of algebraic variables,
nin = number of inputs,
nout = number of outputs,
I = speye(n);
M1 = .5*I+spdiags(-0.2*ones(n,1),2,n,n)+
     spdiags(-0.2*ones(n,1),-2,n,n)+
     spdiags(0.2*ones(n,1),4,n,n)+
     spdiags(0.2*ones(n,1),-4,n,n);
K11 = spdiags(5*ones(n,1),0,n,n)+
         spdiags(-1*ones(n,1),2,n,n)+
         spdiags(-1*ones(n,1),-2,n,n)+
         spdiags(2*ones(n,1),4,n,n)+
         spdiags(2*ones(n,1),-4,n,n);
D1 = mu*M+nu*K11;% mu=0.005,nu=1
K44 = spdiags(-5*ones(l,1),0,l,l)+
      spdiags(ones(l,1),2,l,l)+
      spdiags(ones(l,1),-2,l,l)+
      spdiags(-2*ones(n,1),4,l,l)+
      spdiags(-2*ones(n,1),-4,l,l);
K12 = sprand(n,l,den);%den=density
E1 = [I spalloc(n,n,0);spalloc(n,n,0) M1];
Ja = [spalloc(n,n,0) I;-K11-D1]; Jb = [spalloc(n,l,0);-K12];
Jc = [-K12'spalloc(l,n,0) ]; Jd = -K44;
B1=[spalloc(n,nin,0);spdiags(ones(n,1),0,n,nin)];
B2=spdiags(ones(l,1),0,l,nin);
C1=[spdiags(ones(n,1),0,nout,n) spalloc(nout,n,0)];
C2=spdiags(ones(l,1),0,nout,l);
```

Changing n and l, one can choose the dimension of the system. Likewise nin and nout can change the number of inputs and outputs of the system respectively. And den can be used to change the number of nonzero in the system matrix $A = \begin{bmatrix} J_a & J_b \\ J_c & J_d \end{bmatrix}$. The sparsity patterns of the system matrices

are shown in Figure 4.1. Figure 4.2 shows the convergence histories of



(a) Matrix M_1 (b) Matrix D_1

(c) Matrix K (d) Matrix E_1

(e) Matrix A (f) Matrix J

Figure 4.1: Sparsity patterns of the artificial data

GSLRCF-ADI methods. Here we consider a model of order 10500 (descriptor form), number of shifts is 10, and residual tolerance is 10^{-15}. Figure 4.2a shows the convergence history for the controllability Lyapunov equation,

whereas, Figure 4.2b is for the observability Lyapunov equation. One can notice that very fast convergence is obtained with the iteration steps 15 and 16, respectively in both cases.

Figure 4.3, shows the accuracy of the GSLRCF-ADI method. Considering a model of dimension 2100, we compute the controllability and observability Gramian, using an exact solver and GSLRCF-ADI respectively. The relative error for the controllability Gramian ($\frac{\|X_c - \mathcal{W}_c\|_2}{\|X_c\|_2}$) and the observability Gramian ($\frac{\|X_o - \mathcal{W}_o\|_2}{\|X_o\|_2}$) are 2.7159×10^{-13} and 3.7648×10^{-13} respectively. From Figure 4.3a, it can easily be observed that the largest singular values of the controllability Gramians for both (exact and GSLRCF-ADI) solvers match accurately. Figure 4.3b depicts the same histories for the observability Gramian. In the case of computational time comparison, GSLRCF-ADI method is more efficient than the exact solver (since GSLRCF-ADI method works on the sparse descriptor form of the system). Computational times for the controllability Gramian, for different systems via exact solver and GSLRCF-ADI method respectively are shown in Table 4.1 with the relative error, the comparison is also shown in Figure 4.4.

While comparing GSLRCF-ADI and GLRCF-ADI methods, we observe that both solvers give the same results for a piezo-mechanical model that we have considered. The only difference is in the computational time. Here, we consider a model of dimension 10500 (descriptor form), and compute the time for different iterations when GLRCF-ADI and GSLRCF-ADI is performed for the low rank approximation of the controllability Gramian. The result is also shown in Table 4.2. Figure 4.5 shows that in each step the computational time for GLRCF-ADI is more than seven times that of GSLRCF-ADI. But the scenario is different when we compute the time of a model where the number of nonzero in the system matrices J and A does not vary remarkably. In that case Figure 4.6 shows that GSLRCF-ADI is not more efficient than GLRCF-ADI. This figure is generated from Table 4.3.

To compare the three solvers (exact, GLRCF-ADI, GSLRCF-ADI), we measure the computational time for solving the controllability Lyapunov equation for a model of dimension 2100 with the three solvers. Table 4.4 describes the computational time, and the respective figure is the bar chart in the Figure 4.7.

Computing an appropriate number of ADI shifts is a vital issue to accelerate the rate of convergence. Figure 4.8 illustrates this issue. The iteration with ten shift parameters gives better convergence, while this number is increased, the convergence gets worse. May be less than ten shifts would

dimension	Exact solver (sec)	GSLRCF-ADI (sec)	relative error
200	0.44	0.49	4.8773e-09
400	3.94	0.62	3.0492e-10
600	13.28	0.81	1.9979e-10
800	33.90	1.14	2.1388e-10
1000	65.42	1.38	3.5875e-10
1200	118.02	1.79	5.4358e-10
1400	194.88	2.73	5.1201e-10

Table 4.1: Computational time for different dimensional systems by exact solver and GSLRCF-ADI and the relative errors

iteration	GLRCF-ADI (sec)	GSLRCF-ADI (sec)
5	114.07	14.67
10	194.87	28.38
15	335.67	51.47
20	439.75	63.38
25	596.70	79.61
30	702.53	93.84

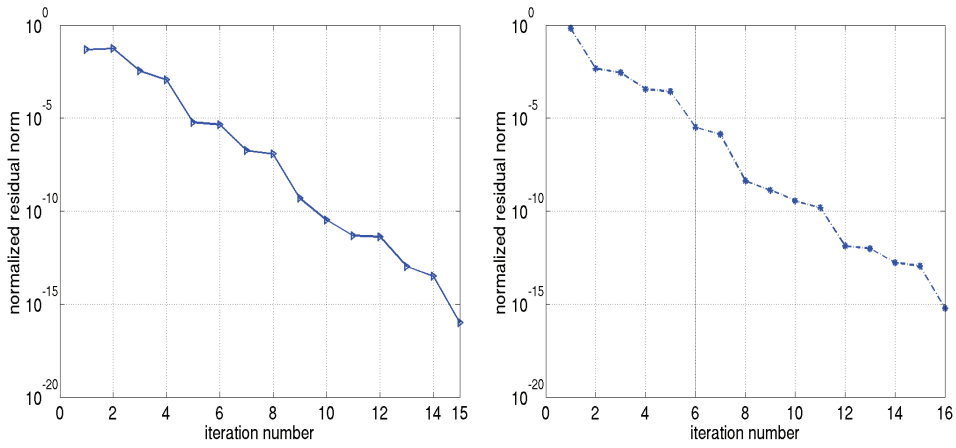
Table 4.2: Computational time of low rank approximate of controllability Gramian by GLRCF-ADI and GSLRCF-ADI (when system matrix J is sparse)

iteration	GLRCF-ADI (sec)	GSLRCF-ADI (sec)
10	11.37	9.79
20	17.94	18.71
30	22.99	25.16
40	35.54	43.60
50	47.01	54.30

Table 4.3: Computational time for low rank factor of controllability Gramian by GLRCF-ADI and GSLRCF-ADI (when system matrix J is not sparse)

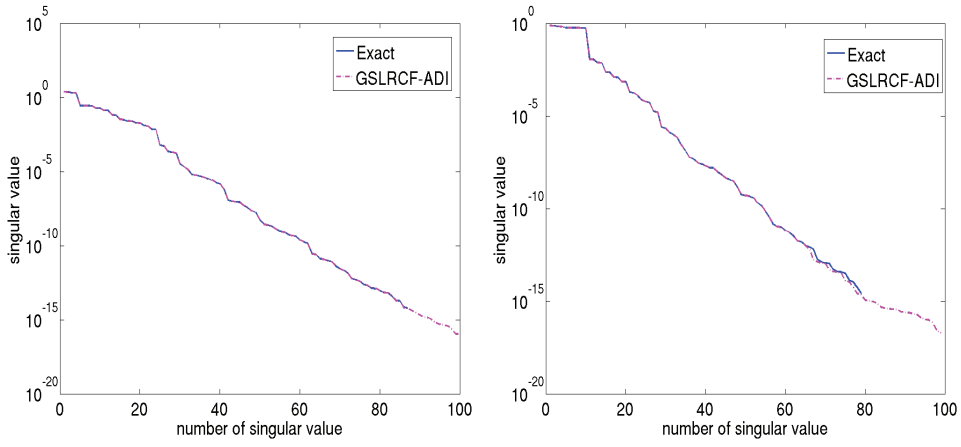
Exact solver (sec)	GLRCF-ADI (sec)	GSLRCF-ADI (sec)	relative error
696.98	92.51	18.38	1.4518e-09

Table 4.4: Computational time of exact solver, GLRCF-ADI and GSLRCF-ADI



(a) Controllability Gramian (b) Observability Gramian

Figure 4.2: GSLRCF-ADI convergence histories (normalized residual norm)



(a) Controllability Gramian (b) Observability Gramian

Figure 4.3: Singular values of the Gramians computed by exact solver and GSLRCF-ADI

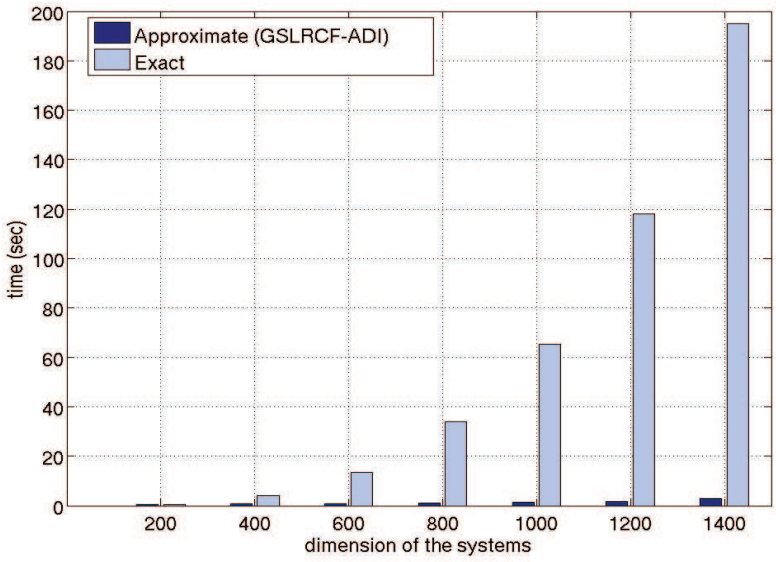


Figure 4.4: Computational time for different dimensional systems by exact solver and GSLRCF-ADI

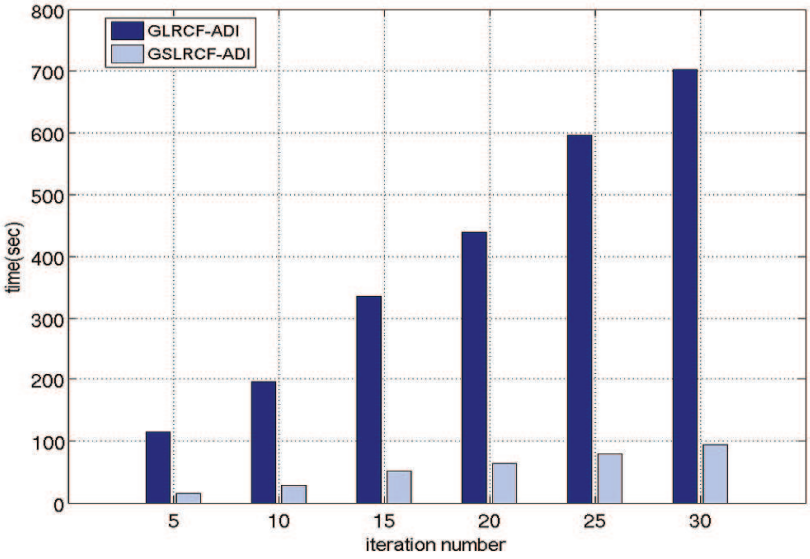


Figure 4.5: Computational time of low rank approximate of controllability Gramian by GLRCF-ADI and GSLRCF-ADI

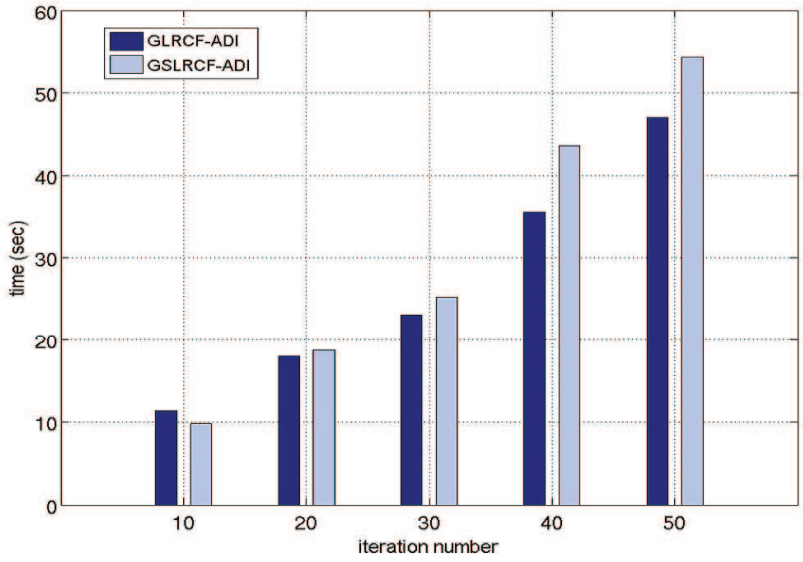


Figure 4.6: Computational time for low rank factor of controllability Gramian by GLRCF-ADI and GSLRCF-ADI (when system matrix J is not sparse)

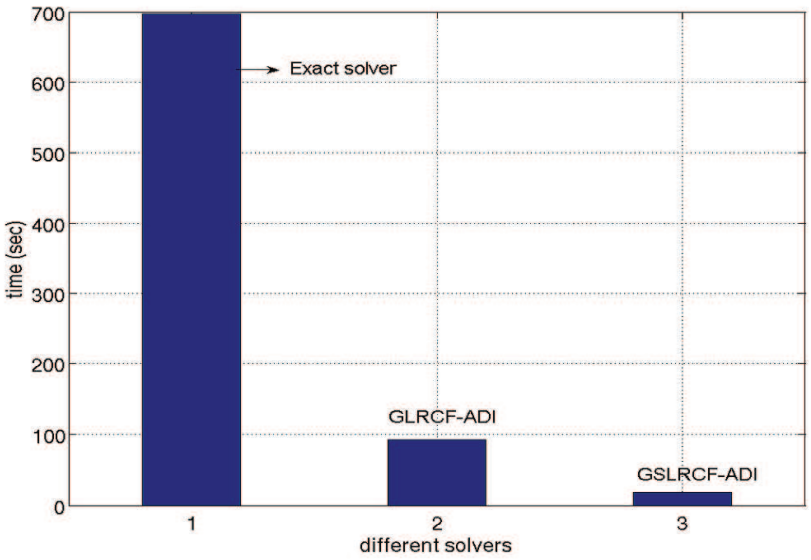


Figure 4.7: Computational time for computing controllability Gramian by the exact solver, GLRCF-ADI and GSLRCF-ADI

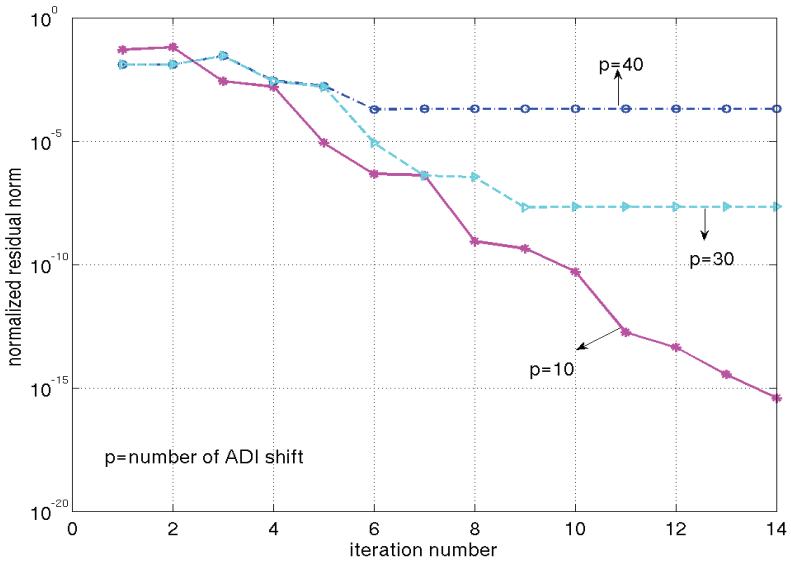


Figure 4.8: Different number of shifts and corresponding convergence history of the low rank approximate solution of the controllability Gramian by GSLRCF-ADI

not show any better result as well.

CHAPTER

FIVE

MODEL REDUCTION OF LARGE SPARSE PIEZO-MECHANICAL SYSTEMS BY BALANCED TRUNCATION

In the first chapter, we have mentioned some very well known methods and respective references to approximate a large-scale continuous-time LTI system. We also have mentioned the main objective of this thesis, i.e., developing an algorithm for MOR of a large sparse peizo-mechanical systems based on Balanced truncation (BT). This chapter includes the main algorithm of this thesis. The first section contains a brief discussion of BT from the literature, for example [1, 28, 8, 71, 11, 17]. While the second section illustrates our algorithm for model order reduction based on BT, for a class of index-1 DAE systems. And in the last section, numerical results obtained by the algorithm are explained.

5.1 Balanced truncation for model reduction

Roughly speaking, truncation methods of model order reduction seek to remove or truncate, unimportant states from the state space models. One can detect these states easily when a system is balanced. A model reduction method is known as *balanced truncation* when it applies the truncation operation to a balanced model. To make it clear, once again, recall a generalized continuous-time LTI system,

$$\begin{aligned} E_1\dot{x}(t) &= J\dot{x}(t) + Bu(t), \\ y(t) &= C\dot{x}(t) + D_a u(t), \end{aligned} \tag{5.1}$$

where $J = J_a - J_b J_d^{-1} J_c$, $B = B_1 - J_b J_d^{-1} B_2$, $C = C_1 - C_2 J_d^{-1} J_c$ and $D_a = -C_2 J_d^{-1} B_2$, are matrices and $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^p$, $y(t) \in \mathbb{R}^m$ are vectors. (5.1) can be put as a descriptor form:

$$\begin{aligned} \underbrace{\begin{bmatrix} E_1 & 0 \\ 0 & 0 \end{bmatrix}}_{E} \begin{bmatrix} \dot{x}(t) \\ \dot{\varphi}(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} J_a & J_b \\ J_c & J_d \end{bmatrix}}_{A} \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix} + \underbrace{\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}}_{F} u(t), \\ y(t) &= \underbrace{\begin{bmatrix} C_1 & C_2 \end{bmatrix}}_{L} \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix}. \end{aligned} \tag{5.2}$$

The controllability Gramian $\mathcal{W}_c$ and and observability Gramian $\mathcal{W}_o$, which are the solutions of two Lyapunov equations play a crucial rule in energy passing of the respective systems, that can be illustrated (in the case of standard syatem) as follows [31]. Define two functions Q_c and Q_o as

$$Q_c = \min_{x(-\infty)=0,\, x(0)=x} \| u(t) \|^2, \qquad t \leq 0, \tag{5.3}$$

$$Q_o = \| y(t) \|^2, \qquad x(0) = x_0, \quad t \geq 0. \tag{5.4}$$

The quantity Q_c is the minimum energy required to drive the system from the zero state at $t = -\infty$ to the state x at $t = 0$ [31]. On the other hand, Q_o is the energy obtained by observing the output with the initial state x_0 under no input [31].

The function Q_c and Q_o are related to the controllability gramian W_c and observability gramian W_o respectively as follows [8]

$$Q_c = x^T \mathcal{W}_c^{-1} x, \tag{5.5}$$

$$Q_o = x_0^T \mathcal{W}_o x_0, \tag{5.6}$$

where $\mathcal{W}_c$ and $\mathcal{W}_o$ be the reachability and observability Gramians, respectively of a asymptotically stable and minimal system. It follows from the above relations that, the states which are difficult to reach, i.e., require a large energy Q_c, are spanned by the eigenvectors of W_c corresponding to the small eigenvalues [31]. Moreover, the states which are difficult to observe, i.e., yield a small observation energy Q_o, are spanned by eigenvectors of W_o corresponding to the small eigenvalues (see [31]). Hence relations (5.5) and (5.6) yield a way to evaluate the degree of reachability and the degree of observability for the states of the given system [31]. We can obtain a reduced model by truncating the states which are difficult to reach or difficult to observe. However, it is not necessary that states which are

difficult to reach are difficult to observe as well, and vice-versa [1]. Hence, the following question arises: for a given system , does there exist a basis where states are difficult to reach and simultaneously difficult to observe? Yes, it is possible, when a realization is balanced. Then reduced order model can be obtained using the following theorem.

Theorem 8 ([31]):

An asymptotically stable and minimal system $\mathbf{\Sigma} = (E_1, J, B, C, D_a)$ has the balanced realization $(E_b, J_b, B_b, C_b, D_a)$ where $E_b = \begin{bmatrix} E_r & E_{12} \\ E_{21} & E_{22} \end{bmatrix}$, $J_b = \begin{bmatrix} A_r & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $B_b = \begin{bmatrix} B_r \\ B_2 \end{bmatrix}$, $C_b = \begin{bmatrix} C_r \\ C_2 \end{bmatrix}$ with $\mathcal{W}_c = \mathcal{W}_o = \text{diag}(\Sigma_1, \Sigma_2)$, where

$$\begin{aligned} \Sigma_1 &= \text{diag}(\sigma_1 I_{m_1}, \cdots, \sigma_k I_{m_k}), \qquad k < n, \\ \text{and} \quad \Sigma_2 &= \text{diag}(\sigma_{k+1} I_{m_{k+1}} \cdots, \sigma_N I_{m_N}). \end{aligned}$$

The multiplicity of σ_i is m_i; $i = 1, 2, \cdots, N$ and $m_1 + m_2 + \cdots + m_N = n$. Then the reduced order model $\mathbf{\Sigma_r} = (E_r, A_r, B_r, C_r, D_a)$ obtained by balance truncation is asymptotically stable, minimal and satisfies

$$\| \mathbf{\Sigma} - \mathbf{\Sigma_r} \|_{\mathcal{H}_\infty} \leq 2(\sigma_{k+1} + \cdots + \sigma_N). \tag{5.7}$$

The equality holds if Σ_2 contains only σ_N. ◊

The above theorem states that if the neglected Hankel singular values [2] are small, then the system $\mathbf{\Sigma}$ and $\mathbf{\Sigma_r}$ are guaranteed to be closed [31]. Note that (5.7) is a *priori* error bound [31]. Therefore considering a suitable error tolerance, we can decide how many states should be truncated to form a reduced model.

This balancing method described above is know as *Lyapunov balancing* [31], since it is involved with the solution of two Lyapunov equations. Besides this, there exist other types of balancing methods, such as *stochastic balancing*, *bounded real balancing*, *positive real balancing* and *frequency weighted balancing*. These methods are discussed in [27, 68, 26, 13, 18, 70, 30]. However all the methods are briefly introduced in [30].

5.2 Square root method (SRM) for reduced order models

For model reduction via balanced truncation, internal balancing of a given model is a preliminary task. If a system is stable and both controllable and

observable, then there exists a transformation such that the transformed controllability and observability Gramians are equal and diagonal [24]. To have such a transformation see [11] in details with examples. Here, we discuss a method, known as square root method (SRM) [64], to compute a reduced order model for a stable non-minimal system. In this method one can directly obtain a reduced balanced system without balancing the whole system before truncation. More precisely, in this way both balancing and truncation are yielded simultaneously. Let $\mathbf{\Sigma} = (J, B, C, D_a)$ be a stable standard state space system, which is not necessarily be both controllable and observable (i.e., minimal). A balanced reduced order model $\mathbf{\Sigma}_r = (J_r, B_r, C_r, D_a)$ can be obtained by performing the following steps:

- *Step-1:* Compute controllability Gramian W_c, and observability Gramian W_o by solving two Lyapunov equations in (4.3a) and (4.3b), respectively.
- *Step-2:* Let $W_c = L_c L_c^T$ and $W_o = L_o L_o^T$ be the Cholesky factorization of two Gramians respectively. This is always possible since W_c and W_o are always symmetric and positive definite matrices.
- *Step-3:* Let singular value decomposition of $L_o^T L_c$ be $L_o^T L_c = USV^T$. The singular values of $L_o^T L_c$ are indeed the Hankel singular values [31]. Hence it can be written as

$$L_o^T L_c = U \Sigma V^T, \tag{5.8}$$

  where $\Sigma = \text{diag}(\sigma_1 I_{m_1}, \sigma_2 I_{m_2}, \cdots \sigma_N I_{m_N})$, N is the number of distinct Hankel singular values with $\sigma_i > \sigma_{i+1} \geq 0$, m_i is the multiplicity of σ_i and $m_1 + m_2 + \cdots + m_N = n$. Let $\Sigma_1 = \text{diag}(\sigma_1 I_{m_1}, \sigma_2 I_{m_2}, \cdots, \sigma_q I_{m_k})$, $k < n$, and $r = m_1 + m_2 + \cdots + m_k$.
- *Step-4:* Define two transformations $T_L = L_o U_1 \Sigma_1^{-\frac{1}{2}}$ and $T_R = L_c V_1 \Sigma_1^{-\frac{1}{2}}$ where U_1 and V_1 are formed by the leading r columns of U and V, respectively.
- *Step-5:* Reduced order model of order r is obtained by setting

$$J_r = T_L^T J T_R, \quad B_r = T_L^T B, \quad \text{and} \quad C_r = C T_R.$$

It is not difficult to check $T_L^T T_R = I_r$ and hence $T_R T_L^T$ is an oblique projector [31]. Again $W_c T_L = T_R \Sigma_1$ and $W_o T_R = T_L \Sigma_1$ gives

$$T_L^T (J W_c + W_c J^T + B B^T) T_L = J_r \Sigma_1 + \Sigma_1 J_r^T + B_r B_r^T, \tag{5.9}$$

$$T_R^T (J^T W_o + W_o J + C^T C) T_R = J_r^T \Sigma_1 + \Sigma_1 J_r + C_r^T C_r, \tag{5.10}$$

which proves that the reduced order model is balanced and satisfies the *Lyapunov inertia theorem* (see [2] for definition and proposition of Lyapunov inertia theorem) [8]. Thus the reduced order model is asymptotically stable. To make everything more clear let us consider the following example

Example: Find a reduced order model of a system $\mathbf{\Sigma} = (J, B, C, D_a)$, where $J = \begin{bmatrix} -1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0.5 \\ 0.005 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0.5 & 0.005 \end{bmatrix}$ and $D_a = \begin{bmatrix} 0 \end{bmatrix}$ using above algorithm.

Solution:

$$W_c = \begin{bmatrix} 0.9235 & 0.2095 & 0.0015 \\ 0.2095 & 0.6628 & 0.0005 \\ 0.0015 & 0.0005 & 0.0000 \end{bmatrix}, \quad W_o = \begin{bmatrix} 0.5000 & 0.5000 & 0.5012 \\ 0.5000 & 0.5625 & 0.6135 \\ 0.5012 & 0.6135 & 0.7058 \end{bmatrix},$$

$$L_c = \begin{bmatrix} 0.9610 & 0 & 0 \\ 0.2180 & 0.1234 & 0 \\ 0.0016 & 0.0013 & 0.0002 \end{bmatrix}, \quad L_o = \begin{bmatrix} 0.7071 & 0 & 0 \\ 0.7071 & 0.2500 & 0 \\ 0.7089 & 0.4490 & 0.0406 \end{bmatrix},$$

$$\begin{aligned} (U, \Sigma, V) =& \text{svd}(L_o^T L_c) \\ =& \begin{bmatrix} -0.9976 & 0.0692 & 0.0000 \\ -0.0692 & -0.9976 & 0.0018 \\ -0.0001 & -0.0018 & 1.0000 \end{bmatrix} \begin{bmatrix} 0.8814 & 0 & 0 \\ 0 & 0.0254 & 0 \\ 0 & 0 & 0.0000 \end{bmatrix} \\ & \begin{bmatrix} -0.9942 & 0.1071 & 0.0002 \\ -0.1071 & -0.9942 & -0.0032 \\ 0.0002 & -0.0032 & 1.0000 \end{bmatrix}, \end{aligned}$$

$$T_L = L_o U(:, 1:2)\Sigma(1:2, 1:2)^{-\frac{1}{2}} = \begin{bmatrix} -0.7690 & 0.3072 \\ -0.7879 & -1.2578 \\ -0.8048 & -2.5033 \end{bmatrix},$$

$$T_L = L_c^T V(:, 1:2)\Sigma(1:2, 1:2)^{-\frac{1}{2}} = \begin{bmatrix} -1.0416 & 0.6459 \\ -0.2507 & -0.6233 \\ -0.0018 & -0.0070 \end{bmatrix},$$

then

$$J_r = \begin{bmatrix} -0.8092 & 0.4780 \\ -0.4780 & -2.197 \end{bmatrix}, B_r = \begin{bmatrix} -1.1670 \\ -0.3343 \end{bmatrix}, \text{ and } C_r = \begin{bmatrix} -1.1670 & 0.3343 \end{bmatrix},$$

therefore the reduced order model is $\mathbf{\Sigma_r} = (J_r, B_r, C_r, D_a)$.

5.3 Algorithm for model reduction of piezo-mechanical systems

The algorithm outlined above is applicable for a standard states pace system to find a reduced order model. However, the idea is extended for a generalized case in [51] (see Chapter 7 in [51] for details). Moreover, in [20] the square root method is used to find reduced order models for Power system models, by preserving the sparsity pattern of the models in the descriptor form. In this section we present two algorithms (separately) based on square root methods for reduced order models of the piezo-mechanical systems in (5.2). Algorithm 5.1 gives a reduced order model in descriptor form, whereas Algorithm 5.2 computes a reduced order model in standard state space form.

5.3.1 How to use Algorithm 5.1 and 5.2 for MOR of piezo-mechanical systems

1 Convert the second order index-1 descriptor system in (3.11) into a first order index-1 descriptor form in 3.15, by preserving the sparsity pattern of the original model (see Chapter 3).

2 To compute the set R_+, Ritz-values with large magnitude, the Arnoldi process is performed with respect to $E_1^{-1}J$. To compute Ritz-values with small magnitude, first the Arnoldi process is performed with respect to $A^{-1}E$ (since inverting J is difficult in this case), and compute a set of Ritz-values R_-. Then inverting R_- we compute the Ritz-values with small magnitude of the matrix $E_1^{-1}J$.

3 The low rank approximation Z_c of the controllability Gramian in Step 2 of the Algorithm 5.1 and 5.2 becomes complex due to complex shifts. But practically, it should be real since Z_c is the Cholesky factor of $W_c \in \mathbb{R}^{n\times n}$. To circumvent this problem one can compute real shifts in Step 1. Otherwise, if the imaginary part of each column of Z_c is negligible, taking real (Z_c) does not effect in finding a reduced order model successfully. Same procedure is applicable for Z_o.

4 In the Step 3, we compute $Z_o = E_1^* Z_o$, because, we know that the controllability Gramian W_c and the observability Gramian W_o of the system 5.1 are the solutions of the Lyapunov equations

$$JW_cE_1^T + E_1W_cJ^T = -BB^T, \quad J^TW_cE_1 + E_1^TW_cJ = -C^TC.$$

Algorithm 5.1 SRM for MOR of index-1 descriptor systems (5.2) (descriptor to descriptor form).

Input: E_1, J_a, J_b, J_c, J_d, B_1, B_2, C_1 and C_2.
Output: $\mathbf{\Sigma}_r = (E_r, A_r, B_r, C_r)$.

1. Compute ADI shift parameters $\{\mu_1, \mu_2, \cdots, \mu_{i_{max}}\}$ using Algorithm 4.4.
2. Compute low rank Cholesky factors Z_c and Z_o, by solving Lyapunov equations in 4.4 via GSLRCF-ADI method (Algorithm 4.4).
3. $Z_o = E_1^* Z_o$.
4. Compute the (thin) singular value decomposition (SVD)

$$U\Sigma V^* = \text{svd}(Z_o^* Z_c).$$

5. Truncate after the kth largest singular values and compute

$$U_1 = U(:, 1:k),\ \Sigma_1 = \Sigma(1:k, 1:k) \text{ and } V_1 = V(:, 1:k).$$

6. Define the matrix transformations T_L and T_R as follows:

$$T_L = Z_o U_1 \Sigma_1^{-\frac{1}{2}} \text{ and } T_R = Z_c V_1 \Sigma_1^{-\frac{1}{2}}.$$

7. Compute

$$\begin{aligned}
\hat{E}_1 &= T_L^T E_1 T_R, \\
\hat{J}_a &= T_L^T J_a T_R, \\
\hat{J}_b &= T_L^T J_b, \\
\hat{J}_c &= J_c T_R, \\
\hat{B}_1 &= T_L^T B_1, \\
\hat{C}_1 &= C_1 T_R.
\end{aligned}$$

8. Form

$$E_r = \begin{bmatrix} \hat{E}_1 & 0 \\ 0 & 0 \end{bmatrix},\ A_r = \begin{bmatrix} \hat{J}_a & \hat{J}_b \\ \hat{J}_c & J_d \end{bmatrix},\ B_r = \begin{bmatrix} \hat{B}_1 \\ B_2 \end{bmatrix},\ C_r = \begin{bmatrix} \hat{C}_1 & C_2 \end{bmatrix}.$$

Algorithm 5.2 SRM for MOR of index-1 descriptor systems (5.2) (descriptor to standard form).

Input: E_1, J_a, J_b, J_c, J_d, B_1, B_2, C_1 and C_2.
Output: $\mathbf{\Sigma}_r = (J_r, B_r, C_r, D_r)$.

1. Compute ADI shift parameters $\{\mu_1, \mu_2, \cdots, \mu_{i_{max}}\}$ using Algorithm 4.4.
2. Compute low rank Cholesky factors Z_c and Z_o, by solving Lyapunov equations in 4.4 via GSLRCF-ADI method (Algorithm 4.4).
3. $Z_o = E_1^* Z_o$.
4. Compute the (thin) singular value decomposition (SVD)

$$U\Sigma V^* = \text{svd}(Z_o^* Z_c).$$

5. Truncate after the kth largest singular values and compute

$$U_1 = U(:, 1:k),\ \Sigma_1 = \Sigma(1:k, 1:k) \text{ and } V_1 = V(:, 1:k).$$

6. Define the matrix transformations T_L and T_R as follows:

$$T_L = Z_o U_1 \Sigma_1^{-\frac{1}{2}} \text{ and } T_R = Z_c V_1 \Sigma_1^{-\frac{1}{2}}.$$

7. Compute

$$\begin{aligned}
\hat{E}_1 &= T_L^T E_1 T_R, \\
\hat{J}_a &= T_L^T J_a T_R, \\
\hat{J}_b &= T_L^T J_b, \\
\hat{J}_c &= J_c T_R, \\
\hat{B}_1 &= T_L^T B_1, \\
\hat{C}_1 &= C_1 T_R.
\end{aligned}$$

8. Form

$$\begin{aligned}
J_r &= \hat{E}_1^{-1}(\hat{J}_a - \hat{J}_b J_d^{-1} \hat{J}_c), \\
B_r &= \hat{E}_1^{-1}(\hat{B}_1 - \hat{J}_b J_d^{-1} B_2), \\
C_r &= (\hat{C}_1 - C_2 J_d^{-1} \hat{J}_c), \\
D_r &= -C_2 J_d^{-1} B_2.
\end{aligned}$$

Now, write 5.1 in standard state space form

$$\dot{x}(t) = \overline{J}x(t) + \overline{B}u(t), \quad y(t) = Cx(t), \tag{5.11}$$

where $\overline{J} = E_1^{-1}J$, $\overline{B} = E_1^{-1}B$. The respective Lyapunov equations for 5.11 are

$$\overline{J}\tilde{W}_c + \tilde{W}_c\overline{J}^T = -\overline{B}\overline{B}^T, \quad \overline{J}^T\tilde{W}_c + \tilde{W}_c\overline{J} = -C^TC.$$

It is observed that $\tilde{W}_c = W_c$, but $\tilde{W}_o = E_1^* W_o E_1$. Therefore $\tilde{Z}_c = Z_c$, but $\tilde{Z}_o = E_1^* Z_o$, where $\tilde{Z}_c$ and $\tilde{Z}_o$ are the Cholesky factors of $\tilde{W}_c$ and $\tilde{W}_o$, respectively.

5 We choose $k << n$ such that the k largest Hankel singular values are preserved in the reduced order model. And it is adapted by using an appropriate error tolerance according to 5.7.

6 We recommend to see [37] for details about the formulation of T_L and T_R in the Step 6.

5.4 Numerical examples

Three sets of results are discussed in this section to assess the performance of the proposed algorithm. One set for the SISO case, another one for the MIMO case of the artificial models presented in Section 4.3 and finally a set of results obtained for the piezo-mechanical original model from [34] are included. All the results are obtained by using MATLAB 7.11.0 (R2010b). But for the first two sets we use an Intel Pentium Dual-Core processor with a 1.8-GHz clock and 2 GB of RAM, whereas the final set of results are carried out with an Intel Xeon Dual-Core CPU with a 3.0-GHz clock and 64 GB of RAM. Note that the dimension of the reduced order model in all cases is controlled by the tolerance $\sqrt{\epsilon}$ where ϵ is the machine precision.

For the different systems, the number of parameters used in the GSLRCF-ADI methods, the iteration number for computing Z_c and Z_o and the dimensions of the reduced-order models are given in the Table 5.1.

5.4.1 SISO piezo-mechanical (artificial) model

Here a SISO piezo-mechanical (artificial) model of dimension 10500 (n=5000, l=500) is considered, where $B_1 = [zeros(n,1); ones(n,1)]$, $B_2 = ones(l,1)$, $C_1 = [ones(1,n) \quad zeros(1,n)]$ and $C_2 = ones(1,l)$ (see Section 4.3). 10 shifts

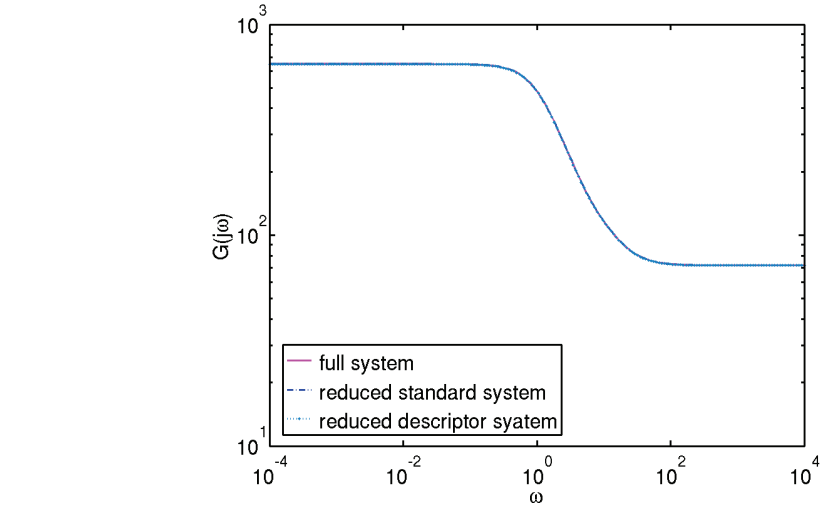
systems	ADI shift no.	K_+	K_-	iteration numbers (Z_c, Z_o)	ROM dimension
SISO (artificial)	10	20	15	15, 16	509 (des), 9 (stand)
MIMO (artificial)	12	20	15	13, 14	83
MIMO (original)	40	60	50	23, 23	38

Table 5.1: Parameters used in the GSLRCF-ADI methods, iteration numbers for low rank approximation of Gramians and dimensions of reduced-order models for different systems

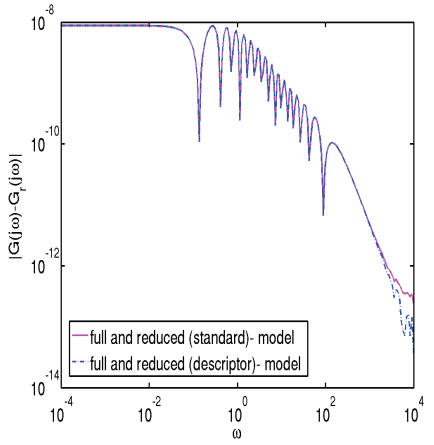
are chosen from $K_+ = 20$ and $K_- = 15$ to compute Z_c and Z_o. The convergence histories of the GSLRCF-ADI method are shown in Figure 4.2 (Section 4.3). Algorithm 5.1 and Algorithm 5.2 yield 509th (descriptor) and 9th (standard)-order reduced models, respectively. Figure 5.1 shows the accuracy of both algorithms. The frequency responses for the full system and its 509th and 9th-order reduced systems with very good matching are shown in Figure 5.1a on the frequency domain 10^{-4} Hz to 10^4 Hz. On the same domain the absolute errors and the relative errors of the frequency responses for full and reduced-order (descriptor and standard) systems are depicted in Figure 5.1b and Figure 5.1c, respectively.

5.4.2 MIMO piezo-mechanical (artificial) model

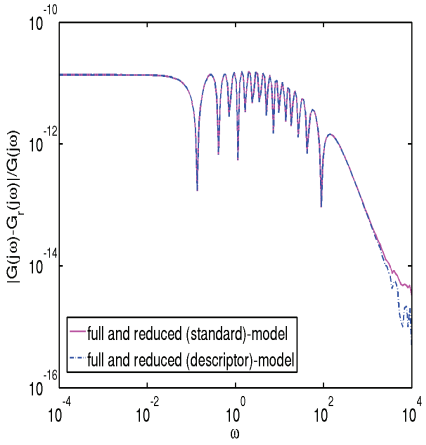
In this case a model of dimension 21000 (n=10000, l=1000) with 25 inputs and 20 outputs is considered (see Section 4.3). While the GSLRCF-ADI method is implemented with 12 ADI shifts (from $K_+ = 20$ and $K_- = 15$), it takes 13 and 14 iteration steps to compute Z_c and Z_o, respectively. The Algorithm 5.2 generate a 83rd-order reduced model. Figure 5.3 compares the full and 83rd-order reduced model on the frequency domain. Figure 5.2a shows the maximum and minimum singular value plots for both full and reduced-order model which indicate a nice match between them. The respective absolute errors and relative errors are shown in Figure 5.2b and Figure 5.2c, respectively. A diagram showing the computational time for different parts of the algorithm is given in Figure 5.3, which shows that 84% of the overall computational time is consumed to compute the low rank approximation of the controllability and observability Gramians.



(a) Frequency responses

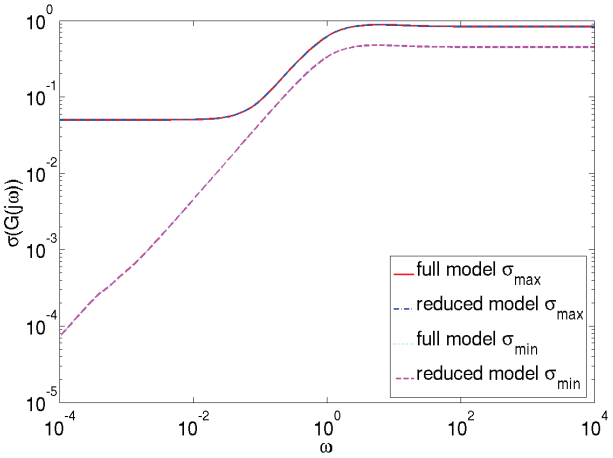


(b) Absolute errors

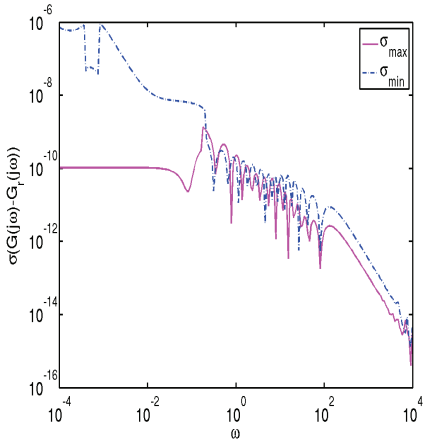


(c) Relative errors

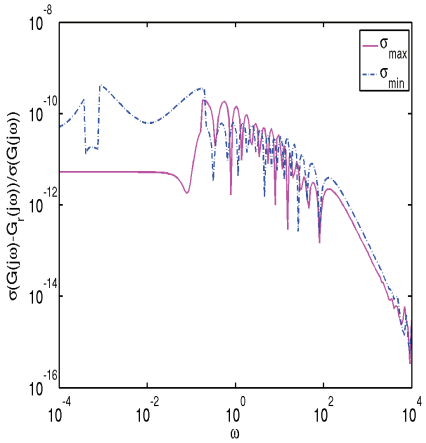
Figure 5.1: Frequency responses and errors in frequency responses for SISO (artificial) piezo-mechanical (full and reduced) model by Algorithm 5.1 and Algorithm 5.2



(a) Sigma plots



(b) Absolute errors



(c) Relative errors

Figure 5.2: Sigma plot (maximum and minimum) and errors for MIMO piezo-mechanical (artificial) full and reduced-order model computed by Algorithm 5.2

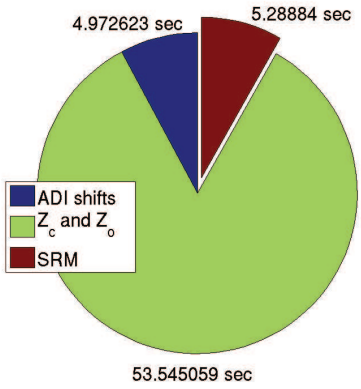


Figure 5.3: Computational time of different parts in computing reduced order model by Algorithm 5.2 for the MIMO piezo-mechanical (artificial) model

5.4.3 MIMO piezo-mechanical (original) model

The piezo-mechanical (original) model considered in this thesis is a second order system (3.11). A data of this model was provided by B. Kranz of the Fraunhofer Institute for Machine Tools and Forming Technology (IWU) in Dresden, Germany ([34]). In this data the dimension of the system is 290137 (second order model). In order to apply Algorithm 5.2 for MOR of this model, first we convert it into a first order model as in (3.15) (see Section 3.3 for details). The formulation represented in (3.13-3.15) is not applicable for the given data due to the numerical error. As a result, we represent the system as the form in (3.5). Then reordering the system matrices, finally the desired first order index-1 descriptor representation given in (3.15) is formed. And in this formulation the dimension of the system become 580274.

To compute Z_c and Z_o the GSLRCF-ADI method are performed with normalized residual norm tolerance $\sqrt{\epsilon}$. The number of ADI (real) shifts used in GSLRCF-ADI and the iteration numbers for computing Z_c and Z_o are shown in the Table 5.1. The convergence histories of GSLRCF-ADI methods are shown in Figure 5.4. Figure 5.4a and Figure 5.4b show the convergence histories for Z_c and Z_o, respectively. Algorithm 5.2 yields 38th-order standard state space reduced model when the truncation tolerance for the MOR is set to $\sqrt{\epsilon}$. The dimension of the reduced order model can be reduced further by increasing the tolerance. This is shown in the Table 5.2. Figure 5.5 shows the sigma plots of full and 38th-order reduced model on a wide (10^0

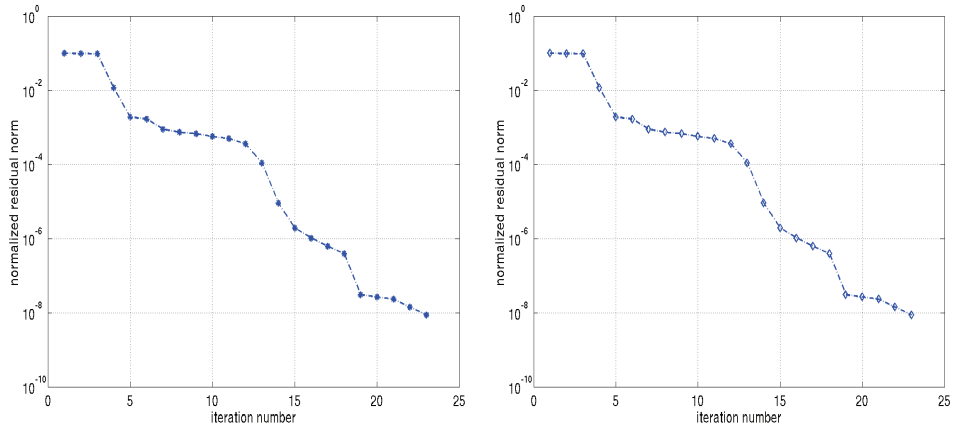
MOR tolerence	ROM dimension
$\sqrt{\epsilon}$ (ϵ is machine precision)	38
10^{-5}	27
10^{-4}	23
10^{-3}	22
10^{-2}	18
10^{-1}	16

Table 5.2: MOR tolerance and the dimensions of the reduced-order models for MIMO piezo-mechanical (original) model using Algorithm 5.2

Hz to 10^8 Hz) frequency range. Figure 5.6 shows the errors between full and reduced model (Figure 5.6a shows absolute errors, Figure 5.6b shows relative errors) of sigma plots where the errors are very negligible.

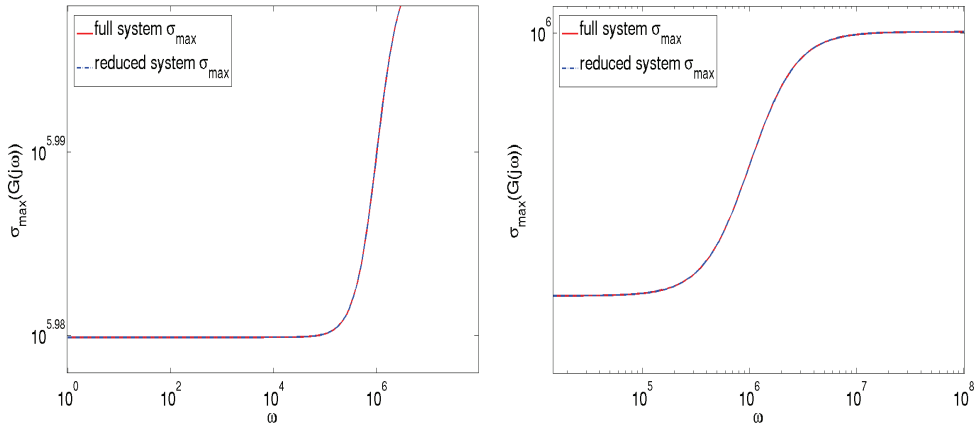
On the other hand Figure 5.7 shows the frequency responses and the deviation of the frequency responses for the individual component of the transfer function. For example Figure 5.7a shows the frequency responses of full and 38th-order reduced model for $G_{1,1}(j\omega)$, and Figure 5.7b shows the deviation between full and reduced model. Similarly, Figure 5.7c and Figure 5.7d show the frequency responses and deviation, respectively for $G_{1,9}(j\omega)$.

From Figure 5.8 it can easily be identified that the reduced order model is stable as it shows that all the poles of reduced order model lie in the open left complex half plane.



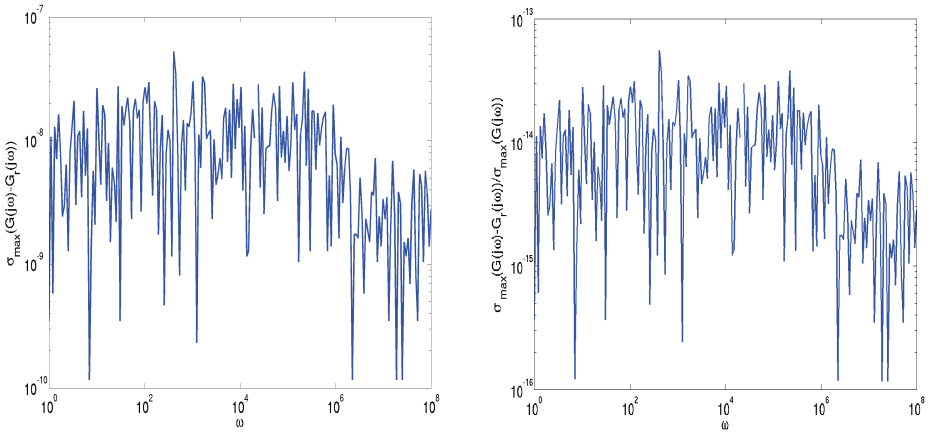
(a) Controllability Gramian (b) Observability Gramian

Figure 5.4: Convergence histories of GSLRCF-ADI methods for MIMO piezo-mechanical (original) model



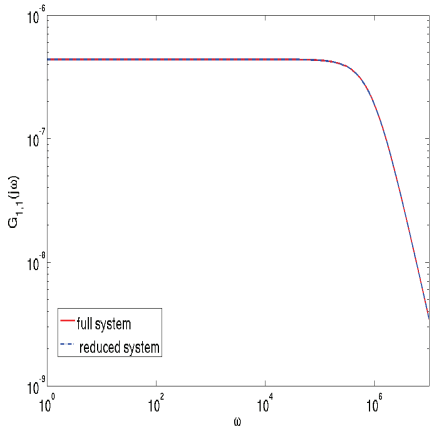
(a) (b)

Figure 5.5: Sigma plot (maximum singular value) of full and 38th-order reduced systems for MIMO piezo-mechanical (original) model

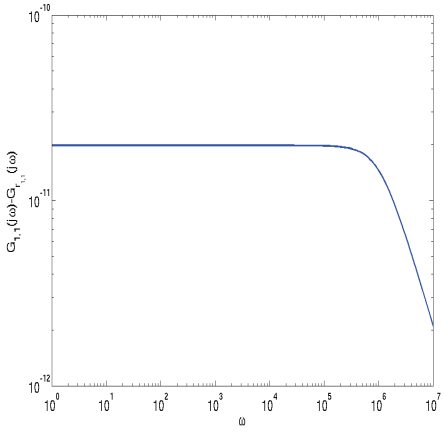


(a) Absolute error (b) Relative error

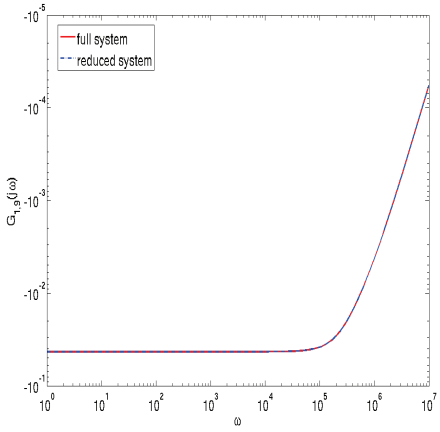
Figure 5.6: Errors in the sigma plot of full and 38th-order reduced systems for MIMO piezo-mechanical (original) model



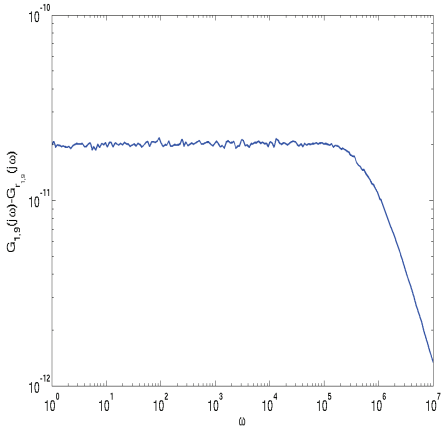
(a) Frequency responses for $G_{1,1}(j\omega)$



(b) Deviation of frequency responses for $G_{1,1}(j\omega)$



(c) Frequency responses for $G_{1,9}(j\omega)$



(d) Deviation of frequency responses for $G_{1,9}(j\omega)$

Figure 5.7: Frequency responses for full and 38th-order reduced system for individual component of $G(j\omega)$ and respective deviations for piezo-mechanical (original) model

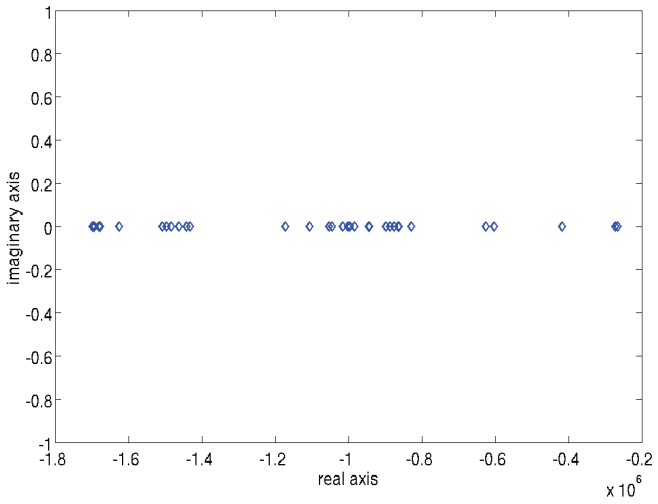


Figure 5.8: Poles of 38th-order reduced system for MIMO piezo-mechanical (original) model

CHAPTER

SIX

CONCLUSION

In this thesis, we have developed a balanced truncation (BT) algorithm for model order reduction (MOR) of piezo-mechanical systems. The algorithm works for a class of large-scale index-1 DAE systems. The efficiency of the algorithm is verified by real-world data describing micro-mechanical piezo-actuators. The algorithm is tremendously efficient for the highly sparse systems. We have discussed the algorithm along with numerical results in the Chapter 5.

The most expensive part of our proposed algorithm is to solve two generalized continuous-time algebraic Lyapunov equations (GCALE), namely controllability GCALE and observability GCALE. GSLRCF-ADI methods described in the Chapter 4 ensure first computation to compute low rank approximation of the Gramians. Since the algorithm is designed for index-1 descriptor systems which are highly sparse, the computation cost is cheap. We have shown the performance of this method by numerical examples in the same chapter.

To compute the ADI parameters, which play an important role in the convergency of ADI methods, we have imitated the technique introduced in [47]. But this technique is reformulated for the index-1 descriptor systems, which is also presented in the Chapter 4. For MOR of piezo-mechanical system we consider real shifts parameters, and using these parameters we obtain a very good reduced order model. To overcome the complexity of computing the real low rank approximation of controllability and observability Gramians, we recommend to compute real shift parameters. To stop the GSLRCF-ADI

methods we use residual based criteria described in the Chapter 4. This technique is less expensive than any other techniques since here we compute the approximate 2-norm of the residuals of the GCALE. Nevertheless ADI method is expensive due to the stopping criteria. And this can be an interesting research topic in future.

Moreover, another important task in this thesis is to preserve the sparsity while converting a second order system into a first order system to be adapted for the algorithm. In Chapter 3, we have presented several techniques to reduce a second order system into a first order equivalent system, whereas the properties of the original system like sparsity, symmetry and stability are preserved.

BIBLIOGRAPHY

[1] A. Antoulas, *Approximation of linear dynamical systems*, Encyclopedia of Electrical and Electronics Engineering, Citeseer, 11 (1999), pp. 403–422. 16, 49, 51

[2] ——, *Approximation of large-scale dynamical systems*, SIAM, Philadelphia, PA, 2005. 1, 12, 14, 15, 51, 53

[3] A. Antoulas, D. Sorensen, and S. Gugercin, *A survey of model reduction methods for large-scale systems*, Contemp. Math., 280 (2001), pp. 193–219. 1

[4] S. Bächle, *Index reduction for differential-algebraic equations in circuit simulation*, TU Berlin, Germany, Tech. Rep. MATHEON, 141 (2004), pp. 1–33. Available from `http://www.matheon.de/preprints/172.pdf`. 10

[5] Z. Bai, *Templates for the solution of algebraic eigenvalue problems: a practical guide*, SIAM, Philadelphia, PA, 2000. 17

[6] R. Bartels and G. Stewart, *Solution of the matrix equation ax+ xb= c [F4]*, Communications of the ACM, 15 (1972), pp. 820–826. 31

[7] P. Benner, J. Li, and T. Penzl, *Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems*, Numerical Linear Algebra with Applications, 15 (2008), pp. 755–777. 31, 34, 35, 38

[8] P. Benner, V. Mehrmann, and D. Sorensen, eds., *Dimension reduction of large-scale systems*, vol. 45 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin/ Heidelberg, Germany, 2005. 1, 9, 38, 49, 50, 53

[9] P. Benner, H. Mena, and J. Saak, *On the parameters selection problem in the Newton-ADI iteration for large scale Riccati equations*, Electronic

Transactions on Numerical Analysis, Kent State University, 29 (2008), pp. 136–149. 38

[10] S. Costa, F. D. Freitas, and A. S. e. Silva, *Design of decentralized controllers for large power systems considering sparsity*, IEEE Transactions on Power Systems, 12 (1997), pp. 144–152. 23, 25

[11] B. Datta, *Numerical methods for linear control systems*, Elsevier Academic Press, London, UK, 2004. 11, 12, 13, 15, 49, 52

[12] M. Davis, A. Gravagne, J. Jackson, and J. Marks, *Controllability, observability, realizability, and stability of dynamic linear system*, Electric Journal of Differential Equation, 37 (2009), pp. 1–32. URL: http://ejde.math.txstate.edu. 14, 15

[13] U. Desai and D. Pal, *A transformation approach to stochastic model reduction*, IEEE Transactions on Automatic Control, 29 (2002), pp. 1097–1100. 51

[14] J. Distefano, A. Stubberud, and I. Williams, *Feedback and control systems: continuous (analog) and discrete (digital)*, Schaums outline series, McGraw-Hill, New York, 1990. 11

[15] D.Kressner, *Numerical methods for general and structured eigenvalue problems*, Springer-Verlag, New York, 2005. 17, 18, 19

[16] W.-G. Drossel and V. Wittstock, *Adaptive spindle support for improving machining operations*, CIRP Annals - Manufacturing Technology, 57 (2008), pp. 395–398. 2, 3

[17] G. E. Dullerud and F. G. Paganini, *A course in robust control theory: A convex approach*, Springer-Verlag, New York, 2000. 12, 13, 14, 15, 49

[18] D. F. Enns, *Model reduction with balanced realizations: An error bound and a frequency weighted generalization*, in The 23rd IEEE Conference, 1984. Decision and Control, vol. 23, IEEE, 2007, pp. 127–132. 51

[19] F. Freitas and A. Costa, *Computationally efficient optimal control methods applied to power systems*, in 20th International Conference on Power Industry Computer Applications, 1997, IEEE, 2002, pp. 287–294. 9, 23, 25

[20] F. Freitas, J. Rommes, and N. Martins, *Gramian-based reduction method applied to large sparse power system descriptor models*, IEEE Transactions on Power Systems, 23 (2008), pp. 1258–1270. v, 23, 25, 28, 31, 33, 35, 36, 54

[21] R. Freund, *Krylov-subspace methods for reduced-order modeling in circuit simulation*, Journal of Computational and Applied Mathematics, 123 (2000), pp. 395–421. 2

[22] P. Fuhrmann, *A polynomial approach to linear algebra*, Springer-Verlag, New York, 1996. 8, 15

[23] M. Gerdin, *Identification and estimation for models described by differential-algebraic equations*, Department of Electrical Engineering, Linköpings Universitet, Linköpings, Sweden, 2006. 9

[24] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds*, IntControl, 39 (1984), pp. 1115–1193. 2, 52

[25] G. Golub and C. Van Loan, *Matrix computations*, Johns Hopkins University Press, 1996. 17, 20

[26] M. Green, *Balanced stochastic realizations*, Linear Algebra and Its Applications, 98 (1988), pp. 211–247. 51

[27] ——, *A relative error bound for balanced stochastic truncation*, IEEE Transactions on Automatic Control, 33 (2002), pp. 961–965. 51

[28] M. Green and D. Limebeer, *Linear robust control*, PrenticeHall, Englewood Cliffs, NJ, 1995. 49

[29] E. J. Grimme, *Krylov projection methods for model reduction*, PhD thesis, University of Illinois at Urbana, Champaign, 1997. 2

[30] S. Gugercin and A. Antoulas, *A survey of model reduction by balanced truncation and some new results*, International Journal of Control, 77 (2004), pp. 748–766. 51

[31] S. Gugercin and J. R. Li, *Smith-type methods for balanced truncation of large sparse systems*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. Sorensen, eds., Lecture Notes in Computational Science and Engineering, Berlin/Heidelberg, Germany, 2005, Springer-Verlag, pp. 49–82. 32, 33, 34, 35, 50, 51, 52

[32] S. Hammarling, *Numerical solution of the stable, non-negative definite Lyapunov equation*, IMA Journal of Numerical Analysis, 2 (1982), p. 303. 31

[33] G. Huang, L. Zhao, and X. Song, *A new bifurcation analysis for power system dynamic voltage stability studies*, Power Engineering Society Winter Meeting, 2002. IEEE, 2 (2002), pp. 882–887. 9, 23, 25

[34] B. Kranz, *Zustandsraumbeschreibung von piezo-mechanischen Systemen auf Grundlage einer Finite-Elemente-Diskretisierung*, in Analysis conference & 27th CADFEM users' meeting 2009, Congress Center Leipzig, Germany, November (18-20) 2009. v, 2, 9, 23, 26, 28, 29, 57, 61

[35] A. Kumar and P. Daoutidis, *Feedback control of nonlinear differential-algebraic-equation systems*, AIChE Journal, 41 (1995), pp. 619–636. 9, 10

[36] P. Kunkel and V. Mehrmann, *Differential-algebraic equations: Analysis and numerical solution*, European Mathematical Society, 2006. 10

[37] A. Laub, M. Heath, C. Paige, and R. Ward, *Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms*, IEEE transactions on automatic control, 32 (2002), pp. 115–122. 57

[38] J. Li, *Model reduction of large linear systems via low rank system gramians*, PhD thesis, Massachusetts Institute of Technology, 2000. 1

[39] J. R. Li and J. White, *Low-rank solution of Lyapunov equations*, SIAM Review, 46 (2004), pp. 693–713. 31, 34, 35, 38

[40] Y. Liu and B. D. O. Anderson, *Singular perturbation approximation of balanced systems*, Proc. 28th IEEE Conf. Decis. Control, 2 (1989), pp. 1355–1360. 2

[41] N. Martins, *Efficient eigenvalue and frequency response methods applied to power system small-signal stability studies*, IEEE Transactions on Power Systems, 1 (2007), pp. 217–224. 9, 23, 25

[42] V. Mehrmann and T. Stykel, *Balanced truncation model reduction for large-scale systems in descriptor form*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. Sorensen, eds., Lecture Notes in Computational Science and Engineering, Berlin/Heidelberg, Germany, 2005, Springer-Verlag, pp. 83–115. 11, 15, 16

[43] B. Moore, *Principal component analysis in linear systems: Controllability, observability, and model reduction*, IEEE Transactions on Automatic Control, 26 (2002), pp. 17–32. 2

[44] R. Neugebauer, W.-G. Drossel, A. Bucht, B. Kranz, and K. Pagel, *Control design and experimental validation of an adaptive spindle support for enhanced cutting processes*, CIRP Annals - Manufacturing Technology, 59 (2010), pp. 373–376. 2

[45] R. Neugebauer, W.-G. Drossel, K. Pagel, and B. Kranz, *Making of state space models of piezo-mechanical systems with exact impedance mapping*

and strain output signals, in Mechatronics and Material Technologies, vol. 2, Zurich, Switzerlan, June 28-30 2010, Swiss Federal Institute of Technology ETH, pp. 73–80. 26

[46] D. Peaceman and H. Rachford Jr, *The numerical solution of parabolic and elliptic differential equations*, SIAM, 3 (1955), pp. 28–41. 33

[47] T. Penzl, *A cyclic low rank Smith method for large, sparse Lyapunov equations with applications in model reduction and optimal control*, Tech. Rep. SFB393/98-6, 09107 Chemnitz, FRG, 1998. Available from `http://www.tu-chemnitz.de/sfb393/sfb98pr.html`. 31, 33, 34, 38, 39, 67

[48] Z. Qu, *Model order reduction techniques: With applications in finite element analysis*, Springer-Verlag, London, 2004. 1

[49] Y. Saad, *Numerical methods for large eigenvalue problems*, Manchester University Press, Manchester, UK, 1992. 17, 18

[50] ——, *Iterative methods for sparse linear systems*, SIAM, Philadelphia, PA, 2003. 22

[51] J. Saak, *Efficient numerical solution of large scale algebraic matrix equations in PDE control and model order reduction*, PhD thesis, Chemnitz University of Technology, Chemnitz, Germany, 2009. 25, 36, 38, 39, 40, 54

[52] B. Salimbahrami, *Structure preserving order reduction of large scale second order models*, PhD thesis, Technische Universität München, 2006. 1, 23, 24

[53] B. Salimbahrami and B. Lohmann, *Structure preserving order reduction of large scale second order systems*, in Proceedings of 10th IFAC/IFORS/IMACS/IFIP Symposium on Large Scale Systems: Theory and Applications, Osaka, Japan, July, 2004, pp. 245–250. 23

[54] W. H. A. Schilders, H. A. Vorst, and J. Rommes, *Model order reduction: theory, research aspects & applications (Mathematics in Industry, European Consortium for Mathematics in Industry)*, vol. 13, Springer-Verlag, Berlin, 2008. 1

[55] G. Simmons and S. Krantz, *Differential equations: Theory, technique, and practice*, McGraw-Hill, 2007. 15

[56] E. Sontag, *Mathematical control theory: Deterministic finite dimensional systems*, Springer-Verlag, New York, 1998. 12, 13

[57] D. Sorensen and A. Antoulas, *On model reduction of structured systems*, in Dimension reduction of large-scale systems, P. Benner, V. Mehrmann, and D. Sorensen, eds., Lecture Notes in Computational Science and Engineering, Berlin/Heidelberg, 2005, Springer-Verlag, pp. 117–130. 13

[58] D. Sorensen and Y. Zhou, *Bounds on eigenvalue decay rates and sensitivity of solutions to Lyapunov equations*, Tech. Rep. TR02-07, Dept. of Comp. Appl. Math., Rice University, Houston, TX, jun 2002. Available online from `http://www.caam.rice.edu/caam/trs/tr02.html#TR02-07`. 38

[59] G. Starke, *Optimal alternating direction implicit parameters for nonsymmetric systems of linear equations*, SIAM Journal on Numerical Analysis, 28 (1991), pp. 1431–1445. 38

[60] T. Stykel, *Analysis and numerical solution of generalized Lyapunov equations*, Ph.D. thesis, Institut für Mathematik, Technische Universität Berlin, Berlin, Germany, 2002. Available from `http://www.math.tu-berlin.de/~stykel/Publications/disser.ps`. 9, 15

[61] ——, *Balanced truncation model reduction of second-order systems*, in Proceedings of the 5th Vienna Symposium on Mathematical Modelling, Vienna, February 2006, pp. 1–9. 23, 25

[62] ——, *Low-rank iterative methods for projected generalized Lyapunov equations*, Electronic Transactions on Numerical Analysis, 30 (2008), pp. 187–202. 31, 38

[63] S. Tan and L. He, *Advanced model order reduction techniques in VLSI design*, Cambridge Univ Press, Cambridge, UK, 2007. 1

[64] M. Tombs and I. Postlethwaite, *Truncated balanced realization of a stable non-minimal state-space system*, International Journal of Control, 46 (1987), pp. 1319–1330. 2, 52

[65] L. Trefethen and D. Bau, *Numerical linear algebra*, SIAM, Philadelphia, PA, USA, 1997. 18, 19, 20

[66] E. Wachspress, *The ADI minimax problem for complex spectra*, Applied Mathematics Letters, 1 (1988), pp. 311–314. 38

[67] ——, *Iterative solution of the Lyapunov matrix equation*, Applied Mathematics Letters, 1 (1988), pp. 87–90. 31, 33, 38

[68] W. Wang and M. Safonov, *A tighter relative-error bound for balanced stochastic truncation*, Systems & Control Letters, 14 (1990), pp. 307–317. 51

[69] D. Watkins, *The matrix eigenvalue problem: GR and Krylov subspace methods*, SIAM, Philadelphia, PA, USA, 2007. 17

[70] K. Zhou, *Frequency-weighted L_∞ norm and optimal Hankel norm model reduction*, IEEE Transactions on Automatic Control, 40 (2002), pp. 1687–1699. 51

[71] K. Zhou and J. Doyle, *Essentials of robust control*, vol. 104, Prentice-Hall Inc., Upper Saddle River, New Jersey, 1998. 12, 14, 21, 49

[72] Y. Zhou, *Numerical methods for large scale matrix equations with applications in LTI system model reduction*, PhD thesis, Rice University, 2002. 1