

Efficient Projection-based Stability Guaranteed Model Order Reduction for Active Circuits

by
Xu Deng

A thesis proposal submitted to the Faculty of Graduate and Postdoctoral Affairs in
partial fulfillment of the requirements for the degree of

Master of Applied Science
in
Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Electronics
Faculty of Engineering
Carleton University
Ottawa, Ontario, Canada

January 2016

Copyright © 2016

Xu Deng

Acknowledgements

I would sincerely like to express my gratitude to my supervisor, Professor Michel Nakhla. Without his guidance, this thesis would have been impossible. I appreciate the critical guidance and the constant support and encouragements he gave me throughout the process.

I would also like to thank my co-supervisor, Dr. Behzad Nouri, for his helpful suggestions and guidance, which was important in many stages of the research for this thesis. I wish to thank him also for his motivation and encouragements.

I give special thanks to my parents for their unconditional love, encouragement, and support. I am eternally grateful to them.

ABSTRACT

Moments matching model reduction method is an important tool for accelerating simulation of large scale system originating in circuit area. But this kind of method cannot guarantee the stability of the reduced model in the case of active circuits. To address this problem, a novel moments matching model reduction method is proposed in this thesis. The contributions in thesis can be summarized as follows.

First, instead of using the same matrix for right and left projection matrix, a different matrix is used as the left projection matrix, and this matrix can be obtained from solving Lyapunov equation for arbitrary negative definite right-hand side.

Second, the problem of solving the Lyapunov equation for arbitrary negative definite right-hand side is converted to a problem of finding eigen pairs. It is shown also that not all eigen pairs are needed. Instead, a restarted Arnoldi method is used to find a subset of the eigen pairs, which is much cheaper than solving the Lyapunov equation directly.

Table of Contents

Acknowledgements	i
ABSTRACT	ii
List of Figures	vii
List of Tables	x
List of Acronyms	xi
List of Symbols	xiii
1 Introduction	1
1.1 Background and Motivation	1
1.2 Objectives and Contributions	4
1.3 Organization of the Thesis	5
2 Background on formulation and characteristics of dynamic circuits	6

2.1	Dynamic systems, concept and definitions	6
2.2	Formulation of circuit equations	7
2.2.1	Linear-circuit equations	8
2.3	Stability and passivity of linear system	10
2.3.1	Stability	10
2.3.2	Passivity	12
3	Model order reduction: methods and properties	14
3.1	Projection Framework for MOR	15
3.2	Non-Krylov-Based Methods	17
3.2.1	Truncated Balanced Realization	17
3.2.2	Hankel optimal model reduction	20
3.3	Krylov-Based Methods	23
3.3.1	Arnoldi method	24
3.3.2	PRIMA	26
3.4	Non-Krylov versus Krylov based MOR	29
3.5	Stability Preservation in Model Order Reduction	30
3.5.1	Using Optimization algorithms	30
3.5.2	Post-processing methods	32
3.6	Summary and Problem Definition	33

4	The proposed stability guaranteed model reduction method	34
4.1	ROM Construction	35
4.1.1	Construction of the Right Projection Matrix	37
4.1.2	Construction of the Left Projection Matrix	38
4.2	Properties of the proposed method	39
4.2.1	Moments Matching	39
4.2.2	Stability Guarantee	47
4.3	Computational Complexity	49
4.4	Computational Results	50
5	Efficient Solution for Generalized Lyapunov equation	54
5.1	Problem Definition	55
5.2	Background and Preliminaries	57
5.3	Existence and Uniqueness of the Solutions of GLE	58
5.4	Proposed Solution Algorithm	63
5.4.1	Formulation of the Proposed Solution Algorithm	66
5.4.2	Construction of the Left Projection Matrix	70
5.4.3	Modified Left Projection Matrix	72
5.5	Computation cost analysis	76
5.5.1	Computational Results	77

6	Summary and future work	87
6.1	Summary	87
6.2	Future Work	88
	References	89

List of Figures

1.1	Moore's Law.	2
1.2	Model Order Reduction.	3
3.1	Projection based model order reduction.	16
3.2	Absolute error $\ \zeta(s)\ _H$ should be minimal.	22
3.3	Arnoldi method, where entries of \mathbf{R} is zero except the last column. . .	25
4.1	A stable active design consisting of coupled transmission lines and amplifier blocks.	50
4.2	Cross sectional geometry of TL structure.	50
4.3	High-frequency equivalent-circuit model for MOS cascade amplifier (the biasing network is not presented).	51
4.4	Comparison of the responses at V_{out2} from the proposed method, the PRIMA-based model and the original circuit.	52

5.1	Poles distribution of the original sparse system with 2000 state variables.	78
5.2	Poles distribution of the reduced system from the random generated 2000 by 2000 sparse system.	79
5.3	First 15 Moments Comparison between the original system and ROM.	79
5.4	A stable active design consisting of four coupled transmission lines and amplifier blocks.	81
5.5	Cross sectional geometry of TL structure.	81
5.6	High-frequency equivalent-circuit model for MOS cascade amplifier (the biasing network is not presented).	82
5.7	Comparison of the frequency responses at V_{out3} (at Far-end of line-3) from the proposed method, the PRIMA-based model and the original circuit.	82
5.8	Comparison of the frequency responses at V_{out4} (at Far-end of line-4) from the proposed method, the PRIMA-based model and the original circuit.	83
5.9	Comparison of the transient responses at V_{out3} (at Far-end of line-3) from the proposed method, the PRIMA-based model and the original circuit.	84

5.10 Comparison of the transient responses at V_{out4} (at Far-end of line-4)	
from the proposed method, the PRIMA-based model and the original	
circuit.	85

List of Tables

4.1	Comparison of the original and reduced models.	51
4.2	Comparison of few dominant poles.	53
5.1	Computation time for calculating left and right projection matrix. . .	80
5.2	Comparison of the original and reduced models.	86
5.3	Comparison of the CPU-time for frequency simulation using original and reduced models.	86

List of Acronyms

Acronyms	Definition
CAD	Computer Aided Design
CLDS	Continuous Time Linear Descriptor Systems
EVD	Eigenvalue Decomposition
GLE	General Lyapunov equation
GLI	Generalized Lyapunov Inequality
HSPICE	A commercial variation of SPICE, H stands for the family name of the creator (Kim Hailey)
LHS	Left Hand Side
LTI	Linear Time Invariant (Dynamical System)
MNA	Modified Nodal Analysis
MOR	Model Order Reduction
MOS	Metaloxidesemiconductor
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
POD	Proper Orthogonal Decomposition
PRIMA	Passive Reduced-order Interconnect Macromodeling Algorithm
PSPICE	Personal Simulation Program with Integrated Circuit Emphasis

Acronyms	Definition
RHS	Right Hand Side
RLC	Resistor, Inductor and Capacitor
ROM	Reduced Order Model
SISO	Single Input and Single Output system
SPD	Symmetric Positive Definite
SPICE	Simulation Program with Integrated Circuit Emphasis
SVD	Singular Value Decomposition
TBR	Truncated Balanced Realization
TL	Transmission Line
VLSI	Very Large Scale Integrated circuit

List of Symbols

Symbols	Definition
\mathbb{R}	The set of all real numbers
\mathbb{C}	The set of all complex numbers
$\mathbb{R}^{n \times m}$	The set of all real matrices of size $n \times m$
$\mathbb{C}^{n \times m}$	The set of all complex matrices of size $n \times m$
\mathbb{R}^n	Shortcut for $\mathbb{R}^{n \times 1}$
\mathbb{C}^n	Shortcut for $\mathbb{C}^{n \times 1}$
\mathbf{G}	Conductance matrices in MNA circuit equation
\mathbf{C}	Susceptance matrices in MNA circuit equation
\mathbf{B}	Input matrices in MNA circuit equation
\mathbf{L}	Output matrices in MNA circuit equation
\mathbf{u}	Inputs at time domain in MNA circuit equation
\mathbf{U}	Inputs at frequency domain in MNA circuit equation
\mathbf{x}	State variables at time domain in MNA circuit equation
\mathbf{X}	State variables at frequency domain in MNA circuit equation
\mathbf{F}	Shortcut for $-\mathbf{G}$, used in Lyapunov equation
\mathbf{A}	Shortcut for $-\mathbf{G}^{-1}\mathbf{C}$
\mathbf{R}	Shortcut for $\mathbf{G}^{-1}\mathbf{B}$
\mathbf{M}_i	The i-th moment of the original system's state variables
\mathcal{M}_i	The i-th moment of the original system's output
$\mathbf{H}(s)$	Transfer function

\mathbf{U}	Left projection matrix
\mathbf{V}	Right projection matrix
\mathbf{P}	The solution of general Lyapunov equation
$\hat{\mathbf{G}}$	Conductance matrices in reduced model
$\hat{\mathbf{C}}$	Susceptance matrices in reduced model
$\hat{\mathbf{B}}$	Input matrices in reduced model
$\hat{\mathbf{L}}$	Output matrices in reduced model
$\hat{\mathbf{u}}$	Inputs at time domain in reduced model
$\hat{\mathbf{U}}$	Inputs at frequency domain in reduced model
$\hat{\mathbf{x}}$	State variables at time domain in reduced model
$\hat{\mathbf{X}}$	State variables at frequency domain in reduced model
$\hat{\mathbf{A}}$	Shortcut for $-\hat{\mathbf{G}}^{-1}\hat{\mathbf{C}}$
$\hat{\mathbf{R}}$	Shortcut for $\hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}$
$\hat{\mathbf{M}}_i$	The i-th moment of the reduced model's state variables
$\hat{\mathcal{M}}_i$	The i-th moment of the reduced model's output
$\mathcal{L}(\mathbf{x})$	Lyapunov function
\mathbf{A}^T	Complex-conjugate transpose of complex matrix \mathbf{A}
\mathbf{A}^T	Complex-conjugate transpose of complex matrix \mathbf{A}
$\mathbf{A}_{n \times n}^{-1}$	The inverse of the square matrix \mathbf{A}
$\mathbf{A} > 0$	\mathbf{A} is a positive definite matrix
$\mathbf{A} \geq 0$	\mathbf{A} is a semi-positive definite matrix
$\text{colsp}(\mathbf{A})$	Column span (also called <i>range</i>) of matrix \mathbf{A}

Chapter 1

Introduction

1.1 Background and Motivation

In the world of today's technology, electronic systems, such as computers and mobile devices, have become essential parts of our daily lives. With the remarkable growth of Very Large-Scale Integration (VLSI) technology these, electronic systems can operate up to gigahertz frequencies with the rise times in the order of nanosecond level. Almost fifty years ago, this trend was predicted by Moore's law stating that the number of transistors in an unit area of integrated circuit doubles every two years.

As the Moore's law continue to dominate the world of integrated circuits, it is important to take the challenge of simulating and optimizing the increasingly complicated integrated circuits by developing more efficient algorithms [1]. Therefore, the

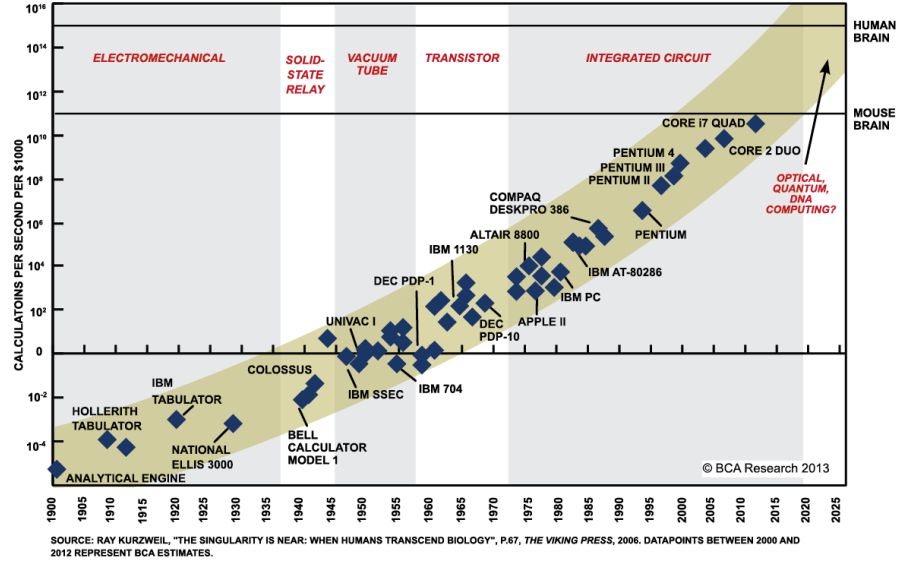


Figure 1.1: Moore's Law.

computer-aided design (CAD) techniques at both chip and packaging level has become an active field of research in last two decades.

Model order reduction is one of the CAD techniques devised to improve the efficiency of the simulation of very-large systems. It finds a Reduced Order Model (ROM) that has similar input-output behavior with the original model (Figure 1.2). Since the reduced model is much smaller than the original model, the analysis of the reduced model is much easier than the analysis of the original model [1]. Normally, a model arising from circuit simulation has an order of thousands to millions, but a ROM usually has an order of less than a hundred. Thus, MOR is attracting lots of attention in VLSI simulation area.

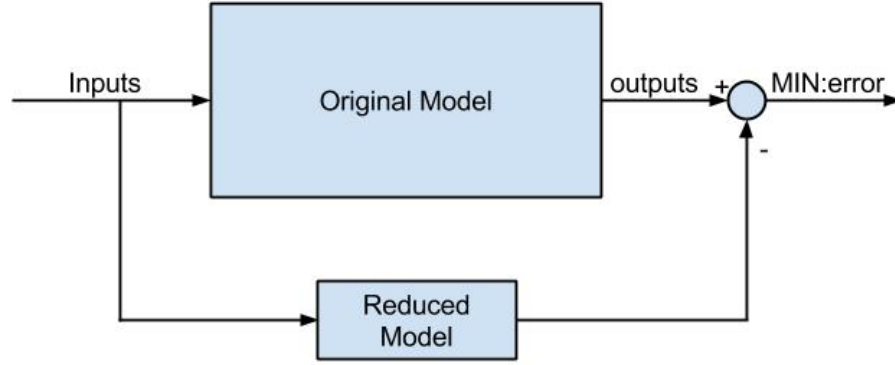


Figure 1.2: Model Order Reduction.

It is required that the resulting ROM from model order reduction must be accurate and at the same time efficient to be incorporated in a simulation. It also should be stable to give reliable system-level simulation results [2]. Unstable models can lead to inaccurate or totally unfeasible time-domain simulation.

Model order reduction (MOR) algorithms have been well established for compact modeling and the analysis of large passive linear circuits [3]. Several algorithms have been proposed in the literature to preserve the passivity and consequently the stability of the reduced model for originally passive circuits [4]. Examples of such passive circuits are lumped and distributed interconnects. However, there are many practical circuits that are stable but not passive. The models for these circuits, such as small-signal device model, usually contains dependent voltage and/or dependent current sources. Using conventional MOR techniques such as PRIMA [5] for MOR of these circuits would not guarantee the stability of the resulting model. However, stability

is a crucial property. If the stability property is not guaranteed in generation of a macromodels, the simulation results of such macromodels and the combined system using them would be unreliable.

The available stability preserving methods in the literature [6–9] are usually based on post-processing. For example, in [9–11] a method is proposed to eliminate the unstable poles of the reduced system by using implicitly and explicitly restarted Arnoldi and Lanczos algorithms. However, the common concept in these methods is to sacrifice accuracy of the reduced model in order to guarantee stability, which may destroy the integrity of the moment-matching algorithm leading to inaccurate reduced model [6]. In addition, numerically restoring stability algorithms are not guaranteed to converge and in general have relatively high computational cost associated with them. Classical Truncated Balanced Realization (TBR) is another method to preserve the stability. However, it is computationally expensive which makes it not suitable for large circuits. In addition, existence of a solution for general circuit formulation is not guaranteed. For systems of higher order, the projection-based methods such as the Krylov-subspace approach are a promising choice.

1.2 Objectives and Contributions

The goal of this thesis is to address the above difficulty through developing an efficient method whose computational complexity scales linearly. This thesis proposes

a novel stability-preserving projection algorithm for model reduction of active linear circuits. To this end, a novel projection framework is presented for constructing stable reduced macromodels for stable active linear circuits. The core idea is to construct the right and the left projection matrices separately. The right projection matrix is constructed such that it guarantees accuracy through matching the moments and the left projection matrix is derived such that Lyapunov stability constraint is satisfied.

1.3 Organization of the Thesis

The rest of this thesis is organized as follows. Chapter 2 presents a concise background of the main subjects relevant to this work such as, dynamical circuits and important physical characteristics such as stability and passivity. Chapter 3 reviews the general concept of model order reductions, where some of the important MOR methods and their application to the active circuits are also reviewed. Chapter 4 describes the details of development of the proposed algorithm and demonstrate the validity of the proposed method. Next, in Chapter 5, an efficient solution for generalized Lyapunov equations and its application in improving the proposed method is presented and validated by numerical example. Finally, Chapter 6 provides the concluding remarks and future research options.

Chapter 2

Background on formulation and characteristics of dynamic circuits

2.1 Dynamic systems, concept and definitions

A dynamic system is an operator \mathcal{S} that maps an input function vector $\mathbf{u}(t)$ to an output function $\mathbf{y}(t)$. It also defines how the states of system are evolving in time. Typically in circuit simulation, the state of the circuit at any time instance is defined by a collection of voltages at nodes, some branch currents, and electrical charges. Depending on the interpretation of time, the system is called ‘continuous’ or ‘discrete’ when $t \in \mathbb{R}$ is continuous or discretely sampled, respectively.

The output of a dynamic system at a given time point does not only depend on

the input but also on the state of the system in the current and past time. Hence, for solving a dynamic system, the initial state of system must be known in order to calculate the system output function for a given input.

Definition 1. Assume that the outputs of a system model \mathcal{S} are $y_1(t)$ and $y_2(t)$ corresponding to inputs $x_1(t)$ and $x_2(t)$ respectively. \mathcal{S} is linear model iff

$$\alpha y_1(t) + \beta y_2(t) = \mathcal{S}\{\alpha x_1(t) + \beta x_2(t)\} \quad (2.1)$$

for all possible values of α and β .

2.2 Formulation of circuit equations

Modified Nodal Analysis (MNA) is widely used for formulating circuit equations in computer-aided network analysis [12]. In SPICE simulators such as HSPICE or PSpICE, the netlist description of any circuit is converted to MNA equation first, in order to simulate the circuit in frequency or time domain [13]. Since a circuit is a composition of components like resistors, capacitors, inductors, etc. The basic idea of this method is making a stamp for every component. Each stamp is representing the effect of the component by manipulating the appropriate entries in MNA matrices, namely \mathbf{G} , \mathbf{C} , and \mathbf{B} . Thus, the MNA formulation is a process of accumulating the stamps of all the components in the appropriate MNA matrices. This process can be

easily implemented by a computer program.

2.2.1 Linear-circuit equations

A general Linear Time Invariant (LTI) circuit can be described using the Modified Nodal Analysis (MNA) [14–16] as follows

$$\begin{aligned} \mathbf{C} \frac{d\mathbf{x}}{dt} &= -\mathbf{G}\mathbf{x} + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y} &= \mathbf{L}\mathbf{x} \end{aligned} \tag{2.2}$$

where $\mathbf{G}, \mathbf{C} \in \mathbb{R}^{n \times n}$ are conductance and susceptance matrices, respectively; $\mathbf{x}(t) \in \mathbb{R}^n$ is the vector of MNA variables (the nodal voltages, some branch currents and electrical charges). $\mathbf{B} \in \mathbb{R}^{n \times m_i}$ and $\mathbf{L} \in \mathbb{R}^{n \times m_o}$ respectively are constant input and output selection matrices, $\mathbf{u}(t)$ is the vector of the m_i input stimuli. n denotes the number of system variables, and m_i and m_o denote the number of system inputs and output, respectively.

Linear systems can be identified in the complex frequency domain as shown in (2.3). This is obtained by applying Laplace transform to (2.2) for zero initial condition.

$$\begin{cases} s\mathbf{C}\mathbf{X} &= \mathbf{F}\mathbf{X} + \mathbf{B}\mathbf{U}(s) \\ \mathbf{Y}(s) &= \mathbf{L}\mathbf{X} \end{cases} \tag{2.3}$$

where $s \in \mathbb{C}$ and $\mathbf{F} \triangleq -\mathbf{G}$.

The state variables denoted by $\mathbf{x}(t)$ is described by the first equation of (2.3) in the form of first order differential equations. Thus, given an initial state, the trajectory of the state can be obtained by solving (2.3) iteratively. This process is called transient analysis in circuit simulation. It is noticed that if a system is not stable, the state could locate at the infinity of the state space. As a result, the simulation results in time domain would be totally unreliable.

When matrix \mathbf{C} is nonsingular, it is called *regular* system. Otherwise, the system is called singular or *descriptor system*. For a regular system, (2.3) can be equivalently rewritten as shown in (2.4).

$$\begin{cases} s\mathbf{X} &= \mathbf{C}^{-1}\mathbf{F}\mathbf{X} + \mathbf{C}^{-1}\mathbf{B}\mathbf{U} \\ \mathbf{Y} &= \mathbf{C}\mathbf{X} \end{cases} \quad (2.4)$$

Regular systems enjoy certain properties making them easier to work with. A regular system of equations representing the circuit with the minimum possible unknowns (minimal realization) is referred to as state-space equation; and the unknowns vector \mathbf{X} is called states-vector. However, for general electrical circuits \mathbf{C} is not invertible.

2.3 Stability and passivity of linear system

Following the discussion of dynamic systems, we introduce the concept of stability and passivity.

2.3.1 Stability

Stability ensures that a system's state trajectory in state space does not go to infinity.

It is essential to preserve the stability property [17] of the original design in any mathematical representation of the circuit.

Poles of circuits

From the frequency domain representation of a linear circuit presented in (2.4), we get

$$(s\mathbf{C} + \mathbf{G})\mathbf{X} = \mathbf{B}\mathbf{U} \quad (2.5)$$

From (2.5), it then follows that

$$\begin{aligned} \mathbf{X} &= (s\mathbf{C} + \mathbf{G})^{-1} \mathbf{B}\mathbf{U} \\ &= (s\mathbf{G}^{-1}\mathbf{C} + \mathbf{I})^{-1} \mathbf{G}^{-1}\mathbf{B}\mathbf{U} \end{aligned} \quad (2.6)$$

From (2.6), the transfer function from the sources $\mathbf{U}(s)$ to the states is defined as

$$\mathbf{H}_x = (s\mathbf{G}^{-1}\mathbf{C} + \mathbf{I})^{-1} \mathbf{G}^{-1}\mathbf{B} \quad (2.7)$$

Let eigenvalues and eigenvectors of the matrix $\mathbf{G}^{-1}\mathbf{C}$ be denoted by $\mathbf{\Lambda}$ and \mathbf{T} , respectively. Then, by substituting $-\mathbf{G}^{-1}\mathbf{C} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$ in (2.7), we have

$$\mathbf{X} = \mathbf{T} (s\mathbf{\Lambda} + \mathbf{I})^{-1} \mathbf{T}^{-1} \mathbf{G}^{-1}\mathbf{B}\mathbf{U} \quad (2.8)$$

From (2.8), the poles of the system are given by

$$p_i = -\frac{1}{\lambda_i} \quad (2.9)$$

where $\lambda_i \neq 0$ is the i -th nonzero diagonal element of $\mathbf{\Lambda}$. Thus, following two theorems can be used to test the stability of a system.

Theorem 1. *System (2.3) is*

1. *asymptotically stable: if and only if all the nonzero eigenvalues of $-\mathbf{G}^{-1}\mathbf{C}$ have negative real parts.*
2. *stable: if and only if all nonzero eigenvalues of $-\mathbf{G}^{-1}\mathbf{C}$ have nonpositive real parts and all pure imaginary eigenvalues have multiplicity one.*

The following General Lyapunov Stability (GLE) can be used to check the stability of any general (linear) descriptor system [7].

Theorem 2 (General Lyapunov Stability). *System (2.2) is stable if there exists symmetric positive definite (SPD) matrix \mathbf{P} satisfying*

$$\mathbf{F}^T \mathbf{P} \mathbf{G} + \mathbf{F}^T \mathbf{P} \mathbf{C} \leq 0 \quad (2.10)$$

where $\mathbf{F} = -\mathbf{G}$

2.3.2 Passivity

A property stricter than stability is passivity. A stable circuit may not be passive [18], but a passive circuit must be stable. Furthermore, when a passive system is connected to other passive systems, the combined system is guaranteed to be passive. But this is not always true for stable systems.

Definition 2. An n-port system is said to be passive if

$$\int_0^t v^T(\tau) i(\tau) d\tau \geq 0 \quad (2.11)$$

for all t and all admissible port voltages $v(t)$ and currents $i(t)$

When a circuit network contains only RLC components, the circuit is passive and

by some adjustment of the MNA [5], the following condition can be satisfied

(a) matrix \mathbf{C} is symmetric positive semi-definite $\mathbf{C} \leq 0$

(b) $(\mathbf{G} + \mathbf{G}^T) \geq 0$

If the above condition in (b) is satisfied for a system it is called contractive system. These conditions are important for preserving stability in the order-reduction process and it will be further considered in the next chapter.

Chapter 3

Model order reduction: methods and properties

This chapter provides a review on the concepts, techniques, and properties of model order reduction (MOR), relevant to the subject of this thesis. The general theme in the MOR is that, given a large-scale dynamical system (linear or nonlinear) with predefined input and output terminals, a small-scale system is found that approximates the behavior of the large system at its terminals. To this end, the concepts and techniques from system theory and mathematics are adapted. The problem of model order reduction of linear and nonlinear dynamical systems has been widely studied in the last two decades and is still considered an active topic. Its applicability to real life problems has made it a popular tool in many branches of science and engineering.

In this chapter, typical non-Krylov and Krylov-based MOR methods are explained to provide a proper background of the proposed method introduced in the next chapter.

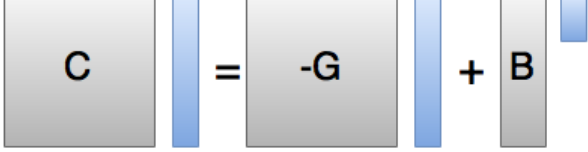
3.1 Projection Framework for MOR

For a system at a particular time point, its state is one of the infinite points in an n -dimensional state-space ($\subseteq \mathbb{R}^n$). The dimension of this space, which is the number of state variables n is called the order of the model.

Given a dynamic system, the movement of its state in the state-space generates a path called system trajectory. When simulating a circuit, practically, we would like to know the trajectory of the state in a given time interval. It is found that the trajectory is not evenly distributed in the state space, and can be considered as a geometrical structure representing the dynamic behavior of the circuit [19]. Projection-based MOR is based on the idea that trajectory can be projected onto some subspace and then recovered to the original space approximately. Figure 3.1 shows the principle of projection-based methods. Both Krylov-based and non-Krylov-based methods can be categorized as projection method.

Steps for Projection-Based Methods

1. Choose a right projection \mathbf{V} , and project original state space onto the column

$$\mathbf{C}\dot{\mathbf{x}} = -\mathbf{G}\mathbf{x} + \mathbf{B}\mathbf{u}$$


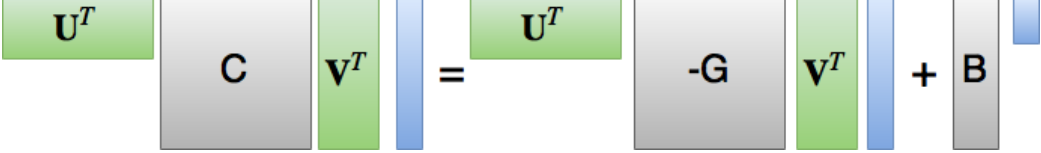
$$\mathbf{U}^T \mathbf{C} \mathbf{V} \dot{\mathbf{x}} = -\mathbf{U} \mathbf{G} \mathbf{V} \mathbf{x} + \mathbf{B} \mathbf{u}$$


Figure 3.1: Projection based model order reduction.

space of \mathbf{V} , where $\mathbf{x} = \mathbf{V}\mathbf{z}$

$$\mathbf{C}\mathbf{V}\dot{\mathbf{z}} = -\mathbf{G}\mathbf{V}\mathbf{z} + \mathbf{B}\mathbf{u} + \mathbf{R}(\mathbf{z}) \quad (3.1)$$

2. Choose a left projection matrix \mathbf{U} , and multiply both sides with the left projection matrix \mathbf{U}

$$\mathbf{U}^T \mathbf{C} \mathbf{V} \dot{\mathbf{z}} = -\mathbf{U}^T \mathbf{G} \mathbf{V} \mathbf{z} + \mathbf{U} \mathbf{B} \mathbf{u} + \mathbf{U}^T \mathbf{R}(\mathbf{z}) \quad (3.2)$$

3. Solving the system letting $\mathbf{R}(\mathbf{z})$ to be orthogonal to the column space of \mathbf{U}

$$\mathbf{U}^T \mathbf{R}(\mathbf{z}) = 0$$

By choosing different right or left projection matrix, we can have different MOR

methods.

3.2 Non-Krylov-Based Methods

The non-Krylov-based methods approximate the original system by seeking an optimal system in 2-induced norms for the system to minimize a certain error between the reduced and original models [20]. The well-known non-Krylov-based methods are presented as follows.

3.2.1 Truncated Balanced Realization

Truncated Balanced Realization (TBR) [21] is one of the commonly used MOR methods in some engineering field such as system control, where the systems are generally defined as

$$\begin{aligned} \mathbf{C} \frac{d\mathbf{x}}{dt} &= \mathbf{F}\mathbf{x} + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y} &= \mathbf{L}\mathbf{x} + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{3.3}$$

$\mathbf{D} \in \mathbb{R}^{n \times m_i}$ is the direct-coupling term.

Definition 3. Given stable system $(\mathbf{C}, \mathbf{F}, \mathbf{B}, \mathbf{L}, \mathbf{D})$, as shown in (3.3), reachability

gramian \mathbf{P} and observability gramian \mathbf{Q} is the solution of:

$$\mathbf{F}\mathbf{P}\mathbf{C}^T + \mathbf{C}\mathbf{P}\mathbf{F}^T + \mathbf{B}\mathbf{B}^T = \mathbf{0} \quad (3.4)$$

$$\mathbf{F}^T\mathbf{Q}\mathbf{C} + \mathbf{C}^T\mathbf{Q}\mathbf{F} + \mathbf{C}^T\mathbf{C} = \mathbf{0} \quad (3.5)$$

Principle of balancing

Based on the Gramians defined above, the minimum energy needed to drive the system from $\mathbf{0}$ to \mathbf{x} is given by $\mathbf{x}^T\mathbf{P}^{-1}\mathbf{x}$ and the largest observation energy by states \mathbf{x} without inputs is given by $\mathbf{x}^T\mathbf{Q}\mathbf{x}$. MOR by truncating balance is to remove the states that require large energy to reach while produce small observation energy. It is noticed that the states that are hard to reach are close to the space of eigenvectors corresponding to the small eigenvalues of \mathbf{P} , while the states that are hard to observe are the states close to the space of eigenvectors corresponding to the small eigenvalues of \mathbf{Q} . If we can make $\mathbf{P} = \mathbf{Q}$, then we can make sure that the states hard to reach are also hard to be observed, and make truncating the original system possible. The system with $\mathbf{P} \neq \mathbf{Q}$ is called unbalanced and vice versa. They are interconvertible by changing the basis of the state space. Truncation can be done for the balanced system.

Procedures of balanced truncating

1. Solve Lyapunov equation to find \mathbf{P} and \mathbf{Q} using (3.4) and (3.5).
2. Do Cholesky decomposition for \mathbf{P} and \mathbf{Q}

$$\mathbf{P} = \mathbf{U}\mathbf{U}^T, \mathbf{U}^T\mathbf{Q}\mathbf{U} = \mathbf{K}\mathbf{\Sigma}^2\mathbf{K}^T \quad (3.6)$$

3. Calculate the new basis \mathbf{T}

$$\mathbf{T} = \mathbf{\Sigma}^{1/2}\mathbf{K}^T\mathbf{U}^{-1} \quad (3.7)$$

4. Compute the balanced system $(\hat{\mathbf{E}}, \hat{\mathbf{F}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}})$.

$$(\hat{\mathbf{C}}, \hat{\mathbf{F}}, \hat{\mathbf{B}}, \hat{\mathbf{L}}, \hat{\mathbf{D}}) = (\mathbf{T}\mathbf{C}\mathbf{T}^{-1}, \mathbf{T}\mathbf{F}\mathbf{T}^{-1}, \mathbf{T}\mathbf{B}, \mathbf{L}\mathbf{T}^{-1}, \mathbf{D}) \quad (3.8)$$

It can be shown that the gramian of the balanced system $\hat{\mathbf{P}} = \hat{\mathbf{Q}} = \mathbf{\Sigma}$ and their eigenvectors are unit vectors, which are ordered by their reachability and observability. So we have next step

5. Partition the balanced system

$$(\hat{\mathbf{C}}, \hat{\mathbf{F}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}}) = \left(\begin{bmatrix} \hat{\mathbf{C}}_{11} & \hat{\mathbf{C}}_{12} \\ \hat{\mathbf{C}}_{21} & \hat{\mathbf{C}}_{22} \end{bmatrix}, \begin{bmatrix} \hat{\mathbf{F}}_{11} & \hat{\mathbf{F}}_{12} \\ \hat{\mathbf{F}}_{21} & \hat{\mathbf{F}}_{22} \end{bmatrix}, \begin{bmatrix} \hat{\mathbf{B}}_1 \\ \hat{\mathbf{B}}_2 \end{bmatrix}, \begin{bmatrix} \hat{\mathbf{L}}_1 & \hat{\mathbf{L}}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{D} \end{bmatrix} \right) \quad (3.9)$$

truncating the system to obtain the reduced system $(\hat{\mathbf{C}}_{11}, \hat{\mathbf{F}}_{11}, \hat{\mathbf{B}}_1, \hat{\mathbf{C}}_1, \mathbf{D})$

Properties of Balanced-truncation

1. Reduced system is balanced and stable if the original system is stable.
2. The error is bound. Assume $\sigma_{i,\dots,q}$ is the singular value of $\mathbf{\Sigma}$ and $\sigma_{1,\dots,k}$ is the singular value of $\hat{\mathbf{\Sigma}}$. Then the transfer function of original system \mathbf{H} and transfer function of ROM $\hat{\mathbf{H}}$ satisfy

$$\|\mathbf{H} - \hat{\mathbf{H}}\|_{\mathcal{H}_\infty} \leq 2(\sigma_{k+1} + \dots + \sigma_q) \quad (3.10)$$

where Hardy norm H_∞ is defined as $\sup_{\sigma>0} (\sup_{\omega} \|\mathbf{H}(\sigma + i\omega)\|)$, where $\|\mathbf{H}(s)\| = \sigma_1(\mathbf{H}(s))$

3. Computation cost for solving the generalized Lyapunov equation is $\mathcal{O}(n^3)$.

3.2.2 Hankel optimal model reduction

TBR has error bound in Hardy norm (H_∞) however, it was not clear whether the truncated system of any size was an optimal approximation of this size. Hankel-norm approximation, which is a method close to TBR in nature, generates ROM that is optimal in the Hankel norm [21]. The Hankel-norm approximation may not be categorized as a projection-based method in an exact sense. But, considering its

capability of generating stability guaranteed ROM, it is reviewed in this chapter.

Definition 4. Hankel operator \mathcal{H} is defined as

$$\mathcal{H} : u \longrightarrow y \triangleq \int_{-\infty}^0 h(t - \tau) u(\tau) d\tau \quad (3.11)$$

where $h(t)$ is the impulse response of the system for $t > 0$.

The maximal gain of this Hankel operator can be calculated as

$$\|\mathcal{H}\|_H = \sup_{u \in \mathcal{L}_2(-\infty, 0]} \frac{\|y\|_2}{\|u\|_2} \quad (3.12)$$

which is called the Hankel norm. For a given system (3.3), Hankel norm is the largest Hankel singularvalue of the system as

$$\|\Sigma\|_H = \sigma_1 = \sqrt{\max(\lambda(\mathbf{P}\mathbf{Q}))} \quad (3.13)$$

where \mathbf{P} and \mathbf{Q} are the observability and controllability Gramians as defined in (3.4) and (3.5), respectively.

The task of Hankel optimal model reduction of a matrix transfer function $\mathbf{H}(s)$ calls for finding a stable reduced system $\hat{\mathbf{H}}(s)$ of order less than a given positive integer q , such that the Hankel norm $\|\zeta(s)\|_H$ of the absolute error $\zeta(s) = \mathbf{H}(s) - \hat{\mathbf{H}}(s)$ is minimal [22].

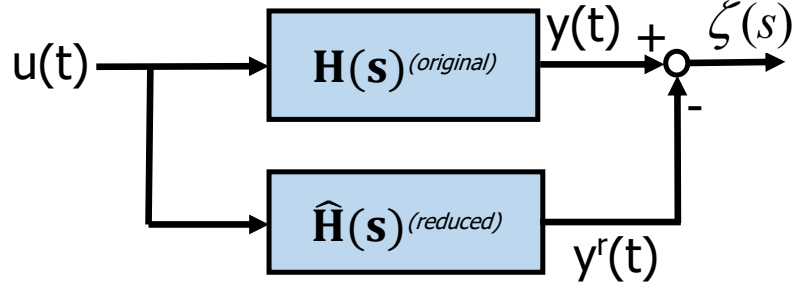


Figure 3.2: Absolute error $\|\zeta(s)\|_H$ should be minimal.

In [22] an algorithm based on a balanced realization approach is given which explicitly generates the reduced-order transfer function $\hat{\mathbf{H}}(s)$ of order q which minimizes the Hankel norm of the approximation error. By applying partial fraction expansion, $\hat{\mathbf{H}}(s)$ is decomposed into the summation of stable and unstable components, $\hat{\mathbf{H}}_{stable}(s)$ and $\hat{\mathbf{H}}_{unstable}(s)$, respectively as

$$\mathbf{H}(s) = \hat{\mathbf{H}}_{stable} + \hat{\mathbf{H}}_{unstable} \quad (3.14)$$

For all stable transfer functions $\hat{\mathbf{H}}_{stable}(s)$ of McMillan degree less than or equal to q , we have

$$\|\mathbf{H} - \hat{\mathbf{H}}\|_H = \sigma_{q+1} \leq \|\mathbf{H} - \hat{\mathbf{H}}_{stable}\|_H \quad (3.15)$$

There are explicit formulation to compute $\hat{\mathbf{H}}_{stable}(s)$ which is very involved. The detailed account of the derivation of $\hat{\mathbf{H}}_{stable}(s)$ falls beyond the scope of this review and can be referred to in [22–24].

Properties of Hankel-norm approximation

1. Requires computation of Gramians with the computation cost of $\mathcal{O}(n^3)$
2. The reduced system can only be optimal in Hankel-norm
3. Original system must be stable, and the reduced system is guaranteed to be stable

3.3 Krylov-Based Methods

The very large scale system arising in circuit simulation requires more efficient algorithms. The Krylov-based methods form an alternative and computationally efficient class of techniques, which achieve accuracy through a moment-matching mechanism.

Given a linear time invariant circuit (2.2), its transfer function in the complex frequency domain is given by

$$\mathbf{H}(s) = \mathbf{L}(\mathbf{G} + s\mathbf{C})^{-1}\mathbf{B} = \mathbf{L} \sum_{j=0}^{\infty} (-1)^j \mathbf{M}_j(s_0) (s - s_0)^j \quad (3.16)$$

where $\mathbf{M}_j \triangleq \mathbf{A}^j \mathbf{R}$ is the moments around s_0 , $\mathbf{A} \triangleq (\mathbf{G} + s_0 \mathbf{C})^{-1} \mathbf{C}$, $\mathbf{R} \triangleq (\mathbf{G} + s_0 \mathbf{C})^{-1} \mathbf{B}$.

It is normally chosen that $s_0 = 0$ in circuit simulation.

Definition 5. Given matrices $\mathbf{A}_{n \times n}$ and $\mathbf{R}_{n \times 1}$ as shown above, Krylov subspace is

defined as

$$\mathcal{K}r_k(\mathbf{A}, \mathbf{R}) = \text{colspan}\{\mathbf{R}, \mathbf{A}\mathbf{R}, \dots, \mathbf{A}^{k-1}\mathbf{R}\} \quad (3.17)$$

If we project the system onto a Krylov-subspace (Definition 5), it is called Krylov-based MOR [2, 25, 26].

The ROM $(\hat{\mathbf{C}}, \hat{\mathbf{G}}, \hat{\mathbf{B}}, \hat{\mathbf{L}})$ with order k should at least match the first k moments of the original system $(\mathbf{C}, \mathbf{G}, \mathbf{B}, \mathbf{L})$.

3.3.1 Arnoldi method

In the last subsection, it is found that moments compose to a Krylov-subspace. Thus, finding Krylov-subspace is crucial for moments matching. Arnoldi method as shown in Figure 3.3 is a numerically reliable way to find Krylov-subspace. It finds an orthogonal basis \mathbf{V}_k which span the column space of $[\mathbf{M}_1, \dots, \mathbf{M}_k]$ and satisfies

$$\mathbf{V}^T \mathbf{F} \mathbf{V} = \mathbf{H} \quad (3.18)$$

Algorithm 3.1: Arnoldi algorithm.

Data: \mathbf{F}, \mathbf{R}

Result: \mathbf{V}, \mathbf{H} , where $\mathbf{V}^T \mathbf{A} \mathbf{V} = \mathbf{H}$, \mathbf{H} is an upper Hessenberg matrix

$\mathbf{q}_1 = \mathbf{R}$;

for j *from* 1 *to* $k - 1$ **do**

$h_{j,k-1} = q_j^* q_k$;
 $q_k = q_k - h_{j,k-1} q_j$;

$h_{k,k-1} = \|q_k\|$;

$q_k = q_k / h_{k,k-1}$

It to be noted that, the algorithm described here is for systems with singular input.

But with slight modification, it could be extended to deal with multiple inputs.

$$\mathbf{A} \mathbf{V} = \mathbf{V} \mathbf{H} + \mathbf{R}$$

Figure 3.3: Arnoldi method, where entries of \mathbf{R} is zero except the last column.

Once the Krylov-subspace \mathbf{V}_k is found, the reduced system can be obtained by

Galerkin projection

$$\begin{aligned} \mathbf{V}_k^T \mathbf{C} \mathbf{V}_k \frac{d\mathbf{z}}{dt} &= -\mathbf{V}_k \mathbf{G} \mathbf{V}_k + \mathbf{V}_k^T \mathbf{B} \mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{L} \mathbf{V}_k \mathbf{z} \end{aligned} \quad (3.19)$$

From (3.19) the reduced system matrices are defined as

$$\hat{\mathbf{C}} \triangleq \mathbf{V}_k^T \mathbf{C} \mathbf{V}_k, \quad \hat{\mathbf{G}} \triangleq \mathbf{V}_k^T \mathbf{G} \mathbf{V}_k, \quad \hat{\mathbf{B}} \triangleq \mathbf{V}_k^T \mathbf{B}, \quad \hat{\mathbf{L}} \triangleq \mathbf{L} \mathbf{V}_k \quad (3.20)$$

Theorem (4) introduced in chapter 3, shows that moments can be matched just by the right projection matrix. There is considerable freedom to choose left projection matrix. Thus, it is possible to find a particular left projection matrix to preserve stability [7].

3.3.2 PRIMA

Passive Reduced-Order Interconnect Macromodeling Algorithm (PRIMA) is a variant of Arnoldi method for maintaining the passivity of the original system, which means the original system only contain passive components (resistors, capacitors, and inductors) while active components (controlled source) must not exist. Otherwise, ROM could be unstable. This issue prevents PRIMA's application to systems outside the class of RLC circuits [27]. The algorithm is composed of a slight modification of

stamps in MNA and applying the Arnoldi algorithm.

1. First modifying the original $(\mathbf{G}, \mathbf{C}, \mathbf{B})$ by multiplying both sides of inductance stamps by -1 , such that

$$\mathbf{C} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & -\mathbf{H} \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \mathbf{N} & \mathbf{E} \\ -\mathbf{E}^T & \mathbf{0} \end{bmatrix} \quad (3.21)$$

And notice that through this modification, the MNA parameters \mathbf{G}, \mathbf{C} satisfies

$$\mathbf{G} + \mathbf{G}^T \geq \mathbf{0} \text{ and } \mathbf{C} = \mathbf{C}^T \geq \mathbf{0}.$$

2. Then apply Arnoldi procedures.

Passivity Preserving

Theorem 3. *The necessary and sufficient conditions for the system admittance matrix $\mathbf{Y}(s) = \hat{\mathbf{L}}(\hat{\mathbf{G}} + s\hat{\mathbf{C}})^{-1}\hat{\mathbf{B}}$ to be passive (and hence stable) are the following [5]*

1. $\mathbf{Y}(s^*) = \mathbf{Y}^*(s)$ for all complex s , where $*$ is the complex conjugate operator.
2. $(\mathbf{Y}(s) + \mathbf{Y}^T(s^*)) \geq \mathbf{0}$ for all complex s with positive real part.

According to this theorem, it is proved that ROM $(\hat{\mathbf{G}}, \hat{\mathbf{C}}, \hat{\mathbf{B}}, \hat{\mathbf{L}})$ by PRIMA is passive.

Proof. the first condition is automatically satisfied since the parameters of the reduced system $(\hat{\mathbf{G}}, \hat{\mathbf{C}}, \hat{\mathbf{B}})$ are all real matrices. The remaining is to prove that $z^{*T}(\mathbf{Y}(s) +$

$\mathbf{Y}^*(s^*))z > 0$. By using the property $\hat{\mathbf{L}} = \hat{\mathbf{B}}^T$ and for any complex vector we have

$$\begin{aligned}
\mathbf{z}^{*T} (\mathbf{Y}(s) + \mathbf{Y}^T(s^*)) \mathbf{z} &= \mathbf{z}^{*T} [\hat{\mathbf{L}}(\hat{\mathbf{G}} + s\hat{\mathbf{C}})^{-1}\hat{\mathbf{B}} + \hat{\mathbf{B}}^T(\hat{\mathbf{G}} + s^*\hat{\mathbf{C}})^{-T}\hat{\mathbf{L}}^T] \mathbf{z} \\
&= \mathbf{z}^{*T} \hat{\mathbf{B}}^T [(\hat{\mathbf{G}} + s\hat{\mathbf{C}})^{-1} + (\hat{\mathbf{G}} + s^*\hat{\mathbf{C}})^{-T}] \hat{\mathbf{B}} \mathbf{z} \\
&= \mathbf{z}^{*T} \hat{\mathbf{B}}^T (\hat{\mathbf{G}} + s\hat{\mathbf{C}})^{-1} [(\hat{\mathbf{G}} + s\hat{\mathbf{C}}) + (\hat{\mathbf{G}} + s^*\hat{\mathbf{C}})^T] (\hat{\mathbf{G}} + s^*\hat{\mathbf{C}})^{-T} \hat{\mathbf{B}} \mathbf{z}
\end{aligned} \tag{3.22}$$

Setting $\boldsymbol{\psi} \triangleq (\hat{\mathbf{G}} + s^*\hat{\mathbf{C}})^{-T} \hat{\mathbf{B}} \mathbf{z}$ and $s = \sigma + j\omega$, we have

$$\begin{aligned}
\mathbf{z}^{*T} (\mathbf{Y}(s) + \mathbf{Y}^T(s^*)) \mathbf{z} &= \boldsymbol{\psi}^{*T} [(\hat{\mathbf{G}} + (\sigma + j\omega)\hat{\mathbf{C}}) + (\hat{\mathbf{G}} + (\sigma - j\omega)\hat{\mathbf{C}})^T] \boldsymbol{\psi} \\
&= \boldsymbol{\psi}^{*T} [(\hat{\mathbf{G}} + \hat{\mathbf{G}}^T) + \sigma(\hat{\mathbf{C}} + \hat{\mathbf{C}}^T)] \boldsymbol{\psi} \\
&= \boldsymbol{\psi}^{*T} \mathbf{V}_k^T [(\mathbf{G} + \mathbf{G}^T) + \sigma(\mathbf{C} + \mathbf{C}^T)] \mathbf{V}_k \boldsymbol{\psi} \tag{3.23}
\end{aligned}$$

let $\boldsymbol{\Phi} \triangleq \mathbf{V}_k \boldsymbol{\psi}$, we get

$$\mathbf{z}^{*T} (\mathbf{Y}(s) + \mathbf{Y}^T(s^*)) \mathbf{z} = \boldsymbol{\Phi}^{*T} [(\mathbf{G} + \mathbf{G}^T) + \sigma(\mathbf{C} + \mathbf{C}^T)] \boldsymbol{\Phi} \tag{3.24}$$

considering that,

$$\text{(a) } \boldsymbol{\Phi}^{*T} [(\mathbf{G} + \mathbf{G}^T)] \boldsymbol{\Phi} \geq 0$$

$$\text{(b) } \boldsymbol{\Phi}^{*T} [(\mathbf{C} + \sigma\mathbf{C}^T)] \boldsymbol{\Phi} = 2\sigma\boldsymbol{\Phi}^{*T} \mathbf{C} \boldsymbol{\Phi} \geq 0$$

This concludes the proof. □

PRIMA is ideal for MOR of linear passive circuits. But it cannot ensure the stability of the ROM even when the original system is stable since the method depends on the peculiar structure of matrices in the original system.

3.4 Non-Krylov versus Krylov based MOR

Efficiency: As explained in the previous section, all presented non-Krylov-based methods require computation time of $\mathcal{O}(n^3)$, which confines their application to small to medium scale systems. On the other hand, the CPU efficiency of the Krylov-based methods makes them a practical choice for model order reduction of the very large scale system arising in circuit simulation. Their main computation cost is to find Krylov-subspace, which can be efficiently accomplished by Arnoldi or Lanczos Method in a CPU cost of $\mathcal{O}(kn)$. This is much computationally inexpensive compared to the non-Krylov methods.

Accuracy: Non-Krylov-based methods usually obtain accuracy by minimizing the norm of system operator as shown in Figure 3.2. The ROM achieved by these methods has an error bound and preserves original system's stability, but the disadvantage is their expensive computation cost since the involvement of computing system Gramians by solving Lyapunov equations for a particular right hand side (RHS) definite

matrix. Meanwhile, ROM obtained by Krylov-based methods obtain accuracy by matching original system's moments.

Stability: Non-Krylov-based methods as reviewed in this chapter preserve the stability of the stable original system by construction. The main disadvantage for the Krylov-based methods is their inability to preserve the stability of general-linear systems. Yet those methods can still preserve stability for some special cases of linear systems.

3.5 Stability Preservation in Model Order Reduction

Considering the fact that the stability of active networks cannot be ensured in the reduction process using Krylov-based methods, it is important to enforce this attribute by some means. This section introduces two major approaches to tackle this issue.

3.5.1 Using Optimization algorithms

A class of the methods, to guarantee the stability of the reduced system, uses optimization algorithm to find a left projection \mathbf{U} instead of using the same right projection \mathbf{V} as in PRIMA. The moments are also matched since the right projection

matrix \mathbf{V} remains untouched.

Corollary 1. *The ROM $(\mathbf{U}^T \mathbf{G} \mathbf{V}, \mathbf{U}^T \mathbf{C} \mathbf{V}, \mathbf{U}^T \mathbf{B})$ is stable if \mathbf{U}, \mathbf{V} satisfy*

$$\mathbf{U}^T \mathbf{C} \mathbf{V} \geq \mathbf{0} \quad (3.25)$$

$$\mathbf{U}^T \mathbf{G} \mathbf{V} + \mathbf{V}^T \mathbf{G}^T \mathbf{U} \geq \mathbf{0} \quad (3.26)$$

Proof. According to Lyapunov theorem, the system is stable if

$$\exists \mathbf{L}(\mathbf{x}) \geq \mathbf{0} : \frac{\partial \mathbf{L}(\mathbf{x})}{\partial t} \leq \mathbf{0}, \forall \mathbf{x} \in \mathbb{R}^N$$

Let $\mathbf{L}(\mathbf{x}) = \mathbf{z}^T \mathbf{U}^T \mathbf{E} \mathbf{V} \mathbf{z}$

$$\frac{\partial \mathbf{L}(\mathbf{x})}{\partial t} = -\mathbf{U}^T \mathbf{G} \mathbf{V} - \mathbf{V}^T \mathbf{G}^T \mathbf{U} \leq \mathbf{0} \quad (3.27)$$

The condition of stability becomes

$$\mathbf{U}^T \mathbf{C} \mathbf{V} \geq \mathbf{0} \quad (3.28)$$

$$\mathbf{U}^T \mathbf{G} \mathbf{V} + \mathbf{V}^T \mathbf{G}^T \mathbf{U} \geq \mathbf{0} \quad (3.29)$$

□

Based on this corollary, we can find U by

$$\min_{\mathbf{U}, \mathbf{Q}_1, \mathbf{Q}_2} \|\mathbf{U} - \mathbf{U}_0\|_s \quad (3.30)$$

$$\mathbf{U}^T \mathbf{C} \mathbf{V} = \mathbf{Q}_1 > \mathbf{0} \quad (3.31)$$

$$\mathbf{U}^T \mathbf{G} \mathbf{V} + \mathbf{V}^T \mathbf{G}^T \mathbf{U} = \mathbf{Q}_2 > \mathbf{0} \quad (3.32)$$

Although this method can match moments and preserving stability, it uses optimization and hence, the existence of a solution is not promised.

3.5.2 Post-processing methods

The commonly available stability preserving methods in the literature [6–9] are usually based on post-processing. For example, in [9–11] a method is proposed to eliminate the unstable poles of the reduced system by using implicitly and explicitly restarted Arnoldi and Lanczos algorithms. However, the common concept in these methods is to sacrifice accuracy of the reduced model in order to guarantee stability, which may destroy the integrity of the moment-matching algorithm leading to inaccurate reduced model [6]. In addition, numerically restoring stability algorithms are not guaranteed to converge and in general, have relatively high computational cost associated with them.

3.6 Summary and Problem Definition

We have reviewed several typical MOR methods in this chapter. It is showed that, all those methods are facing some limitations in preserving stability in general cases. In one hand, The non-Krylov-based methods such as Truncated Balanced Realization (TBR) and Hankel-norm approximation based methods cannot be applied to our circuit model due to their prohibitive CPU cost. On the other hand, The existing Krylov-based methods despite being CPU efficient, they fail to preserve system stability for general linear systems. There are attempts reported in the literature to address this issue such as the post processing and optimization methods. However, they have not been proven successful for the general industrial class of problems.

To address the above difficulties, in the next two chapters, a stability-preserving projection algorithm for model reduction of active linear circuits is proposed. The core idea is that, the right and the left projection matrices are computed separately. The right projection matrix is constructed such that it guarantees accuracy through matching the moments and the left projection matrix is derived such that Lyapunov stability constraint is satisfied.

Chapter 4

The proposed stability guaranteed model reduction method

In this chapter, a novel projection-based method of constructing stable MOR from general linear model is presented. As explained in chapter 2, stability preservation is necessary for reliable time-domain simulation. Based on efficiently computing of the left and right projection matrices, the core idea is to compute the right and left projection matrices separately. The right-projection matrix is constructed to guarantee accuracy through moments matching and the left-projection matrix is computed to satisfy Lyapunov stability constraint. To this end, it is first proved that moments of the original system can be matched by only the right projection matrix. Then according to Lyapunov theorem, a form of left projection matrix is found. Finally, we show

how to solve Lyapunov equation cleverly. Thus, an efficient algorithm was devised for directly computing the ROM without using any post-processing or optimization method.

4.1 ROM Construction

Given a general LTI system described by Modified Nodal Analysis (MNA)

$$\begin{aligned}\mathbf{C} \frac{d}{dt} \mathbf{x}(t) &= -\mathbf{G} \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{L} \mathbf{x}(t)\end{aligned}\tag{4.1}$$

where \mathbf{C} and $\mathbf{G} \in \mathbb{R}^{n \times n}$ are susceptance and conductance matrices respectively, $\mathbf{x}(t) \in \mathbb{R}^n$ is the vector of MNA variables. $\mathbf{B} \in \mathbb{R}^{n \times r}$ and $\mathbf{L} \in \mathbb{R}^{s \times n}$ are the input and output matrices respectively. n denotes the number of original system's order, r denotes the number of inputs and s denotes the number of outputs.

We would like to construct a ROM

$$\begin{aligned}\hat{\mathbf{C}} \frac{d}{dt} \mathbf{z}(t) &= -\hat{\mathbf{G}} \mathbf{z}(t) + \hat{\mathbf{B}} \mathbf{u}(t) \\ \mathbf{y}(t) &\approx \hat{\mathbf{L}} \mathbf{z}(t)\end{aligned}\tag{4.2}$$

where $\hat{\mathbf{C}}$ and $\hat{\mathbf{G}} \in \mathbb{R}^{q \times q}$ are susceptance and conductance matrices of the ROM respec-

tively, $\mathbf{z}(t) \in \mathbb{R}^q$ is the vector of MNA variables. $\hat{\mathbf{B}} \in \mathbb{R}^{q \times r}$ and $\hat{\mathbf{L}} \in \mathbb{R}^{s \times q}$ are the input and output matrices respectively. q denotes the order of the reduced system.

The proposed method obtains the ROM by a Petrov-Galerkin projection [28] with a right projection matrix \mathbf{V} and a left projection \mathbf{U} . Petrov-Galerkin method first projects the original state variables to a smaller subspace as $\mathbf{X} = \mathbf{V}\mathbf{Z}$. A left projection matrix $\mathbf{U} \in \mathbb{R}^{m \times n}$ with full-column rank is then defined. The matrix \mathbf{U} introduced satisfies the orthogonality condition as $\mathbf{U}^T \mathbf{R}(s) = \mathbf{0}$, where $\mathbf{R}(s)$ is the residual error due to the approximation. Applying both projection matrices to (4.1), the reduced model is written as

$$\begin{aligned} \mathbf{U}^T \mathbf{C} \mathbf{V} \frac{d}{dt} \mathbf{z}(t) &= -\mathbf{U}^T \mathbf{G} \mathbf{V} \mathbf{z}(t) + \mathbf{U}^T \mathbf{B} \mathbf{u}(t) \\ \mathbf{y}(t) &\approx \mathbf{L} \mathbf{V} \mathbf{z}(t). \end{aligned} \quad (4.3)$$

which means

$$\begin{aligned} \hat{\mathbf{C}} &= \mathbf{U}^T \mathbf{C} \mathbf{V}, & \hat{\mathbf{G}} &= \mathbf{U}^T \mathbf{G} \mathbf{V} \\ \hat{\mathbf{B}} &= \mathbf{U}^T \mathbf{B}, & \hat{\mathbf{L}} &= \mathbf{L} \mathbf{V} \end{aligned} \quad (4.4)$$

Different projection-based MOR methods differ in the choice of matrices \mathbf{U} and \mathbf{V} . In the case of Galerkin projection, $\mathbf{U} = \mathbf{V}$ (i.e. a congruence transformations). Here, for the proposed method, $\mathbf{U} \neq \mathbf{V}$.

4.1.1 Construction of the Right Projection Matrix

The right projection matrix \mathbf{V} is constructed such that it guarantees accuracy through matching the moments. It can be computed through any Krylov-space based methods such as Block-Lanczos or Block-Arnoldi process. In this work, the orthonormal basis of the Krylov subspace are selected as the right projection matrix. And Block-Arnoldi process is used for this purpose.

Block-Arnoldi Algorithm

For a system with r inputs, Block-Arnoldi algorithm [29] calculates a projection matrix $\mathbf{V} \in \mathbb{R}^{m \times n}$, whose columns span a Krylov subspace defined as

$$\text{colspan}\{\mathbf{V}\} = Kr(\mathbf{A}, \mathbf{R}, k) = \text{span}\{\mathbf{R}, \mathbf{A}\mathbf{R}, \dots, \mathbf{A}^{(k-1)}\mathbf{R}\} \quad (4.5)$$

where $m = k \times r$. Specifically, the block Arnoldi algorithm reduces the matrix $\mathbf{A} \triangleq -\mathbf{G}^{-1}\mathbf{C}$ to a block upper Hessenberg matrix $\mathbf{H}_m \in \mathbb{R}^{m \times m}$. The algorithm involves successively filling in the columns of \mathbf{V} in the relation $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{H}_m$ from the starting matrix \mathbf{R} , subject to $\mathbf{V}^T\mathbf{V} = \mathbf{I}_m$, where \mathbf{I}_m is an $m \times m$ identity matrix. The orthonormal matrix \mathbf{V} is selected as the right-projection matrix in this work.

4.1.2 Construction of the Left Projection Matrix

The left projection matrix \mathbf{U} is different from the right projection matrix and plays the role of guaranteeing stability of the ROM through the satisfaction of the Lyapunov constraint. To this end, given a right projection matrix $\mathbf{V} \in \mathbb{R}^{n \times q}$, the left projection matrix is computed as

$$\mathbf{U}^T = \mathbf{V}^T \mathbf{C}^T \mathbf{P} \in \mathbb{R}^{m \times n}, \quad (4.6)$$

where \mathbf{P} is any symmetric positive-definite matrix satisfying the Lyapunov equation

$$\mathbf{C}^T \mathbf{P} \mathbf{F} + \mathbf{F}^T \mathbf{P} \mathbf{C} = -\mathbf{Q}_A \quad (4.7)$$

where $\mathbf{F} = -\mathbf{G}$. The right-hand-side matrix $\mathbf{Q}_A \in \mathbb{R}^{n \times n}$ can be any arbitrarily decided positive definite matrix $\mathbf{Q}_A > \mathbf{0}$. One of the possible choices in general practical problems is the identity matrix $\mathbf{Q}_A = \mathbf{I}_{n \times n}$.

Having the right projection matrix \mathbf{V} from Arnoldi process and a Lyapunov solution matrix \mathbf{P} , ROM can be written as

$$\begin{aligned} (\mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V}) \dot{\mathbf{z}} &= -(\mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{G} \mathbf{V}) \mathbf{z} + (\mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{B}) \mathbf{u}(t) \\ \mathbf{y} &= (\mathbf{L} \mathbf{V}) \mathbf{z} \end{aligned} \quad (4.8)$$

Hence, for the stable macromodel, the reduced-order MNA matrices are

$$\begin{aligned}\hat{\mathbf{C}} &= \mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V}, & \hat{\mathbf{G}} &= \mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{G} \mathbf{V}, \\ \hat{\mathbf{B}} &= \mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{B}, \text{ and } & \hat{\mathbf{L}} &= \mathbf{L} \mathbf{V}.\end{aligned}\tag{4.9}$$

4.2 Properties of the proposed method

The reduced model from the proposed method has two important properties: moments matching and stability guarantee. This section gives the proof for these two properties.

4.2.1 Moments Matching

Accuracy of our ROM is guaranteed by moments matching, where moments are the components of the transfer function. The transfer function of input to output of (4.1) is given by

$$\mathbf{H}(s) = \mathbf{L} (\mathbf{G} + s\mathbf{C})^{-1} \mathbf{B}\tag{4.10}$$

and it can be expanded to Taylor series around $s = 0$

$$\mathbf{H}(s) = \mathbf{L} \sum_{j=0} \mathbf{M}_j s^j\tag{4.11}$$

The j -th moment \mathcal{M}_j of system (4.1) is defined as

$$\mathcal{M}_j \triangleq \mathbf{L}\mathbf{M}_j = \mathbf{L}\mathbf{A}^j \mathbf{R} \quad (4.12)$$

where $\mathbf{A} \triangleq -\mathbf{G}^{-1}\mathbf{C}$ and $\mathbf{R} \triangleq \mathbf{G}^{-1}\mathbf{B}$.

Similarly, we define the j -th moment $\hat{\mathcal{M}}_j$ of reduced system (4.1) as

$$\hat{\mathcal{M}}_j \triangleq \hat{\mathbf{L}}\hat{\mathbf{M}}_j = \hat{\mathbf{L}}\hat{\mathbf{A}}^j\hat{\mathbf{R}} \quad (4.13)$$

where $\hat{\mathbf{A}} \triangleq -\hat{\mathbf{G}}^{-1}\hat{\mathbf{C}}$ and $\hat{\mathbf{R}} \triangleq \hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}$.

Since moments of a function is the coefficients for its Taylor expansion, to represent a function, we don't need all moments, we can choose a integer q and use only the first q moments and ignore the tail moments. The value of q depends on how much accuracy we want. The proposed method can preserve the same number of moments as the order of the ROM.

We will see that the reduced macromodel (4.8) preserves the first $k = \lfloor m/r \rfloor$ block moments of the original system, which is the same as the classical block Arnoldi reduction. The block moments of the original system, as shown in (4.12), for $s_0 = 0$ (without the loss of generality) are

$$\mathcal{M}_j = \mathbf{L}\mathbf{M}_j = \mathbf{L}\mathbf{A}^j \mathbf{R} \quad (4.14)$$

where $\mathbf{A} \triangleq -\mathbf{G}^{-1}\mathbf{C}$ and $\mathbf{R} \triangleq \mathbf{G}^{-1}\mathbf{B}$.

Similarly, the moments of the reduced-order system are given by

$$\hat{\mathcal{M}}_j = \hat{\mathbf{L}}\hat{\mathbf{M}}_j = \hat{\mathbf{L}}\hat{\mathbf{A}}^j\hat{\mathbf{R}} \quad (4.15)$$

where $\hat{\mathbf{A}} \triangleq -\hat{\mathbf{G}}^{-1}\hat{\mathbf{C}}$ and $\hat{\mathbf{R}} \triangleq \hat{\mathbf{G}}^{-1}\hat{\mathbf{B}}$ and matrices $\hat{\mathbf{G}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{B}}$, and $\hat{\mathbf{L}}$ are defined in (4.9).

To show that the first k block moments in (4.1) and (4.8) are identical, we first define the following Proposition 1 and Lemma 1, which are used later in the proof of Theorem 4.

Proposition 1. *Given a reduced-order full column-rank projection matrix $\mathbf{V} \in \mathbb{R}^{n \times m}$, there exists an $N \times m$ matrix \mathbf{E}_j , such that*

$$\mathbf{A}^j\mathbf{R} = \mathbf{V}\mathbf{E}_j, \quad 0 \leq j < k. \quad (4.16)$$

Proof. From the definition of Krylov subspace (4.5) it is deduced that, associated with every (j -th) moment, there exists a projection of it in the Krylov subspace. Let a $n \times m$ matrix \mathbf{E}_j be the projection of the block moment \mathbf{M}_j into Krylov subspace (4.5). We thus have

$$\mathbf{M}_j = \mathbf{V}\mathbf{E}_j. \quad (4.17)$$

Substituting $\mathbf{M}_j = \mathbf{A}^j\mathbf{R}$ in (4.17) proves the relationship in (4.16). \square

Lemma 1. *Given the right projection matrix \mathbf{V} as specified in Proposition 1 and \mathbf{U} in (4.6), a $n \times n$ matrix as*

$$\mathbf{\Phi} \triangleq -\mathbf{V} \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \mathbf{U}^T \mathbf{C} \quad (4.18)$$

satisfies the relation

$$\mathbf{\Phi}^j \mathbf{R} = \mathbf{A}^j \mathbf{R}, \quad 0 \leq j < k. \quad (4.19)$$

Proof. The proof is possible by induction on j . First, it is trivial to prove (4.19) for $j = 0$. Next, assume that (4.19) is true for any $j - 1$ when $0 < j < k$, that is

$$\mathbf{\Phi}^{j-1} \mathbf{R} = \mathbf{A}^{j-1} \mathbf{R}. \quad (4.20)$$

Multiplying both sides of (4.20) by $\mathbf{A} \triangleq -\mathbf{G}^{-1} \mathbf{C}$ yields

$$-\mathbf{G}^{-1} \mathbf{C} \times \mathbf{\Phi}^{j-1} \mathbf{R} = \mathbf{A}^j \mathbf{R}. \quad (4.21)$$

Substituting (4.16) from Proposition 1 in (4.20), we have

$$-\mathbf{G}^{-1} \mathbf{C} \times \mathbf{\Phi}^{j-1} \mathbf{R} = \mathbf{V} \mathbf{E}_j. \quad (4.22)$$

Multiplying both sides of (4.22) with an $n \times n$ matrix $\mathbf{V}(\mathbf{U}^T \mathbf{G} \mathbf{V})^{-1} \mathbf{U}^T \mathbf{G}$, we obtain

$$-\mathbf{V}(\mathbf{U}^T \mathbf{G} \mathbf{V})^{-1} \mathbf{U}^T \mathbf{G} \times \mathbf{G}^{-1} \mathbf{C} \times \Phi^{j-1} \mathbf{R} = \mathbf{V}(\mathbf{U}^T \mathbf{G} \mathbf{V})^{-1} (\mathbf{U}^T \mathbf{G} \mathbf{V}) \mathbf{E}_j \quad (4.23)$$

then,

$$-\mathbf{V}(\mathbf{U}^T \mathbf{G} \mathbf{V})^{-1} \mathbf{U}^T \mathbf{C} \times \Phi^{j-1} \mathbf{R} = \mathbf{V} \mathbf{E}_j. \quad (4.24)$$

Using (4.16) and (4.18), (4.24) leads to $\Phi^j \mathbf{R} = \mathbf{A}^j \mathbf{R}$ which is the desired relation in (4.19). \square

Following the Proposition 1 and Lemma 1 established above, the following Theorem 4 is proved.

Theorem 4. *Using the proposed left and right projection matrices \mathbf{U} and \mathbf{V} , the reduced system in (4.8) matches the first k block moments of the original system.*

Proof. In this proof we establish that

$$\mathcal{M}_j = \hat{\mathcal{M}}_j, \quad 0 \leq j < k. \quad (4.25)$$

which is logically equivalent to

$$\mathbf{M}_j = \mathbf{V} \hat{\mathbf{M}}_j, \quad 0 \leq j < k. \quad (4.26)$$

I) For $j = 0$: To show that the first two block moments are equal, for $j = 0$, from Proposition 1, we have

$$\mathbf{R} = \mathbf{V}\mathbf{E}_0. \quad (4.27)$$

Multiplying both sides of (4.27) by the matrix $\mathbf{V}(\mathbf{U}^T\mathbf{G}\mathbf{V})^{-1}\mathbf{U}^T\mathbf{G}$, used in the proof of Lemma 1, we have

$$\mathbf{V}(\mathbf{U}^T\mathbf{G}\mathbf{V})^{-1}\mathbf{U}^T\mathbf{G} \times \mathbf{R} = \mathbf{V}(\mathbf{U}^T\mathbf{G}\mathbf{V})^{-1}\mathbf{U}^T\mathbf{G}\mathbf{V}\mathbf{E}_0. \quad (4.28)$$

Substituting $\mathbf{R} \triangleq \mathbf{G}^{-1}\mathbf{B}$,

$$\mathbf{V}(\mathbf{U}^T\mathbf{G}\mathbf{V})^{-1}\mathbf{U}^T\mathbf{G} \times \mathbf{G}^{-1}\mathbf{B} = \mathbf{V}(\mathbf{U}^T\mathbf{G}\mathbf{V})^{-1}(\mathbf{U}^T\mathbf{G}\mathbf{V})\mathbf{E}_0. \quad (4.29)$$

it will be

$$\mathbf{V}(\mathbf{U}^T\mathbf{G}\mathbf{V})^{-1}(\mathbf{U}^T\mathbf{B}) = \mathbf{V}\mathbf{E}_0. \quad (4.30)$$

Next, substituting (4.27) in (4.30), we get

$$\mathbf{V}\hat{\mathbf{G}}^{-1}\hat{\mathbf{B}} = \mathbf{R} \quad (4.31)$$

or equivalently

$$\mathbf{V}\hat{\mathbf{R}} = \mathbf{R} \quad (4.32)$$

According to the definition of the moments for the original and reduced system in (4.14) and (4.15), respectively, (4.32) can be rewritten as

$$\mathbf{M}_0 = \mathbf{V}\hat{\mathbf{M}}_0. \quad (4.33)$$

By multiplying both sides by output matrix \mathbf{L} , we get

$$\mathbf{L}\mathbf{M}_0 = \mathbf{L}\mathbf{V}\hat{\mathbf{M}}_0 = \hat{\mathbf{L}}\hat{\mathbf{M}}_0 \quad (4.34)$$

resulting in $\mathcal{M}_0 = \hat{\mathcal{M}}_0$ that establishes the matching of the first block moments.

II) For $0 < j < k$: From (4.15), we have

$$\mathbf{V}\hat{\mathbf{M}}_j = \mathbf{V}\hat{\mathbf{A}}^j\hat{\mathbf{R}} = \mathbf{V} \times (-1)^j (\hat{\mathbf{G}}^{-1}\hat{\mathbf{C}})^j \times \hat{\mathbf{R}}. \quad (4.35)$$

Using the definitions of the reduced matrices in (4.9) it is rewritten as

$$\begin{aligned}
\mathbf{V}\hat{\mathbf{M}}_j &= \mathbf{V} \times (-1)^j \left(\left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \left(\mathbf{U}^T \mathbf{C} \mathbf{V} \right) \right)^j \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \mathbf{U}^T \mathbf{B} \\
&= \mathbf{V} \times \left(- \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \left(\mathbf{U}^T \mathbf{C} \mathbf{V} \right) \right) \times \left(- \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \left(\mathbf{U}^T \mathbf{C} \mathbf{V} \right) \right) \times \dots \\
&\quad \times \left(- \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \left(\mathbf{U}^T \mathbf{C} \mathbf{V} \right) \right) \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \mathbf{U}^T \mathbf{B} \\
&= \left(-\mathbf{V} \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \mathbf{U}^T \mathbf{C} \right) \times \left(-\mathbf{V} \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \mathbf{U}^T \mathbf{C} \right) \times \dots \\
&\quad \times \left(-\mathbf{V} \left(\mathbf{U}^T \mathbf{G} \mathbf{V} \right)^{-1} \mathbf{U}^T \mathbf{C} \right) \mathbf{V} \hat{\mathbf{R}}.
\end{aligned} \tag{4.36}$$

Next, substituting substituting (4.18) and (4.32) in(4.35) we get

$$\mathbf{V}\hat{\mathbf{M}}_j = \boldsymbol{\Phi} \times \boldsymbol{\Phi} \times \dots \times \boldsymbol{\Phi} \times \mathbf{R} = \boldsymbol{\Phi}^j \mathbf{R}. \tag{4.37}$$

Using (4.19) and (4.14) from (4.37), we have

$$\mathbf{V}\hat{\mathbf{M}}_j = \boldsymbol{\Phi}^j \mathbf{R} = \mathbf{A}^j \mathbf{R} = \mathbf{M}_j. \tag{4.38}$$

Multiplying both sides by output matrix \mathbf{L} and considering (4.15), we have

$$\mathbf{L}\mathbf{V}\hat{\mathbf{M}}_j = \hat{\mathbf{L}}\hat{\mathbf{M}}_j = \mathbf{L}\mathbf{M}_j \tag{4.39}$$

Hence, it is

$$\hat{\mathcal{M}}_j = \mathcal{M}_j \quad (4.40)$$

which concludes the proof. \square

4.2.2 Stability Guarantee

Theorem 5 (Lyapunov function). *Let $\mathcal{L}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ be a positive definite function, i.e., $\mathcal{L}(\mathbf{0}) = 0$ and $\mathcal{L}(\mathbf{x}) > 0$, for $\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}$, if $\frac{d\mathcal{L}(\mathbf{x})}{dt} < 0$, for $\mathbf{x} \in \mathbb{R}^n$, the system (4.1) is asymptotically stable and $\mathcal{L}(\mathbf{x})$ is called Lyapunov function [30].*

Lemma 2. *Using the proposed left and right projection matrices, the reduced system in (4.8) preserves the stability of the original system.*

Proof. For the resulting reduced models, a natural choice for the Lyapunov function can be in the following quadratic form

$$\mathcal{L}(z(t)) = \mathbf{z}^T \mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V} \mathbf{z}, \quad (4.41)$$

where $\mathbf{z}(t)$ is the autonomous response of the reduced model for any initial condition

$\mathbf{x}_0 \in \mathbb{R}^m$. Using (4.41)

$$\begin{aligned} \frac{d\mathcal{L}(z(t))}{dt} &= \frac{d\mathbf{z}^T}{dt} \mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V} \mathbf{z} + \mathbf{z}^T \mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V} \frac{d\mathbf{z}}{dt} \\ &= (\mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V} \frac{d\mathbf{z}}{dt})^T \mathbf{z} + \mathbf{z}^T (\mathbf{V}^T \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{V} \frac{d\mathbf{z}}{dt}) \end{aligned} \quad (4.42)$$

Using (4.3) in the autonomous form ($\mathbf{u}(t) = \mathbf{0}$) and (4.41), (4.42) can be rewritten as

$$\begin{aligned} \frac{d\mathcal{L}(z(t))}{dt} &= \mathbf{z}^T \mathbf{V}^T (-\mathbf{G})^T \mathbf{C}^T \mathbf{P} \mathbf{V} \mathbf{z} + \mathbf{z}^T \mathbf{V}^T \mathbf{C}^T \mathbf{P} (-\mathbf{G}) \mathbf{V} \mathbf{z} \\ &= \mathbf{z}^T \mathbf{V}^T (\mathbf{F}^T \mathbf{P} \mathbf{C} + \mathbf{C}^T \mathbf{P} \mathbf{F}) \mathbf{V} \mathbf{z} \end{aligned} \quad (4.43)$$

Since \mathbf{P} in (4.6) satisfies the Lyapunov equation in (4.7), the kernel in (4.43) is negative definite. And the congruence transformations preserves the definiteness of the matrices, the derivative of Lyapunov function for the reduced system in (4.43) is still negative definite. Hence, according to the Theorem 5, for an asymptotically stable original system, the reduced models of any order are ensured to be asymptotically stable.

□

4.3 Computational Complexity

The main computational complexity in the proposed method (over the Arnoldi-process used to form \mathbf{V}) is associated with finding a symmetric positive-definite matrix \mathbf{P} to satisfy the Lyapunov equation in (4.7). It is noticed that the proposed method does not depend on the solution of the Lyapunov equation with specific right hand-side matrix, which is the case for Truncated Balanced Realization (TBR). Moreover, in the presence of singular matrix \mathbf{C} of the model, which is inevitable for circuit equations, both existence and uniqueness of the solution for a Lyapunov equation with specific right-hand side is not guaranteed. However, in the proposed method, we only need to find one solution of the Lyapunov equation with any right-hand side definite matrix.

Formulation in TBR is based on the adequately accurate solutions of two Lyapunov equations for a certain right-hand side matrices. This fact makes application of TBR on circuit simulation problems very challenging and burdensome and uncertain. Whereas, for the proposed method, any matrix \mathbf{P} that can hold the Lyapunov inequality is adequate.

In the next chapter, I developed a computationally inexpensive algorithms to solve the sparse inequality in (4.7) with large matrices.

equivalent circuit is shown in Figure 4.3.

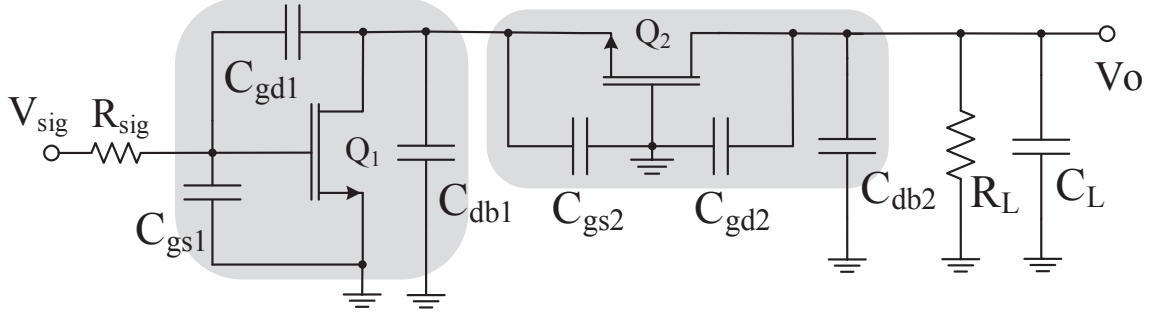


Figure 4.3: High-frequency equivalent-circuit model for MOS cascade amplifier (the biasing network is not presented).

Applying the proposed method, a stable reduced-order model for the original (stable) network in Figure 4.1 is obtained. The accuracy of the resulting reduced model is demonstrated in Figure 4.4, where the output response V_{out2} of the reduced model is compared with the simulation results of the original (unreduced) model and the PRIMA-based reduced model. Both reduced models are of the same order.

The sizes and stability properties of the original and reduced models are compared in Table 4.1.

Table 4.1: Comparison of the original and reduced models.

	Dimension	Stability
Original Circuit	4519	Yes
PRIMA-based Reduced Model	30	No
Proposed Reduced Model	30	Yes

In Table 4.2, first few dominant poles of the purposed reduced model are compared with those obtained using a PRIMA-based approach. It depicts that, the poles for the reduced model obtained through the proposed method lie in the left-half complex

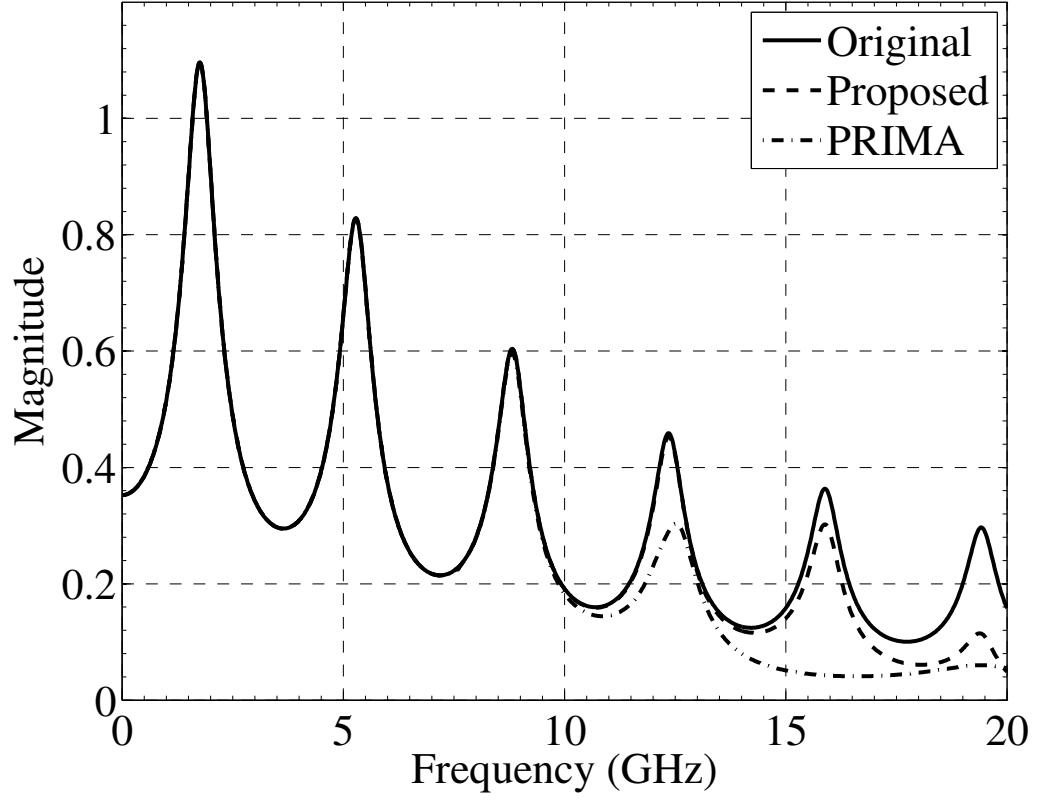


Figure 4.4: Comparison of the responses at V_{out2} from the proposed method, the PRIMA-based model and the original circuit.

plane. While the Galerkin-based scheme (i.e. PRIMA-Based) for reduction does not result in stable reduced model.

Frequency-domain response of the original circuit and the proposed reduced model are in excellent agreement within the intended frequency range compared to the PRIMA-based model. This is achieved besides the fact that, the proposed reduced model provided an accurate time-domain response, whereas the time-domain simulation of the PRIMA model failed to converge because of the unstable poles.

Table 4.2: Comparison of few dominant poles.

PRIMA-based	Proposed
+2.0113e08, -8.7566e06	-8.7566e06
-9.8022e06 ± 2.2684e10i	-9.9983e06 ± 2.2684e10i
-1.9901e09 ± 7.6647e10i	-1.5962e08 ± 4.7654e10i
-2.1766e09 ± 1.1209e10i	-2.1766e09 ± 1.1209e10i
-2.2110e09 ± 5.6743e10i	-2.2265e09 ± 3.3313e10i

Chapter 5

Efficient Solution for Generalized Lyapunov equation

In the previous chapter, we explored the framework of the proposed moment-matching-based model order reduction (MOR) algorithm, which constructs a guaranteed stable reduced order model (ROM) by solving a Lyapunov equation. Also, it was pointed out that the dominant computational complexity in the proposed method is associated with solving the Lyapunov equation in (4.7) to find a symmetric positive-definite solution matrix \mathbf{P} .

5.1 Problem Definition

The only method in the MOR literature that can construct stable ROM for any original stable systems is truncated balanced realization (TBR). To obtain a reduced model, TBR requires the solution of two dual Lyapunov equations of the form

$$\mathbf{C}\mathbf{W}_c\mathbf{F}^T + \mathbf{F}\mathbf{W}_c\mathbf{C}^T = -\mathbf{B}\mathbf{B}^T \quad (5.1)$$

$$\mathbf{C}^T\mathbf{W}_o\mathbf{F} + \mathbf{F}^T\mathbf{W}_o\mathbf{C} = -\mathbf{L}\mathbf{L}^T \quad (5.2)$$

where \mathbf{W}_c defines the “Controllability Gramian” and \mathbf{W}_o the “Observability Gramian” of system (4.1) for $\mathbf{F} \triangleq -\mathbf{G}$ [20]. Although good direct numerical solvers are available for (5.1) and (5.2), TBR for circuit simulation applications still suffers from certain drawbacks such as (a) for large circuits where n becomes large solving (5.1)-(5.2) might be prohibitively CPU-time and memory demanding, (b) the matrix \mathbf{C} is generally singular and hence, the existence and uniqueness of the solution matrices, \mathbf{W}_c and \mathbf{W}_o are not guaranteed.

In the proposed method, presented in the previous chapter, the left-projection matrix was expeditiously obtained using Arnoldi algorithm, while the right-projection matrix is formed through solving a single Lyapunov equation under a relaxed condition as shown in (5.3).

$$\mathbf{C}^T\mathbf{P}\mathbf{F} + \mathbf{F}^T\mathbf{P}\mathbf{C} = -\mathcal{X}. \quad (5.3)$$

It is relaxed in the sense that, the proposed method does not depend on the solution of the Lyapunov equation for any particular right-hand-side (RHS) matrix; instead, any well-suited symmetric positive-definite matrix can be used as RHS matrix.

Superiority: The following facts are also to be noted, indicating the superiority of the proposed method compared to the existing balanced-realization-based method.

- (a) The steps of the proposed method includes solving only one Lyapunov equation, which immediately results in an at-least two times speed-up in model generation time compared to TBR
- (b) $\mathcal{X} = \mathbf{I}_{n \times n}$ can be a trivial choice in many practical applications
- (c) Any efficient algorithm that can compute an approximate solution $\mathbf{P} = \tilde{\mathbf{P}}$ for (5.3) can be employed to form the right-projection matrix
- (d) In general, any approximated positive-definite matrix $\tilde{\mathbf{P}}$ satisfying the Lyapunov matrix inequality as shown in (5.4) can be used as a solution for the proposed method.

$$\mathbf{C}^T \tilde{\mathbf{P}} \mathbf{F} + \mathbf{F}^T \tilde{\mathbf{P}} \mathbf{C} < \mathbf{0} \quad (5.4)$$

- (e) Using an approximate solution $\tilde{\mathbf{P}}$ does not lead to any accuracy degradation in the resulting reduced model

- (f) Uniqueness of the possible solutions is not a required condition for the proposed method

Beside all the strength discussed above, a fact should not be ignored that, the dominant CPU-expenses associated with the method is still the cost of solving one (relaxed) Lyapunov equation. Any improvement in solving Lyapunov equation directly meliorate the efficiency of the proposed method.

In the rest of this chapter, an efficient algorithm to solve the Lyapunov matrix inequality as shown in (5.5) with large sparse matrices is presented.

$$\mathbf{C}^T \mathbf{P} \mathbf{F} + \mathbf{F}^T \mathbf{P} \mathbf{C} < \mathbf{0} \quad (5.5)$$

5.2 Background and Preliminaries

We consider LTI circuits realized using MNA formulation in the form of continuous time linear descriptor systems (CLDS) as given in (4.1) and is repeated here for the eases of reference.

$$\mathbf{C} \frac{d\mathbf{x}(t)}{dt} = \mathbf{F} \times \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t) \quad (5.6)$$

where $\mathbf{F} = -\mathbf{G}$, \mathbf{C} and $\mathbf{G} \in \mathbb{R}^{n \times n}$ are susceptance and conductance sparse matrices respectively, and $\mathbf{x}(t) \in \mathbb{R}^n$ is the vector of 'n' MNA variables.

Definition 6 (Regular system). CLDS system (5.6) is regular if there exists $\lambda \in \mathbb{C}$

such that $\det(\lambda \mathbf{C} - \mathbf{F}) \neq 0$.

Theorem 6 (Asymptotic stability of CLDS). *The circuit represented in (5.6) is asymptotically stable if and only if all finite eigenvalues of the matrix pencil $\lambda \mathbf{C} - \mathbf{F}$ lie in \mathbb{C}^- (the open left half-plane).*

A complex value $\lambda \neq \infty$ is said to be a *generalized finite eigenvalue* of the matrix pencil $(\lambda \mathbf{C} - \mathbf{F})$ if $\det(\lambda \mathbf{C} - \mathbf{F}) = 0$. The pencil $(\lambda \mathbf{C} - \mathbf{F})$ has *infinite eigenvalue* if and only if the matrix \mathbf{C} is singular.

When matrix \mathbf{C} is singular in (5.6), it is called a *generalized state-space system* or a *singular* or *descriptor* system [32], which is the case for the general circuit (MNA) equations. The generalized Lyapunov equations corresponding to the circuit in (5.6) is written as

$$\mathbf{C}^T \mathbf{P} \mathbf{F} + \mathbf{F}^T \mathbf{P} \mathbf{C} = -\mathbf{Q}_A \quad (5.7)$$

where matrix \mathbf{C} is generally singular.

5.3 Existence and Uniqueness of the Solutions of GLE

Case-I. Nonsingular Systems: For nonsingular systems where matrix \mathbf{C} is full-rank ($\text{rank}(\mathbf{C}) = n$) and hence invertible, (5.7) can be equivalently rewritten in the

normal form (for state-space systems) as

$$\mathbf{P}(\mathbf{F}\mathbf{C}^{-1}) + (\mathbf{F}\mathbf{C}^{-1})^T \mathbf{P} = -\mathbf{C}^{-T} \mathbf{Q}_A \mathbf{C}^{-1} \quad (5.8)$$

Hence, the classical result about the positive definite solution of the stable Lyapunov equation [33] remains valid for the generalized equations. Therefore, for the stable circuit in (5.6) where \mathbf{C} is nonsingular, a unique matrix $\mathbf{P} > \mathbf{0}$ always exist for any $\mathbf{Q}_A > \mathbf{0}$. This fact for nonsingular systems is distilled in the following corollary which is directly established based on the Theorem 1 in [34].

Theorem 7 (GLE with nonsingular \mathbf{C}). *Let $(\lambda\mathbf{C} - \mathbf{F})$ be a regular pencil. Then, equation (5.8) has a unique-symmetric, positive-definite solution \mathbf{P} for every positive definite matrix \mathbf{Q}_A if and only if all eigenvalues of the pencil $(\lambda\mathbf{C} - \mathbf{F})$ are finite and lie in the open left half-plane .*

Case-II. Singular Systems: Consider the GLE (5.7), where the matrix $\mathbf{F} = -\mathbf{G}$ is invertible and \mathbf{C} is rank-deficient ($\text{rank}(\mathbf{C}) = r < n$) and hence non-invertible. The complex nature of singular systems causes many difficulties in analytical and numerical treatment of GLE compared to the nonsingular systems. In particular, existence and uniqueness of the solution even for the asymptotically stable systems are not always guaranteed. This fact can be illustrated by the following example from [32].

Example 1. The GLE (5.7) with

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{F} = -\mathbf{I}_2, \text{ and } \quad \mathbf{Q}_A = \mathbf{I}_2$$

has no solutions.

A series of results are obtained by constructing the different Lyapunov functions. A summary of the existing approaches which are related to the subject of this theses are organized in the two following categories.

Case-II.a. Specific Right-Hand Side:

Theorem 8 ([35]). *A system (5.6) is regular, impulse free¹ and asymptotically stable if and only if there exists matrix \mathbf{P} satisfying:*

$$\begin{cases} \mathbf{P}^T \mathbf{F} + \mathbf{F}^T \mathbf{P} = -\mathbf{Q}_A & (5.9a) \\ \mathbf{C}^T \mathbf{P} = \mathbf{P}^T \mathbf{C} & (5.9b) \end{cases}$$

for any positive definite \mathbf{Q}_A .

Theorem 9 ([35]). *A system (5.6) is regular, impulse free and asymptotically stable if and only if there exist a unique solution $\mathbf{P} \geq \mathbf{0}$ to the GLE*

$$\mathbf{C}^T \mathbf{P} \mathbf{F} + \mathbf{F}^T \mathbf{P} \mathbf{C} = -\mathbf{C}^T \mathbf{Q}_A \mathbf{C}. \quad (5.10)$$

¹System is referred to as *impulse free* if it exhibits no impulsive transient behavior.

satisfying

$$\text{rank}(\mathbf{C}^T \mathbf{P} \mathbf{C}) = \text{rank}(\mathbf{P}) = r < n. \quad (5.11)$$

Furthermore, the following theorem states the uniqueness of the solution while relaxing the constraint on the rank of the matrices.

Theorem 10 ([36,37]). *Let the system in (5.6) be regular and consider the following GLE:*

$$\mathbf{C}^T \mathbf{P} \mathbf{F} + \mathbf{F}^T \mathbf{P} \mathbf{C} = -\mathbf{C}^T \mathbf{Q}_A \mathbf{C}. \quad (5.12)$$

It can be stated

- (a) *if the system is impulse free and asymptotic stable, then for any arbitrary $\mathbf{Q}_A > \mathbf{0}$ there exists a symmetric matrix $\mathbf{P} > \mathbf{0}$ solution to the GLE (5.10). Also, $\mathbf{C}^T \mathbf{P} \mathbf{C} \geq \mathbf{0}$ is unique for each $\mathbf{Q}_A > \mathbf{0}$.*
- (b) *if there exist matrices $\mathbf{P} \geq \mathbf{0}$ and $\mathbf{Q}_A > \mathbf{0}$ satisfying GLE (5.10), then the system (5.6) is impulse free and stable.*

The presence of \mathbf{C} in the right-hand side guarantees the solvability of the GLE (5.10) under some restrictive assumptions on the circuit such as (5.10) as well as, (a) the pencil $(\lambda \mathbf{C} - \mathbf{F})$ should be regular, (b) its all finite eigenvalues of should lie in the open left half-plane, and (c) its index does not exceed two [32].

Case-II.b. Lyapunov Matrix Inequality:

Theorem 11 ([37]). *The system (5.6) is regular, impulse free and asymptotically stable if and only if a Lyapunov function defined as*

$$\mathcal{V}(\mathbf{Cx}(t)) = \mathbf{x}^T(t)\mathbf{C}^T\mathbf{Px}(t) \quad (5.13)$$

satisfies

$$\frac{d\mathcal{V}(\mathbf{Cx}(t))}{d\mathbf{x}(t)} < 0, \quad \forall \mathbf{x}(t) \neq \mathbf{0} \quad (5.14)$$

where \mathbf{P} satisfies

$$\mathbf{PC} + \mathbf{C}^T\mathbf{P} \leq \mathbf{0}, \text{ and } \text{rank}(\mathbf{C}^T\mathbf{P}) = \text{rank}(\mathbf{C}) \quad (5.15)$$

Theorem 12 ([35]). *A system (5.6) is regular, impulse free and asymptotically stable if and only if there exists matrix \mathbf{P} satisfying:*

$$\begin{cases} \mathbf{P}^T\mathbf{F} + \mathbf{F}^T\mathbf{P} < \mathbf{0} \\ \mathbf{C}^T\mathbf{P} = \mathbf{P}^T\mathbf{C} \leq \mathbf{0} \end{cases} \quad (5.16a)$$

$$(5.16b)$$

The algorithms for solving nonsingular GLE when $\mathbf{C} = \mathbf{I}$ or it is invertible as shown in (5.8) are well established [38–40]. The associated computational costs are $\mathcal{O}(n^3)$ for dense \mathbf{G} matrix.

Furthermore, Theorems 8-12 present the sufficient and necessary conditions for the existence and uniqueness of the solution of the singular GLE for stable descriptor

systems. However, it is important to be noted that, all the existing (conventional) approaches for solving Lyapunov equations based on the above theorems suffers from some restrictive assumptions on the systems which can not be always satisfied for the circuit equations. Therefore, the numerical solutions of the generalized Lyapunov equations with singular \mathbf{C} matrices are known to be prohibitively complicated in most cases. Such equations may not have solutions even if all finite eigenvalues of the pencil $(\lambda\mathbf{C} - \mathbf{F})$ lie in the open left half-plane. Moreover, even if a solution exists, it is, in general, not unique [41].

5.4 Proposed Solution Algorithm

Based on the aforementioned considerations, solution of singular GLE for the very-large scaled circuit simulation problems is not practical in an engineering scene. In this section an efficient and systematic methodology is developed to address the issue. To ensure the efficiency of the proposed method, the cumbersome task of solving a GLE for any particular RHS is avoided by devising a systematic method to compute a symmetric matrix $\mathbf{P} \geq \mathbf{0}$ satisfying the Generalized Lyapunov Inequality (GLI) condition.

To this end, the following theorem is considered to establish stability based on the definition of a constraint *generalized Lyapunov function*.

Definition 7. As it has been previously stated, the system (5.6) asymptotically stable if there exists symmetric, positive-definite matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ satisfying the GLE in (5.7), where \mathbf{Q}_A is symmetric and positive definite. For such a solution matrix \mathbf{P} the matrix $\mathbf{C}^T \mathbf{P} \mathbf{E}$ is symmetric, positive-definite; as the congruent transformations preserve definiteness of the matrix. Then the following quadratic generalized Lyapunov function is defined as

$$\mathcal{L}(\mathbf{x}(t)) = \mathbf{x}^T(t) \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{x}(t) > 0, \quad t \in [0, \infty) \quad (5.17)$$

for all nonzero solutions $\mathbf{x}(t) \neq \mathbf{0}$ of equation (5.6).

As it is stated in the last chapter and is repeated here for the sake of ease of access, we consider the following stability theorem which studies the stability of the system based on the existence and property of the Lyapunov function.

Theorem 13 (Lyapunov Generalized function). *Let $\mathcal{L}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ be a positive definite function, i.e., $\mathcal{L}(\mathbf{0}) = 0$ and $\mathcal{L}(\mathbf{x}) > 0$, for $\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}$, if $\frac{d\mathcal{L}(\mathbf{x})}{dt} < 0$, for $\mathbf{x} \in \mathbb{R}^n$, the system (5.6) is asymptotically stable [30].*

Following the Theorem 13 and using (5.17) for a stable system (5.6), we get the

following.

$$\begin{aligned}\frac{d\mathcal{L}(\mathbf{x}(t))}{dt} &= \frac{d\mathbf{x}^T(t)}{dt} \mathbf{C}^T \mathbf{P} \mathbf{C} \mathbf{x}(t) + \mathbf{x}^T(t) \mathbf{C}^T \mathbf{P} \mathbf{C} \frac{d\mathbf{x}(t)}{dt} \\ &= \left(\mathbf{C} \frac{d\mathbf{x}(t)}{dt} \right)^T \mathbf{P} \mathbf{C} \mathbf{x}(t) + \mathbf{x}^T(t) \mathbf{C}^T \mathbf{P} \left(\mathbf{C} \frac{d\mathbf{x}(t)}{dt} \right) < 0\end{aligned}\quad (5.18)$$

Using (5.6) in the autonomous form ($u(t) = 0$),

$$\begin{aligned}\frac{d\mathcal{L}(\mathbf{x}(t))}{dt} &= (\mathbf{F}\mathbf{x}(t))^T \mathbf{P} \mathbf{C} \mathbf{x}(t) + \mathbf{x}^T(t) \mathbf{C}^T \mathbf{P} (\mathbf{F}\mathbf{x}(t)) \\ &= \mathbf{x}^T(t) [\mathbf{F}^T \mathbf{P} \mathbf{C} + \mathbf{C}^T \mathbf{P} \mathbf{F}] \mathbf{x}(t) < 0\end{aligned}\quad (5.19)$$

The inequality in (5.19) is held for any possible transient response $\mathbf{x}(t)$ only when the kernel matrix is negative-definite. Therefore, the following corollary is immediately established.

Corollary 2. *System (5.6) is asymptotically stable if there exists symmetric, positive-definite matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ satisfying the following generalized Lyapunov Inequality (GLI)*

$$\mathbf{F}^T \mathbf{P} \mathbf{C} + \mathbf{C}^T \mathbf{P} \mathbf{F} < 0 \quad (5.20)$$

Based on the aforementioned conclusion, we show next how to solve this equation efficiently by solving a generalized eigenvalue problem.

5.4.1 Formulation of the Proposed Solution Algorithm

This section presents the formulation of the proposed efficient algorithm for solving GLE. The crux of the methodology can be formalized as presented in the following Theorem 14.

Theorem 14. $\mathbf{P} = \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1}$ is a solution for GLI (5.20), where

$$(\mathbf{F}^{-1} \mathbf{C})^T \mathbf{\Gamma} = \mathbf{\Gamma} \mathbf{\Lambda}, \quad \mathbf{\Gamma} \in \mathbb{C}^{n \times r}, \text{ and } \mathbf{\Lambda} = \text{diag}([\lambda_1, \dots, \lambda_r]) \quad (5.21)$$

column vectors of \mathbf{L} are the r eigenvectors corresponding to the r stable eigenvalues of $(\mathbf{F}^{-1} \mathbf{C})^T$ such that $\lambda_i \in \mathbb{C}$, $\text{Re}(\lambda_i) < 0$.

The following Lemma 3 will be used in the proof of Theorem 14.

Lemma 3 ([42]). *If matrix \mathbf{A} is real, then its eigenvalues and eigenvectors come in complex conjugate pairs.*

Proof: That is, if λ is an eigenvalue of \mathbf{A} with corresponding eigenvector $\boldsymbol{\gamma}$ then, λ^* is an eigenvalue of \mathbf{A} and $\boldsymbol{\gamma}^*$ is the corresponding eigenvector.

First, \mathbf{A} being real means $(\mathbf{A} - \lambda \mathbf{I})^* = \mathbf{A} - \lambda^* \mathbf{I}$. Let $\boldsymbol{\gamma}$ be an eigenvector of \mathbf{A} corresponding to the λ , then $\boldsymbol{\gamma} \neq \mathbf{0}$ and $(\mathbf{A} - \lambda \mathbf{I}) \boldsymbol{\gamma} = \mathbf{0}$. It follows that,

$$((\mathbf{A} - \lambda \mathbf{I}) \boldsymbol{\gamma})^* = \mathbf{0}^* \Rightarrow (\mathbf{A} - \lambda \mathbf{I})^* \boldsymbol{\gamma}^* = (\mathbf{A} - \lambda^* \mathbf{I}) \boldsymbol{\gamma}^* = \mathbf{0}$$

hence (λ^*, γ^*) is also an eigen-pair of \mathbf{A} . ■

Theorem 14 can be established as presented below.

Proof: First, using Theorem 3 it is straightforward to prove that the solution matrix $\mathbf{P} = \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1}$ is a real matrix. This is important for ensuring the real-valued transient response for the reduced model which will be constructed using this solution \mathbf{P} .

Next, by substituting the solution matrix \mathbf{P} in GLI (5.20) we get

$$\begin{aligned}
 \mathbf{F}^T \mathbf{P} \mathbf{C} + \mathbf{C}^T \mathbf{P} \mathbf{F} &= \mathbf{F}^T \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1} \mathbf{C} + \mathbf{C}^T \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1} \mathbf{F} \\
 &= \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1} \mathbf{C} + \mathbf{C}^T \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \\
 &= \mathbf{\Gamma} \left((\mathbf{F}^{-1} \mathbf{C})^{*T} \mathbf{\Gamma} \right)^{*T} + \left((\mathbf{F}^{-1} \mathbf{C})^T \mathbf{\Gamma} \right) \mathbf{\Gamma}^{*T} \quad (5.22)
 \end{aligned}$$

Considering that for a real matrix i.e. $(\mathbf{F}^{-1} \mathbf{C})^{*T} = (\mathbf{F}^{-1} \mathbf{C})^T$ and by substituting (5.21) in (5.22), we have

$$\begin{aligned}
 \mathbf{F}^T \mathbf{P} \mathbf{C} + \mathbf{C}^T \mathbf{P} \mathbf{F} &= \mathbf{\Gamma} (\mathbf{\Gamma} \mathbf{\Lambda})^{*T} + (\mathbf{\Gamma} \mathbf{\Lambda}) \mathbf{\Gamma}^{*T} \\
 &= \mathbf{\Gamma} (\mathbf{\Lambda}^{*T} + \mathbf{\Lambda}) \mathbf{\Gamma}^{*T} = \mathbf{\Gamma} \begin{bmatrix} 2\mathcal{R}e(\lambda_1) & & \\ & \ddots & \\ & & 2\mathcal{R}e(\lambda_r) \end{bmatrix} \mathbf{\Gamma}^{*T} \quad (5.23)
 \end{aligned}$$

Now, $\forall \mathbf{Y} \in \mathbb{C}^{n \times 1}$, it is

$$\mathbf{Y}^{*T} \mathbf{\Gamma} \begin{bmatrix} 2\mathcal{R}e(\lambda_1) & & \\ & \ddots & \\ & & 2\mathcal{R}e(\lambda_r) \end{bmatrix} \mathbf{\Gamma}^{*T} \mathbf{Y} = \tilde{\mathbf{Y}}^{*T} \begin{bmatrix} 2\mathcal{R}e(\lambda_1) & & \\ & \ddots & \\ & & 2\mathcal{R}e(\lambda_r) \end{bmatrix} \tilde{\mathbf{Y}} \quad (5.24)$$

where $\tilde{\mathbf{Y}} = [\tilde{y}_1, \dots, \tilde{y}_r]^T \triangleq \mathbf{\Gamma}^{*T} \mathbf{Y} \in \mathbb{C}^{n \times 1}$, and $\tilde{y}_i \in \mathbb{C}$ for $i = 1, \dots, r$.

Then (5.24) can be equivalently rewritten as

$$\begin{aligned} \tilde{\mathbf{Y}}^{*T} \begin{bmatrix} 2\mathcal{R}e(\lambda_1) & & \\ & \ddots & \\ & & 2\mathcal{R}e(\lambda_r) \end{bmatrix} \tilde{\mathbf{Y}} &= [\tilde{y}_1^*, \dots, \tilde{y}_r^*] \begin{bmatrix} 2\mathcal{R}e(\lambda_1) & & \\ & \ddots & \\ & & 2\mathcal{R}e(\lambda_r) \end{bmatrix} \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_r \end{bmatrix} \\ &= [\tilde{y}_1^*, \dots, \tilde{y}_r^*] \begin{bmatrix} 2\mathcal{R}e(\lambda_1)\tilde{y}_1 \\ \vdots \\ 2\mathcal{R}e(\lambda_r)\tilde{y}_r \end{bmatrix} = 2\mathcal{R}e(\lambda_1)\tilde{y}_1^* \tilde{y}_1 + \dots + 2\mathcal{R}e(\lambda_r)\tilde{y}_r^* \tilde{y}_r \\ &= 2\mathcal{R}e(\lambda_1)|\tilde{y}_1|^2 + \dots + 2\mathcal{R}e(\lambda_r)|\tilde{y}_r|^2 < 0 \quad (5.25) \end{aligned}$$

The inequality in (5.25) is valid since all λ_i are stable eigenvalues of, where $\mathcal{R}e(\lambda_i) < 0$. Having 5.25, the matrix in (5.23) is proved to be symmetric, negative-

definite. Therefore, from (5.22) and (5.25), it is concluded,

$$\mathbf{F}^T \mathbf{P} \mathbf{C} + \mathbf{C}^T \mathbf{P} \mathbf{F} < \mathbf{0} \quad (5.26)$$

■

Of course, solving an eigenvalue problem to find the eigen-pairs of the matrix $(\mathbf{F}^{-1} \mathbf{C})^T$ is not always practical due to the very large matrices in circuit simulation applications. As an efficient alternative is that, eigenvalue problem of $(\mathbf{F}^{-1} \mathbf{C})^T$ can be converted to a general eigenvalue problem of a pencil, which can be solved very efficiently.

The generalized eigenvalue problem can be compactly defined as follows.

Generalized eigenvalue problem [43]: Let \mathcal{C} and \mathcal{F} two n-by-n real matrices; the eigenvalues of the pencil $(\lambda \mathcal{C} - \mathcal{F})$ are defined by $\Lambda = \{\lambda_i \in \mathbb{C} \mid \det(\lambda_i \mathcal{C} - \mathcal{F}) = 0\}$ and

$$\mathcal{F} \gamma_i = \lambda_i \mathcal{C} \gamma_i, \quad \gamma_i \neq \mathbf{0} \quad (5.27)$$

then γ_i is referred to as an eigenvector of pencil. For the diagonal eigenvalue matrix Λ and eigenvector matrix Γ , it is

$$\mathcal{F} \Gamma = \mathcal{C} \Gamma \Lambda \quad (5.28)$$

Based on the matrices in (5.21), let a new matrix be defined as $\tilde{\mathbf{\Gamma}} \triangleq \mathbf{F}^{-T}\mathbf{\Gamma}$ and by substituting it in (5.21), we get

$$\begin{aligned} (\mathbf{F}^{-1}\mathbf{C})^T \mathbf{\Gamma} &= \mathbf{\Gamma} \mathbf{\Lambda} \\ \mathbf{C}^T \mathbf{F}^{-T} \mathbf{\Gamma} &= \mathbf{\Gamma} \mathbf{\Lambda} \\ \mathbf{C}^T \tilde{\mathbf{\Gamma}} &= \mathbf{F}^T \tilde{\mathbf{\Gamma}} \mathbf{\Lambda} \end{aligned} \tag{5.29}$$

As shown in (5.29), $\tilde{\mathbf{\Gamma}}$ is calculated by general eigen-decomposition of \mathbf{C}^T with \mathbf{F}^T . Then, $\mathbf{\Gamma}$ can be obtained as

$$\mathbf{\Gamma} = \mathbf{F}^T \tilde{\mathbf{\Gamma}} \tag{5.30}$$

Through this approach, finding $\mathbf{\Gamma}$ does not require matrix inversion (applying LU decomposition) and all operands are sparse when matrix \mathbf{C} and \mathbf{F} are sparse.

5.4.2 Construction of the Left Projection Matrix

As presented in the section 4.1.2 of previous chapter, having a right projection matrices $\mathbf{V} \in \mathbb{R}^{n \times q}$ which is obtained from Arnoldi algorithm, the left projection matrices is computed as shown in and repeated here,

$$\mathbf{U} = \mathbf{P}^T \mathbf{C} \mathbf{V} \in \mathbb{R}^{n \times m}, \tag{5.31}$$

Having the solution of the Lyapunov equation $\mathbf{P} = \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^* \mathbf{F}^{-1}$ in Theorem 14 from (5.31) we get

$$\mathbf{U} = \mathbf{F}^{-T} \mathbf{\Gamma}^* \mathbf{\Gamma}^T \mathbf{F}^{-1} \mathbf{C} \mathbf{V} \in \mathbb{R}^{n \times m}, \quad (5.32)$$

By substituting (5.30) in (5.32) we get

$$\begin{aligned} \mathbf{U} &= \mathbf{F}^{-T} (\mathbf{F}^T \tilde{\mathbf{\Gamma}})^* (\mathbf{F}^T \tilde{\mathbf{\Gamma}})^T \mathbf{F}^{-1} \mathbf{C} \mathbf{V} \\ &= \mathbf{F}^{-T} \mathbf{F}^T \tilde{\mathbf{\Gamma}}^* \tilde{\mathbf{\Gamma}}^T \mathbf{F} \mathbf{F}^{-1} \mathbf{C} \mathbf{V} \\ &= \tilde{\mathbf{\Gamma}}^* \tilde{\mathbf{\Gamma}}^T \mathbf{C} \mathbf{V} \end{aligned} \quad (5.33)$$

The steps of the above approach are summarized in Algorithm 2, presented using Matlab-like pseudo-code.

Algorithm 5.1: The proposed method based on the efficient GLE solver.

Data: Original system's matrices $\mathbf{G}, \mathbf{C}, \mathbf{B}, \mathbf{L}$ and reduced system size q , and an integer p greater than q

Result: Reduced system's matrices $\hat{\mathbf{G}}, \hat{\mathbf{C}}, \hat{\mathbf{B}}, \hat{\mathbf{L}}$

Find left projection \mathbf{V} with Arnoldi method;

$[\tilde{\mathbf{\Gamma}}, \mathbf{\Lambda}] = \text{eigs}(\mathbf{C}^T, -\mathbf{G}^T, r);$

$\mathbf{\Lambda} = \text{diag}(\mathbf{\Lambda});$

// select eigenvectors corresponding to the stable eigenvalues:

$i = \text{real}(\mathbf{\Lambda}) < 0;$

$\mathbf{\Lambda} = \mathbf{\Lambda}(i);$

$\tilde{\mathbf{\Gamma}} = \tilde{\mathbf{\Gamma}}(:, i);$

// select largest $m \leq r$ real-part eigenvalues

$[\sim, i] = \text{sort}(\text{real}(\mathbf{\Lambda}));$

$i = i(1:q);$

$\mathbf{\Lambda} = \mathbf{\Lambda}(i);$

$\tilde{\mathbf{\Gamma}} = \tilde{\mathbf{\Gamma}}(:, i);$

$\mathbf{U} = \tilde{\mathbf{\Gamma}}^* \tilde{\mathbf{\Gamma}}^T \mathbf{C} \mathbf{V};$

$\hat{\mathbf{C}} = \mathbf{U}^T \mathbf{C} \mathbf{V};$

$\hat{\mathbf{G}} = \mathbf{U}^T \mathbf{G} \mathbf{V};$

$\hat{\mathbf{B}} = \mathbf{U}^T \mathbf{B};$

$\hat{\mathbf{L}} = \mathbf{L} \mathbf{V};$

5.4.3 Modified Left Projection Matrix

We can further simplify the left projection matrix based on the efficient solution of GLE and through exploiting the following lemma.

Lemma 4. *Let matrices $\mathbf{U} \in \mathbb{R}^{n \times m}$ and $\bar{\mathbf{U}} \in \mathbb{R}^{n \times m}$ having the same column space be given. Using \mathbf{U} and $\bar{\mathbf{U}}$ as left projection matrices lead to the same ROM.*

Proof. Since \mathbf{U} and $\bar{\mathbf{U}}$ have the same column space, we can write

$$\bar{\mathbf{U}} = \mathbf{U} \mathbf{U}_s \tag{5.34}$$

where $\mathbf{U}_s \in \mathbb{R}^{m \times m}$ is an invertible square matrix.

The ROMs generated by \mathbf{U} or $\bar{\mathbf{U}}$ with the same right projection matrix \mathbf{V} are given by

$$\mathbf{U}^T \mathbf{C} \mathbf{V} \dot{\mathbf{x}} + \mathbf{U}^T \mathbf{G} \mathbf{V} \mathbf{x} = \mathbf{U}^T \mathbf{B} \mathbf{u} \quad (5.35)$$

and

$$\bar{\mathbf{U}}^T \mathbf{C} \mathbf{V} \dot{\mathbf{x}} + \bar{\mathbf{U}}^T \mathbf{G} \mathbf{V} \mathbf{x} = \bar{\mathbf{U}}^T \mathbf{B} \mathbf{u} \quad (5.36)$$

By substituting (5.32) in (5.34), we get

$$\mathbf{U}_s^T \mathbf{U}^T \mathbf{C} \mathbf{V} \dot{\mathbf{x}} + \mathbf{U}_s^T \mathbf{U}^T \mathbf{G} \mathbf{V} \mathbf{x} = \mathbf{U}_s^T \mathbf{U}^T \mathbf{B} \mathbf{u} \quad (5.37)$$

Since \mathbf{U}_s is an invertible matrix, by multiplying both sides of (5.37) by \mathbf{U}_s^{-T} , it is seen that both ROMs (5.35) and (5.37) are the same. \square

From (5.31), we have the left projection matrix $\mathbf{U} = \mathbf{P} \mathbf{C} \mathbf{V}$, where $\mathbf{P} = \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1}$ is a dense matrix. Dense matrix multiplication cost the computation time of $\mathcal{O}(n^3)$, but this cost can be avoided by using an equivalent left projection matrix.

The column space of \mathbf{U} is the same column space of \mathbf{P} which is given by according to Theorem 14

$$\mathbf{P} = \mathbf{F}^{-T} \mathbf{\Gamma} \mathbf{\Gamma}^{*T} \mathbf{F}^{-1} = \tilde{\mathbf{\Gamma}} \tilde{\mathbf{\Gamma}}^{*T} \quad (5.38)$$

It is to be noted that, \mathbf{P} has the same column space as matrix $\tilde{\mathbf{\Gamma}}$. However, since

$\tilde{\Gamma}$ is the general eigenvectors of some real matrices, it is a complex matrix. If we use $\tilde{\Gamma}$ directly as the left projection matrix, the resulted ROM would contain complex parameters. However, as the eigenvectors of real matrices always comes in conjugate pair, $\tilde{\Gamma}\tilde{\Gamma}^{*T}$ can be written as

$$\begin{aligned}\mathbf{P} = \tilde{\Gamma}\tilde{\Gamma}^{*T} &= (\mathcal{R}e(\tilde{\Gamma}) + j\mathcal{I}m(\tilde{\Gamma}))(\mathcal{R}e(\tilde{\Gamma}^T) - j\mathcal{I}m(\tilde{\Gamma}^T)) \\ &= (\mathcal{R}e(\tilde{\Gamma})\mathcal{R}e(\tilde{\Gamma}^T) + \mathcal{I}m(\tilde{\Gamma})\mathcal{I}m(\tilde{\Gamma}^T)) + j(\mathcal{I}m(\tilde{\Gamma})\mathcal{R}e(\tilde{\Gamma}^T) - \mathcal{R}e(\tilde{\Gamma})\mathcal{I}m(\tilde{\Gamma}^T))\end{aligned}\tag{5.39}$$

As proved in Theorem 14, $\mathbf{P} \in \mathbb{R}^{n \times n}$ is a real matrix and hence, the imaginary part in any representation of \mathbf{P} such as (5.39) to be zero, as

$$(\mathcal{I}m(\tilde{\Gamma}))\mathcal{R}e(\tilde{\Gamma}^T) - \mathcal{R}e(\tilde{\Gamma})\mathcal{I}m(\tilde{\Gamma}^T) = 0\tag{5.40}$$

Therefore, we have

$$\mathbf{P} = (\mathcal{R}e(\tilde{\Gamma})\mathcal{R}e(\tilde{\Gamma}^T) + \mathcal{I}m(\tilde{\Gamma})\mathcal{I}m(\tilde{\Gamma}^T))\tag{5.41}$$

It is important to be noted that in (5.41), it is

$$\text{Span} \left\{ \mathcal{R}e(\tilde{\mathbf{\Gamma}}) \mathcal{R}e(\tilde{\mathbf{\Gamma}}^T) \right\} \subseteq \text{colspan} \left\{ \mathcal{R}e(\tilde{\mathbf{\Gamma}}) \right\} \quad (5.42)$$

$$\text{Span} \left\{ \mathcal{I}m(\tilde{\mathbf{\Gamma}}) \mathcal{I}m(\tilde{\mathbf{\Gamma}}^T) \right\} \subseteq \text{colspan} \left\{ \mathcal{I}m(\tilde{\mathbf{\Gamma}}) \right\} \quad (5.43)$$

From (5.42) and (5.43), it is concluded that

$$\text{colspan} \left\{ \mathbf{P} \right\} \subseteq \text{colspan} \left\{ \mathcal{R}e(\tilde{\mathbf{\Gamma}}) + \mathcal{I}m(\tilde{\mathbf{\Gamma}}) \right\} \quad (5.44)$$

The interpenetration of (5.44) is that \mathbf{P} has the same column space as real matrix $(\mathcal{R}e(\tilde{\mathbf{\Gamma}}) + \mathcal{I}m(\tilde{\mathbf{\Gamma}}))$, which means that also $(\mathcal{R}e(\tilde{\mathbf{\Gamma}}) + \mathcal{I}m(\tilde{\mathbf{\Gamma}}))$ has the same column space as \mathbf{PCV} . Applying Lemma 4, the left projection matrix \mathbf{PCV} and $(\mathcal{R}e(\tilde{\mathbf{\Gamma}}) + \mathcal{I}m(\tilde{\mathbf{\Gamma}}))$ generates the same ROM. Thus, we choose the left projection matrix to be

$$\mathbf{U} = \mathcal{R}e(\tilde{\mathbf{\Gamma}}) + \mathcal{I}m(\tilde{\mathbf{\Gamma}}) \quad (5.45)$$

The steps of the improved proposed method are summarized in Algorithm 3, presented using Matlab-like pseudo-code.

Algorithm 5.2: The proposed method with modified left projection matrix.

Data: Original system's matrices $\mathbf{G}, \mathbf{C}, \mathbf{B}, \mathbf{L}$ and reduced system size q , and an integer p greater than q

Result: Reduced system's matrices $\hat{\mathbf{G}}, \hat{\mathbf{C}}, \hat{\mathbf{B}}, \hat{\mathbf{L}}$

Find left projection \mathbf{V} with Arnoldi method;

$[\tilde{\mathbf{\Gamma}}, \mathbf{\Lambda}] = \text{eigs}(\mathbf{C}^T, -\mathbf{G}^T, p);$

$\mathbf{\Lambda} = \text{diag}(\mathbf{\Lambda});$

// select eigenvectors corresponding to the negative eigenvalues

$i = \text{Re}(\mathbf{\Lambda}) < 0;$

$\mathbf{\Lambda} = \mathbf{\Lambda}(i);$

$\tilde{\mathbf{\Gamma}} = \tilde{\mathbf{\Gamma}}(:, i);$

// select largest m real-part eigenvalues

$[\sim, i] = \text{sort}(\text{Re}(\mathbf{\Lambda}));$

$i = i(1:m);$

$\mathbf{\Lambda} = \mathbf{\Lambda}(i);$

$\tilde{\mathbf{\Gamma}} = \tilde{\mathbf{\Gamma}}(:, i);$

$\mathbf{U} = \text{Re}(\tilde{\mathbf{\Gamma}}) + \text{Im}(\tilde{\mathbf{\Gamma}});$

$\hat{\mathbf{C}} = \mathbf{U}^T \mathbf{C} \mathbf{V};$

$\hat{\mathbf{G}} = \mathbf{U}^T \mathbf{G} \mathbf{V};$

$\hat{\mathbf{B}} = \mathbf{U}^T \mathbf{B};$

$\hat{\mathbf{L}} = \mathbf{L} \mathbf{V};$

5.5 Computation cost analysis

The cost of the proposed model order reduction method mainly comes from computing the left and right projection matrices. The right projection matrix is obtained by block Arnoldi process which takes the computation cost $\mathcal{O}(mn)$ considering that matrix \mathbf{G} and \mathbf{C} are sparse, where m is the order of the reduced system and n denotes the order of the original system [43].

In the proposed method, the left projection matrix is given by $(\text{Re}(\tilde{\mathbf{\Gamma}}) + \text{Im}(\tilde{\mathbf{\Gamma}}))$. The dominant cost for computing the left projection matrix is associated with find-

ing the eigenvectors corresponding to r number of the stable eigenvalues through a generalized eigenvalue analysis.

Recently various methods for solving generalized eigenvalue problems have been developed. It has been established that, the r eigen-pairs of a generalized eigenvalue problem can be found in r steps of an Arnoldi iteration. With state-of-the-art approaches, the computational cost of each Arnoldi iteration step for solving a generalized eigenvalue problem has also been reduced to $\mathcal{O}(n)$ [44–46]. Therefore, the cost for computing all r real and complex-conjugate eigen-pairs is upper-bounded to $\mathcal{O}(rn)$. Considering that $m \leq r$, the CPU cost for the proposed MOR method including forming both right and left projection matrices can be kept as low as $\mathcal{O}(mn)$.

Comparing the cost of the proposed method with the CPU-cost required for a direct solution of Lyapunov equation which is $\mathcal{O}(n^2)$ -to- $\mathcal{O}(n^3)$ [47] proves the superiority of the proposed method in terms of the computational complexity.

5.5.1 Computational Results

In this section, we will present two examples. The first example is an illustrative example which serves the purpose of illustrating the core ideas of the proposed method and to study the properties of the proposed method. The second example is a practical large active circuit to demonstrate efficiency and accuracy of the proposed method.

Example 1:

This is a random sparse system. All computation is done in MATLAB. Matrix \mathbf{G}, \mathbf{C} are generated randomly with function ‘sprand(2000,2000,0.05)’ , which specifies that they have the size of 2000×2000 and the sparsity is 95%. $\mathbf{L}^T = \mathbf{B}$ where \mathbf{B} is a 2000-by-1 matrix whose first element equals to 1. The original system is reduced to a ROM with the size of 15 using the proposed method with modified left projection matrix (Algorithm 3).

Figure (5.1) shows the poles of the original system in complex domain. The poles scatters evenly across the LHS and RHS complex domain, which indicates that the original system is highly unstable. Figure (5.2) shows the poles of the reduced system in complex domain. The poles only scatters at the LHS of the complex domain, which means the reduced system is stable.

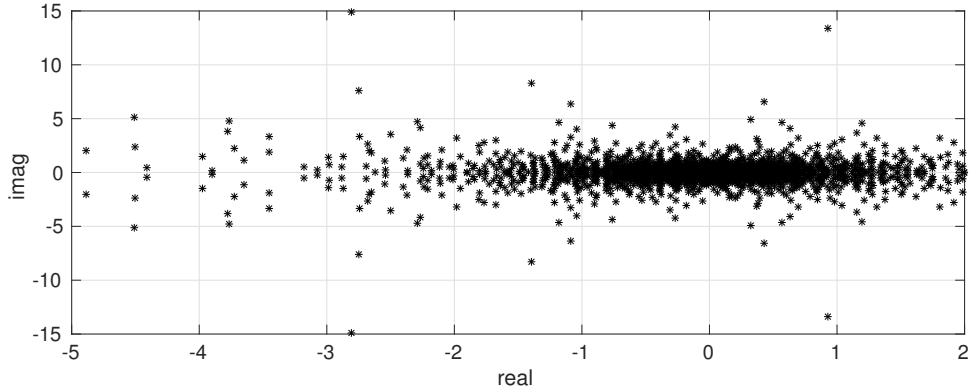


Figure 5.1: Poles distribution of the original sparse system with 2000 state variables.

Figure 5.3 shows the first 15 (\mathcal{M}_{0-14}) moments of the original and reduced system

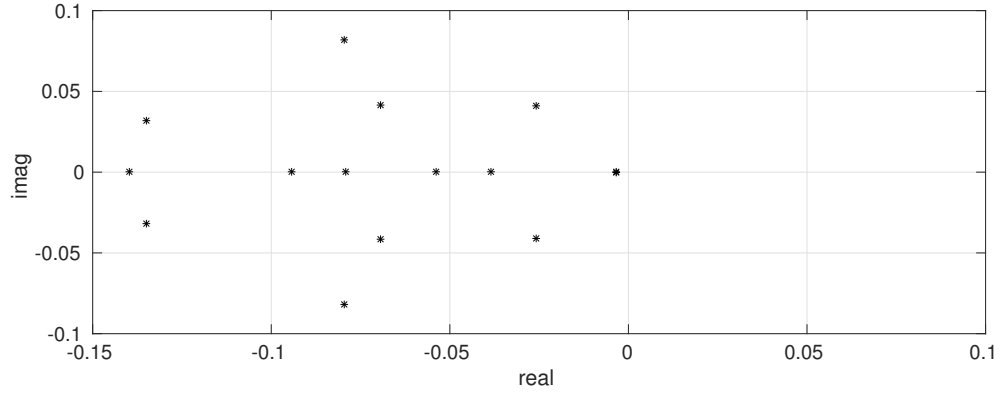


Figure 5.2: Poles distribution of the reduced system from the random generated 2000 by 2000 sparse system.

with y-axis in log scale. We use the absolute value of \mathcal{M}_i as our magnitude, where the moments of the original system is marked as 'o' and the moments of the reduced system is marked as 'x'. The moments of the reduced system matches with the moments of the original system.

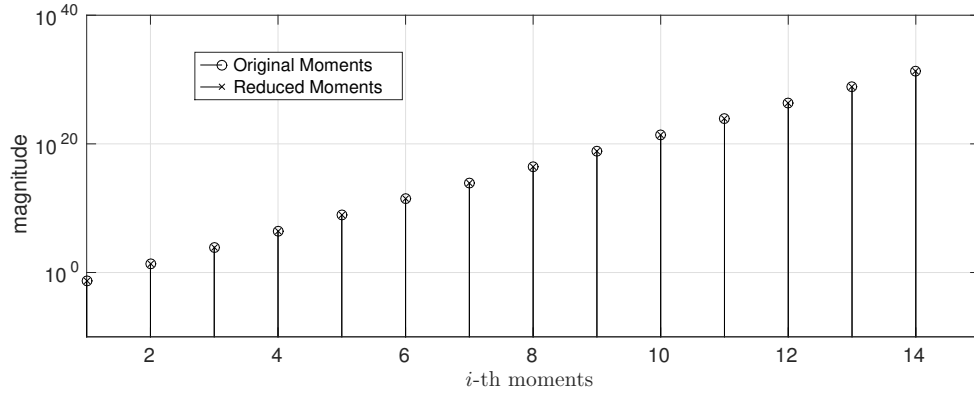


Figure 5.3: First 15 Moments Comparison between the original system and ROM.

The computation time for calculating \mathbf{U} and \mathbf{V} for each reduced model are compared in Table 5.1. The ROM is achieved by applying Algorithm 3 to the sparse

system matrices.

Table 5.1 shows that the computation time for calculating the left projection matrix using the proposed method is less than calculating the right projection matrix. This is achieved mainly by the two improvements explained in this chapter. First, the generalized eigenvalue analysis is used to obtain $\tilde{\mathbf{T}}$. Second, the left projection matrix is computed as $\mathbf{U} = \mathcal{R}e(\tilde{\mathbf{T}}) + \mathcal{I}m(\tilde{\mathbf{T}})$.

Table 5.1: Computation time for calculating left and right projection matrix.

Original Order	Reduced Order	U (Second)	V (Second)
2000	5	1.5517	2.9701
	10	1.7420	5.4356
	15	1.7257	8.1952

Example 2:

We consider an amplifier circuit shown in Figure 5.4. The interconnect structure consists of four coupled lines of length $L=2\text{cm}$.

The RLGC parameters of the lines whose geometry is shown in Figure 5.5, which are calculated using the field solver in HSPICE[®]. The transmission line is represented with 1000 segments of T-equivalent lumped representation [31]. The amplifier equivalent circuit is shown in Figure 5.6.

Applying the proposed method, a stable reduced-order model for the original (stable) network in Figure 5.4 is obtained. The accuracy of the resulting reduced model is demonstrated in Figures 5.7-5.8, where the output voltage responses of the

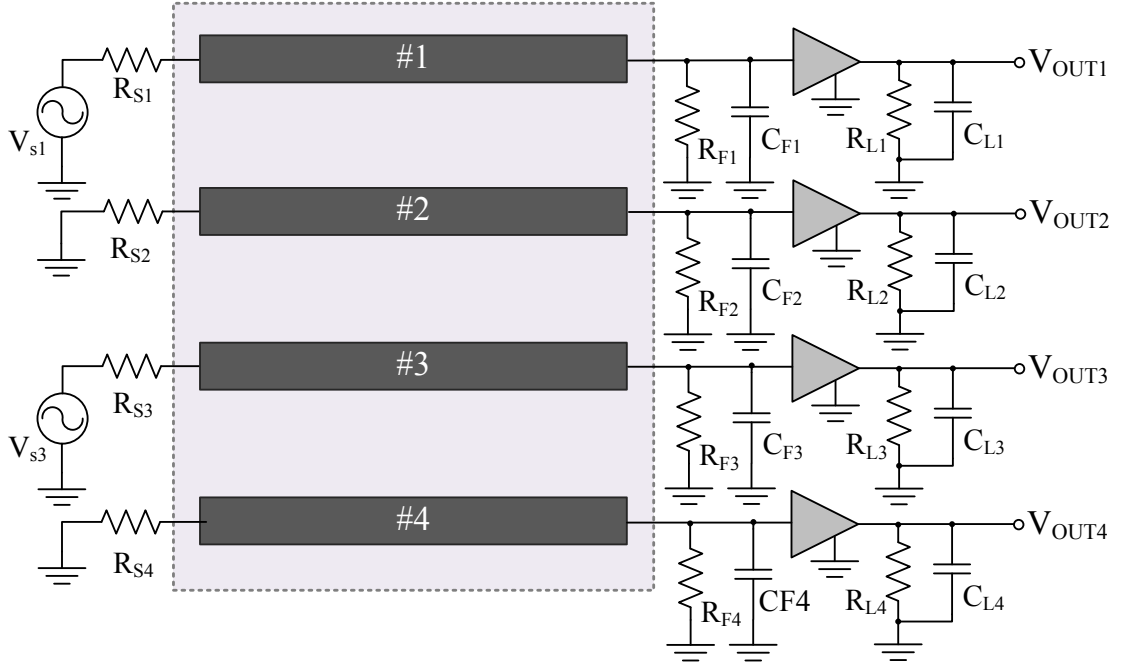


Figure 5.4: A stable active design consisting of four coupled transmission lines and amplifier blocks.

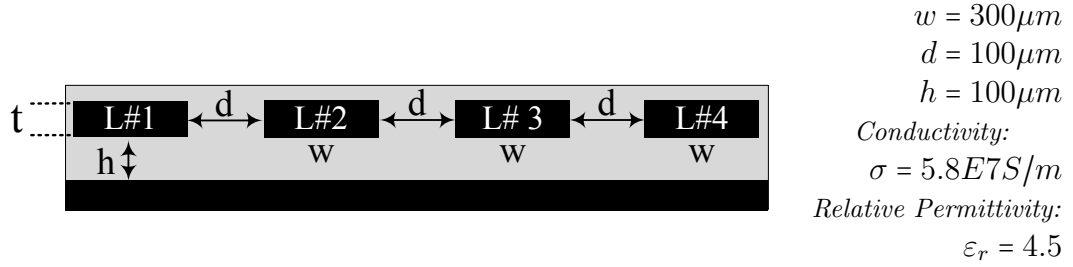


Figure 5.5: Cross sectional geometry of TL structure.

reduced model is compared with the simulation results of the original (unreduced) model and the PRIMA-based reduced model. Both reduced models are of the same order.

Frequency-domain response of the original circuit and the proposed reduced model are in excellent agreement within the intended frequency range compared to the

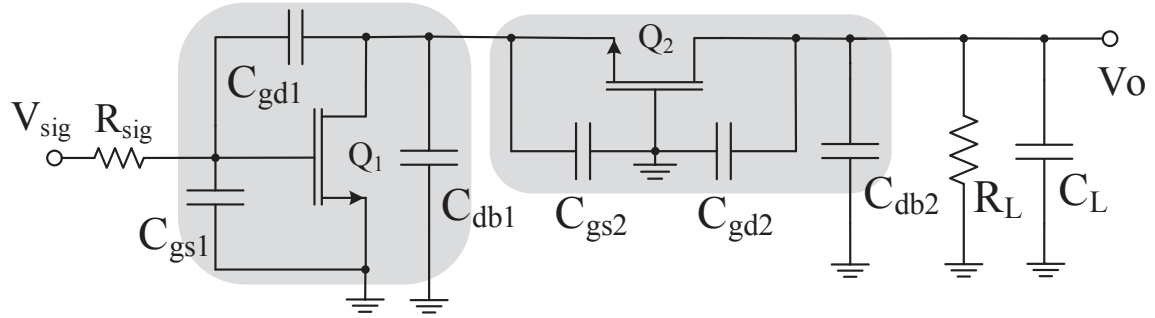


Figure 5.6: High-frequency equivalent-circuit model for MOS cascade amplifier (the biasing network is not presented).

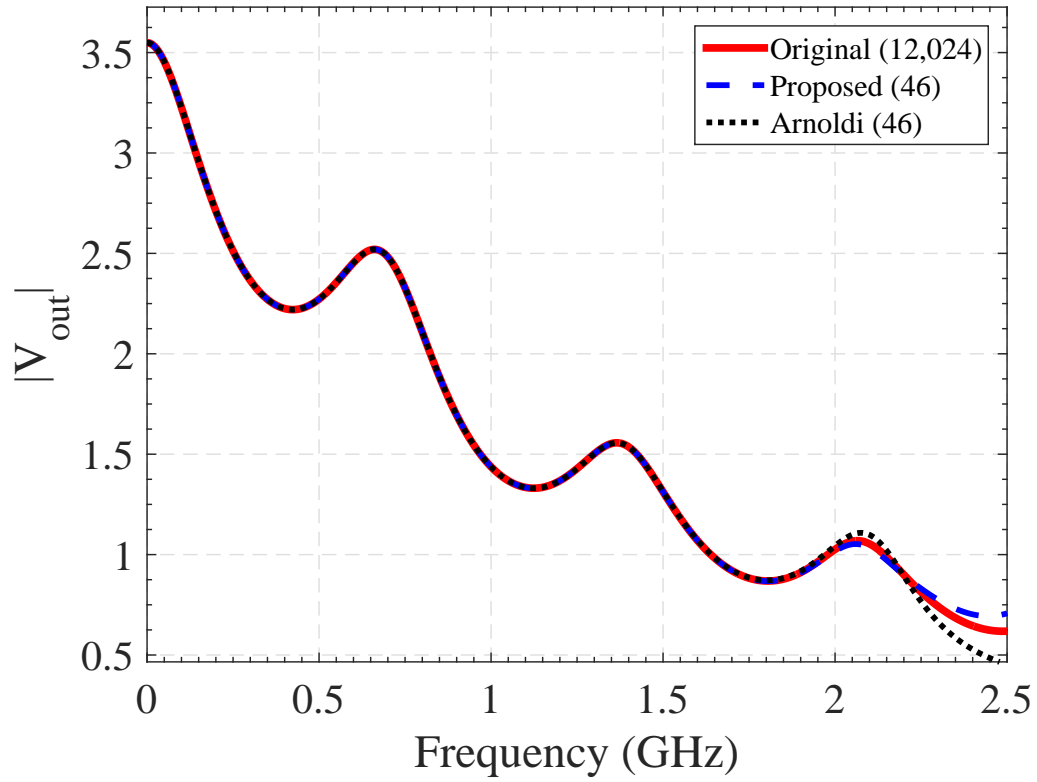


Figure 5.7: Comparison of the frequency responses at V_{out3} (at Far-end of line-3) from the proposed method, the PRIMA-based model and the original circuit.

PRIMA-based model.

The accuracy of the resulting reduced model is further demonstrated in Fig-

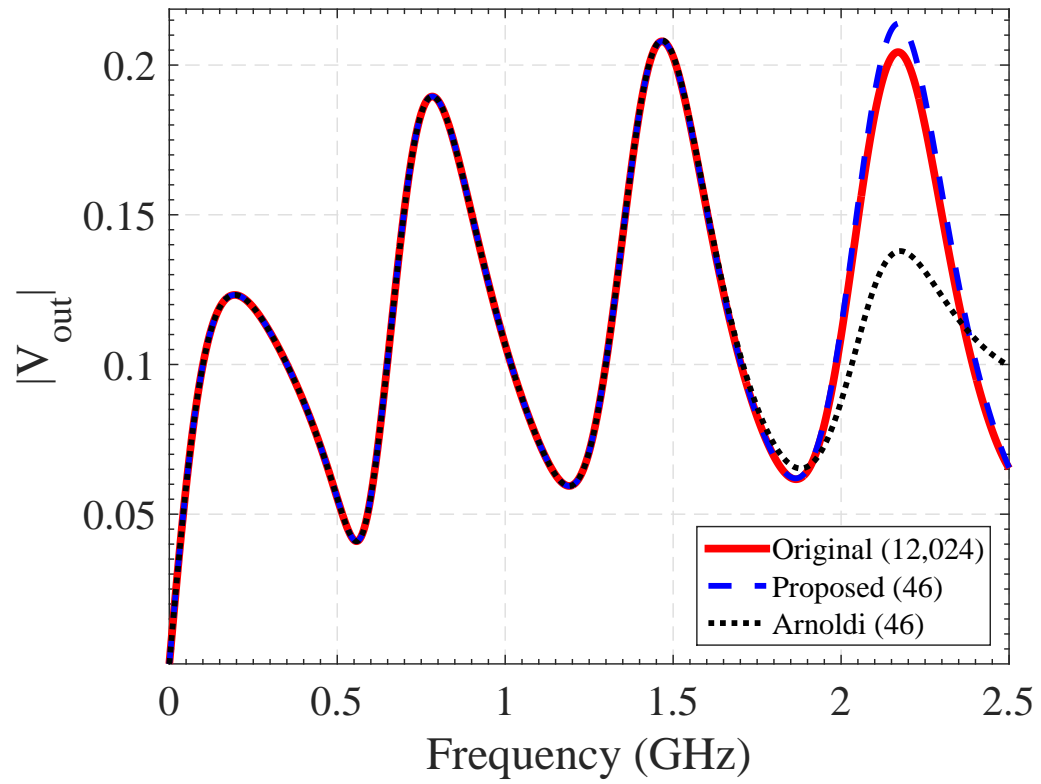


Figure 5.8: Comparison of the frequency responses at V_{out4} (at Far-end of line-4) from the proposed method, the PRIMA-based model and the original circuit.

ures 5.9-5.10, where the transient voltage responses of the reduced model is compared with the simulation results of the original (unreduced) model. The input signal is trapezoidal-pulse input signals with the delay time $t_d = 1ns$, pulse-width $t_{pw} = 5ns$, and the rise and fall times $t_r = 0.25ns$ and $t_f = 0.25ns$, respectively.

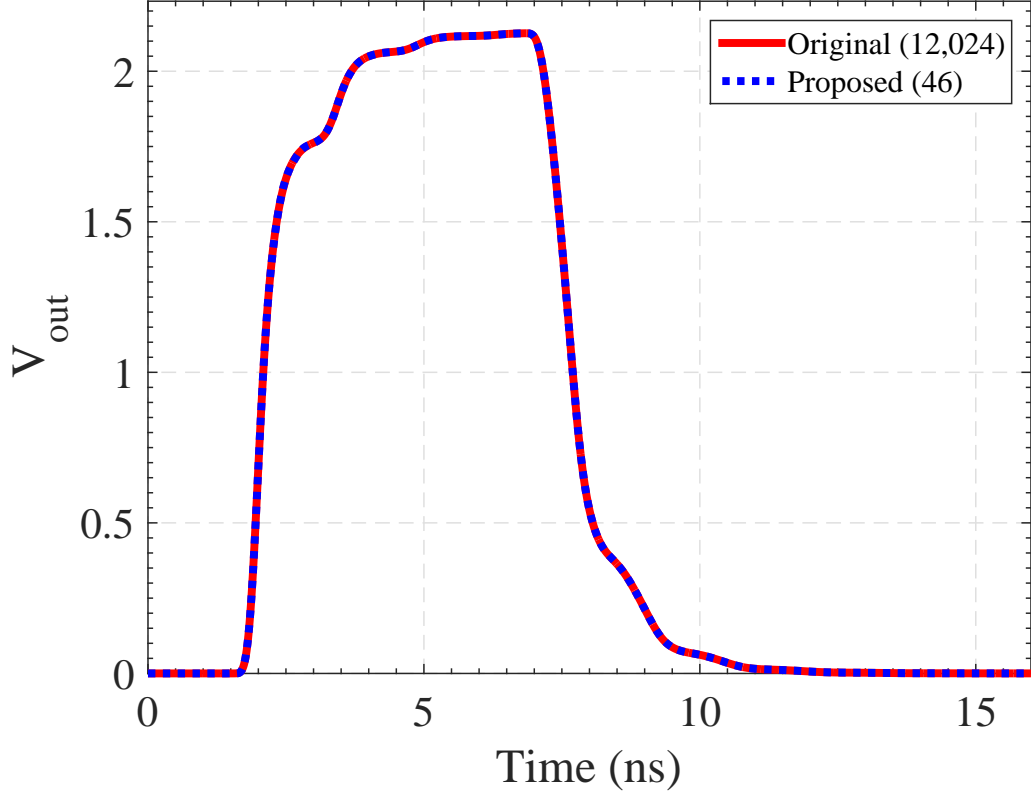


Figure 5.9: Comparison of the transient responses at V_{out3} (at Far-end of line-3) from the proposed method, the PRIMA-based model and the original circuit.

The proposed reduced model provided an accurate time-domain response, whereas the time-domain simulation of the PRIMA model failed to converge because of the unstable poles.

The sizes and stability properties of the original and reduced models are compared

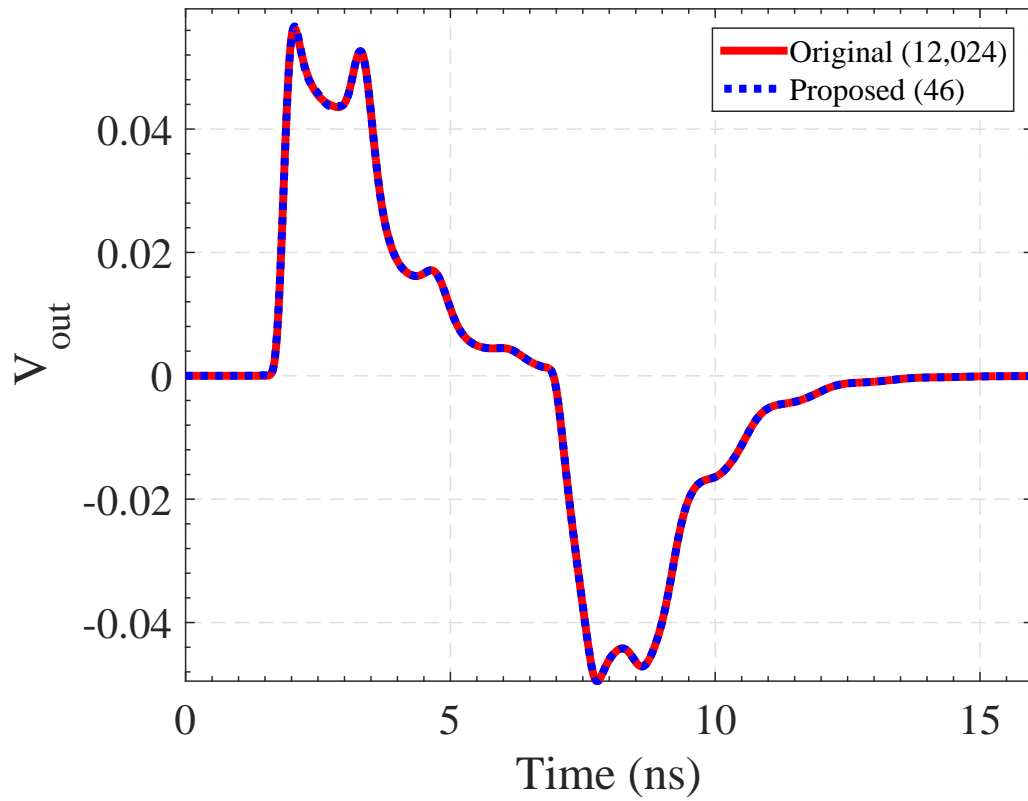


Figure 5.10: Comparison of the transient responses at V_{out4} (at Far-end of line-4) from the proposed method, the PRIMA-based model and the original circuit.

in Table 5.2.

Table 5.2: Comparison of the original and reduced models.

	Dimension	Stability
Original Circuit	12,024	Yes
PRIMA-based Reduced Model	46	No
Proposed Reduced Model	46	Yes

The CPU-time for frequency domain simulation of the original is compared with the proposed reduced model in Table 5.3.

Table 5.3: Comparison of the CPU-time for frequency simulation using original and reduced models.

	Original	Reduced
Order	12,024	46
Simulation Time	33.306	0.195
Speed-up		≈ 171

Chapter 6

Summary and future work

6.1 Summary

A good model reduction method should satisfy the following three conditions. First, the approximated outputs should be accurate. Second, the ROM preserves important properties of the original system such as stability. Third, the reduced method is efficient for computation. There exist many state-of-art methods that can maintain good properties of the original system and the approximation response is very accurate, such as balanced truncating and Hankel norm approximation. However, these methods are not efficient in computation, which limits their application to small systems. There also exist methods that are computationally efficient, and ROM can match an arbitrary number of moments of the original system. But the problem for this kind

of method is that it cannot preserve the stability of the original system.

In this thesis, a new MOR method is proposed as an attempt to find a ideal model reduction method that satisfies all three conditions. The main idea of the solution is to do Petrov-Galerkin projection instead of Galerkin projection, by generating the left projection matrix from the solution of Lyapunov equation. It is proved that the proposed method can preserve both the moments and the stability of the circuit.

6.2 Future Work

- Combining the proposed method with clustering technique [48], to address the difficulty associated with the reduction of networks with a large number of input/output terminals, while guarantee the stability for the ROM.
- Extension of the proposed method to guarantee the stability of parameterized MOR of active circuits.

References

- [1] W. H. Schilders, H. A. Van der Vorst, and J. Rommes, *Model order reduction: theory, research aspects and applications*. Springer, 2008, vol. 13.
- [2] E. J. Grimme, “Krylov projection methods for model reduction,” Ph.D. dissertation, Citeseer, 1997.
- [3] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2005.
- [4] S. X.-D. Tan and L. He, *Advanced Model Order Reduction Techniques in VLSI Design*. Cambridge, MA: Cambridge University Press, 2007.
- [5] A. Odabasioglu, M. Celik, and L. T. Pileggi, “PRIMA: Passive reduced-order interconnect macromodeling algorithm,” in *Proceedings of the 1997 IEEE/ACM International Conference on Computer-aided Design*, ser. ICCAD ’97. Washington, DC, USA: IEEE Computer Society, 1997, pp. 58–65. [Online]. Available: <http://dl.acm.org/citation.cfm?id=266388.266423>

- [6] R. C. Selga, B. Lohmann, and R. Eid, “Stability preservation in projection-based model order reduction of large scale systems,” *European Journal of Control*, vol. 18, no. 2, pp. 122–132, 2012.
- [7] B. N. Bond and L. Daniel, “Guaranteed stable projection-based model reduction for indefinite and unstable linear systems,” in *Proc. IEEE/ACM International Conference on Computer-Aided Design*, San Jose, CA, Nov. 2008, pp. 728–735.
- [8] P. Liu, Z. Qi, A. Aviles, and S. X. Tan, “A general method for multi-port active network reduction and realization,” in *Proc. IEEE International Conference on Behavioral Modeling and Simulation Workshop*, Sep. 2005, pp. 7–12.
- [9] I. Kalashnikova, B. van Bloemen Waanders, S. Arunajatesan, and M. Barone, “Stabilization of projection-based reduced order models for linear time-invariant systems via optimization-based eigenvalue reassignment,” *Computer Methods in Applied Mechanics and Engineering*, vol. 272, pp. 251–270, Apr. 2014.
- [10] I. M. Jaimoukha and E. M. Kasenally, “Implicitly restarted krylov subspace methods for stable partial realizations,” *SIAM J. Matrix Anal. Appl.*, vol. 18, no. 3, pp. 633–652, 1997.
- [11] V. Papakos and I. Jaimoukha, “A deflated implicitly restarted lanczos algorithm for model reduction,” in *Proc. 42nd IEEE Conference on Decision and Control*, vol. 3, Maui, HI, Dec. 2003, pp. 2902–2907.

- [12] C.-W. Ho, A. E. Ruehli, and P. A. Brennan, “The modified nodal approach to network analysis,” *IEEE Transactions on Circuits and Systems*, vol. 22, no. 6, pp. 504–509, Jun 1975.
- [13] M. Celik, C. Cangelaris, and A. Yagnour, “An all-purpose transmission-line model for interconnect simulation in spice,” *Microwave Theory and Techniques, IEEE Transactions on*, vol. 45, no. 10, pp. 1857–1867, 1997.
- [14] F. N. Najm, *Circuit Simulation*. Hoboken, NJ: Wiley-IEEE Press, 2010.
- [15] C.-W. Ho, A. Ruehli, and P. Brennan, “The modified nodal approach to network analysis,” *IEEE Transactions on Circuits and Systems*, vol. 22, no. 6, pp. 504–509, Jun. 1975.
- [16] J. Vlach and K. Singhal, *Computer Methods for Circuit Analysis and Design*, 2nd ed. Boston, MA: Kluwer Academic Publishers, 2003.
- [17] W. J. Rugh, *Linear system theory*. prentice hall Upper Saddle River, NJ, 1996, vol. 2.
- [18] P. Triverio, S. Grivet-Talocia, M. Nakhla, F. Canavero, and R. Achar, “Stability, causality, and passivity in electrical interconnect models,” *IEEE Transactions on Advanced Packaging*, vol. 30, no. 4, pp. 795–808, Nov 2007.

- [19] B. Nouri, M. S. Nakhla, and R. Achar, “Optimum order estimation of reduced macromodels based on a geometric approach for projection-based MOR methods,” *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 3, no. 7, pp. 1218–1227, Jul. 2013.
- [20] A. C. Antoulas, *Approximation of large-scale dynamical systems*. Siam, 2005, vol. 6.
- [21] S. Gugercin and A. C. Antoulas, “A survey of model reduction by balanced truncation and some new results,” *International Journal of Control*, vol. 77, no. 8, pp. 748–766, 2004.
- [22] K. Glover, “All optimal hankel-norm approximations of linear multivariable systems and their \mathcal{L}_∞ -error bounds,” *International Journal of Control*, vol. 36, no. 6, pp. 1115–1193, Jun. 1984.
- [23] K. Zhou, J. C. Doyle, K. Glover *et al.*, *Robust and optimal control*. Prentice hall New Jersey, 1996, vol. 40.
- [24] P. Benner, E. S. Quintana-Ortí, and G. Quintana-Ortí, “Computing optimal hankel norm approximations of large-scale systems,” in *Proc. 43rd IEEE Conference on Decision and Control*, vol. 3, Dec. 2004, pp. 3078–3083.

- [25] C. W. Rowley, T. Colonius, and R. M. Murray, “Model reduction for compressible flows using pod and galerkin projection,” *Physica D: Nonlinear Phenomena*, vol. 189, no. 1, pp. 115–129, 2004.
- [26] J. R. Phillips, “Projection frameworks for model reduction of weakly nonlinear systems,” in *Proceedings of the 37th Annual Design Automation Conference*. ACM, 2000, pp. 184–189.
- [27] J. R. Phillips, L. Daniel, and L. M. Silveira, “Guaranteed passive balancing transformations for model order reduction,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 22, no. 8, pp. 1027–1041, 2003.
- [28] K. Carlberg, C. Bou-Mosleh, and C. Farhat, “Efficient non-linear model reduction via a least-squares petrov–galerkin projection and compressive tensor approximations,” *International Journal for Numerical Methods in Engineering*, vol. 86, no. 2, pp. 155–181, 2011.
- [29] A. Odabasioglu, M. Celik, and L. Pileggi, “Prima: passive reduced-order interconnect macromodeling algorithm,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 17, no. 8, pp. 645–654, Aug 1998.
- [30] H. K. Khalil, “Nonlinear systems, 3rd,” *New Jersey, Prentice Hall*, vol. 9, 2002.

- [31] C. R. Paul, *Analysis of multiconductor transmission lines*. John Wiley & Sons, 2008.
- [32] T. Stykel, “Analysis and numerical solution of generalized lyapunov equations,” Ph.D. dissertation, Technischen universität Berlin, Berlin, Germany, Jan. 2002.
- [33] T. Kailath, *Linear Systems*. New Jersey: Prentice-Hall, 1980.
- [34] T. Penzl, “Numerical solution of generalized lyapunov equations,” *Advances in Computational Mathematics*, vol. 8, no. 1, pp. 33–48, Jan. 1998.
- [35] B. Men, Q. Zhang, X. Li, and C. Yang, “The stability of linear descriptor systems,” *International Journal of Information and Systems*, vol. 2, no. 3, pp. 362–374, 2006.
- [36] J. Ishihara and M. Terra, “On the lyapunov theorem for singular systems,” *IEEE Transactions on Automatic Control*, vol. 47, no. 11, pp. 1926–1930, Nov. 2002.
- [37] D. L. J. Debeljkovic and I. M. Buzurovic, “The stability of linear continuous singular systems in the sense of lyapunov: An overview,” *International Journal of Information and Systems Sciences*, vol. 7, no. 2-3, pp. 247–268, 2007.
- [38] R. H. Bartels and G. Stewart, “Solution of the matrix equation $ax + xb = c$ [f4],” *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.

- [39] A. Barraud, “A numerical algorithm to solve $\|x - x_0\| = q$,” in *Decision and Control including the 16th Symposium on Adaptive Processes and A Special Symposium on Fuzzy Set Theory and Applications, 1977 IEEE Conference on*. IEEE, 1977, pp. 420–423.
- [40] S. Hammarling, “Numerical solution of the stable, non-negative definite lyapunov equation lyapunov equation,” *IMA Journal of Numerical Analysis*, vol. 2, no. 3, pp. 303–323, 1982.
- [41] T. Stykel, “Numerical solution and perturbation theory for generalized lyapunov equations,” *Linear Algebra and its Applications*, vol. 349, no. 1-3, pp. 155–185, Jul. 2002.
- [42] L. Turyn, *Advanced Engineering Mathematics*. Boca Raton, FL: CRC Press, 2014.
- [43] G. H. Golub and C. F. van Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins University Press, 1996.
- [44] F. Sheng and D. Jiao, “A deterministic-solution based fast eigenvalue solver with guaranteed convergence for finite-element based 3-d electromagnetic analysis,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 7, pp. 3701–3711, Jul. 2013.

- [45] J. Lee, D. Chen, V. Balakrishnan, C.-K. Koh, and D. Jiao, “A quadratic eigenvalue solver of linear complexity for 3-d electromagnetics-based analysis of large-scale integrated circuits,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 3, pp. 380–390, Mar. 2012.
- [46] J. Lee, V. Balakrishnan, C.-K. Koh, and D. Jiao, “A linear-time complex-valued eigenvalue solver for full-wave analysis of large-scale on-chip interconnect structures,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 8, pp. 2021–2029, Aug. 2009.
- [47] P. Benner and J. Saak, “Numerical solution of large and sparse continuous time algebraic matrix riccati and lyapunov equations: A state of the art survey,” *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 32–52, Aug. 2013.
- [48] B. Nouri, M. S. Nakhla, and R. Achar, “Efficient reduced-order macromodels of massively coupled interconnect structures via clustering,” *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 3, no. 5, pp. 826–840, 2013.