UNIVERSITY OF THESSALY

DOCTORAL THESIS

Order reduction of large thermal and electrical models with system-theoretic techniques and matrix equation algorithms

Author:
George FLOROS

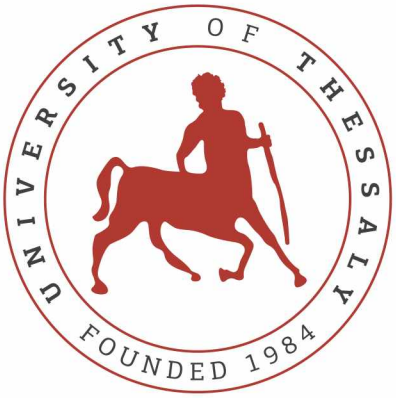
Supervisors:
Prof. Nestor EVMORFOPOULOS
Prof. George STAMOULIS
Prof. Gerasimos POTAMIANOS

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Electronics Lab
Department of Electrical and Computer Engineering

November 30, 2019



UNIVERSITY OF
THESSALY

Declaration of Authorship

I, George FLOROS, declare that this thesis titled, "Order reduction of large thermal and electrical models with system-theoretic techniques and matrix equation algorithms " and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

"*Madness has no purpose. Or reason. But it may have a goal...*"

Spock (Leonard Nimoy)

UNIVERSITY OF THESSALY

Abstract

Department of Electrical and Computer Engineering

Doctor of Philosophy

Order reduction of large thermal and electrical models with system-theoretic techniques and matrix equation algorithms

by George FLOROS

Computer simulation of modern IC subsystems, like power grids, interconnect structures and substrate regions, possesses a fundamental role in the EDA industry. The electric models of the aforementioned systems are enormous, and their functional simulation requires solving systems of equations with dimension that can reach several millions or even billions. Model Order Reduction (MOR) techniques provide attractive ways to reduce these highly complex models, replacing them by models with much smaller internal dimensions whose behavior at the input/output ports approximates that of the original models.

Balanced Truncation (BT) type MOR methods of internal states that are at the same time insufficiently controllable at the inputs and insufficiently observable at the outputs have the advantage of very reliable estimates for the accuracy of the reduced model. However, they have significant computational and storage costs for producing the reduced-order models, as they require the solution of Lyapunov matrix equations which is an intensive computational procedure, and also involve the storage of dense matrices, making them applicable only to models of few thousand states. To eliminate computational costs and storage needs, low-rank factorization methods have been developed that allow the use of these methods in real-world applications.

The emphasis in the present work is given in circuit and thermal models derived from integrated circuits and the calculation of the reduced-order models on specific frequency or time windows, as well as on the efficient implementation of Balanced Truncation type methods. More specifically, in most applications the circuit is intended to operate only in specific frequency windows, which means that the reduced-order model can become unnecessarily large to achieve approximation over all frequencies. Correspondingly, for the thermal models there is usually a final time in which all thermal effects can be assumed to have reached steady state. The focus is on model order reduction methods in specific frequency or time windows, combined with the efficient implementation of sparse methods for calculating Lyapunov-type matrix equations in order to manipulate large-scale input models.

Greek Abstract

Μείωση τάξης μεγάλων θερμικών και ηλεκτρικών μοντέλων με τεχνικές θεωρίας συστημάτων και αλγορίθμους εξισώσεων πινάκων

Η αριθμητική προσομοίωση των σύγχρονων υποσυστημάτων των ολοκληρωμένων κυκλωμάτων, όπως τα ηλεκτρικά δίκτυα, οι αγωγοί διασύνδεσης και οι περιοχές υποστρώματος, έχουν θεμελιώδη σημασία για την βιομηχανία ολοκληρωμένων κυκλωμάτων. Τα ηλεκτρικά μοντέλα των προαναφερθέντων συστημάτων είναι τεράστια και η λειτουργική τους προσομοίωση απαιτεί την επίλυση συστημάτων εξισώσεων με διαστάσεις που μπορούν να φτάσουν αρκετά εκατομμύρια. Οι τεχνικές υποβιβασμού τάξης μοντέλου (model order reduction - MOR) παρέχουν ελκυστικούς τρόπους για τη μείωση αυτών των πολύπλοκων μοντέλων, αντικαθιστώντας τα με μοντέλα με πολύ μικρότερες εσωτερικές διαστάσεις των οποίων η συμπεριφορά στις θύρες εισόδου / εξόδου (ports) προσεγγίζει αυτή των αρχικών μοντέλων.

Οι μέθοδοι MOR τύπου "εξισορρόπησης και αποκοπής" (Balanced Truncation - BT) των εσωτερικών καταστάσεων που είναι ταυτόχρονα μη επαρκώς ελέγξιμες από τις εισόδους και μη επαρκώς παρατηρήσιμες στις εξόδους, διαθέτουν το πλεονέκτημα των πολύ αξιόπιστων εκτιμήσεων για την ακρίβεια προσέγγισης του μειωμένου μοντέλου. Εμπεριέχουν, όμως, σημαντικό υπολογιστικό και αποθηκευτικό κόστος για την εξαγωγή των μειωμένων μοντέλων, καθώς απαιτούν την επίλυση δαπανηρών εξισώσεων πυκνών πινάκων τύπου Lyapunov, γεγονός που τις καθιστά άμεσα εφαρμόσιμες μόνο σε μοντέλα της τάξης λίγων χιλιάδων καταστάσεων. Για να απαλειφθεί το υπολογιστικό κόστος και οι ανάγκες αποθήκευσης, έχουν αναπτυχθεί μέθοδοι παραγοντοποίησης χαμηλού βαθμού (low-rank) που επιτρέπουν την χρήση αυτών των μεθόδων σε πραγματικές εφαρμογές.

Η έμφαση στην παρούσα εργασία δίνεται σε κυκλωματικά και θερμικά μοντέλα που προκύπτουν από ολοκληρωμένα κυκλώματα και στον υπολογισμό του ελαττωμένου μοντέλου σε συγκεκριμένα παράθυρα συχνοτήτων ή χρόνου, καθώς και στην αποδοτική υλοποίηση μεθόδων τύπου Balanced Truncation. Στις περισσότερες εφαρμογές το κύκλωμα προορίζεται να λειτουργεί μόνο σε συγκεκριμένα παράθυρα συχνοτήτων, πράγμα που σημαίνει ότι το το μοντέλο μειωμένης τάξης μπορεί να γίνει άσκοπα μεγάλο για να επιτευχθεί προσέγγιση σε όλες τις συχνότητες. Αντίστοιχα στα θερμικά μοντέλα υπάρχει συνήθως ένας τελικός χρόνος κατά τον οποίο όλες οι θερμικές επιδράσεις μπορούν να θεωρηθούν ότι έχουν φθάσει σε σταθερή κατάσταση. Η εστίαση δίνεται σε μεθόδους υποβιβασμού τάξης μοντέλου σε συγκεκριμένα παράθυρα συχνοτήτων ή χρόνου, σε συνδυασμό με την αποδοτική υλοποίησης αραιών μεθόδων για τον υπολογισμό των εξισώσεων πινάκων τύπου Lyapunov με σκοπό τον χειρισμό μεγάλων μοντέλων εισόδου.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Nestor Evmorfopoulos for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. George Stamoulis, and Prof. Gerasimos Potamianos, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives. My sincere thanks also goes to the other committee members for accepting the invitation.

I thank my fellow labmates in for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last four years. Last but not the least, I would like to thank anyone who supported me spiritually throughout writing this thesis and my life in general.

Contents

List of Figures

List of Tables

List of Abbreviations

IC	**I**nterconnect **C**ircuit
VLSI	**V**ery **L**arge **S**cale **I**ntegration
MOR	**M**odel **O**rder **R**eduction
CAD	**C**omputer **A**ided **D**esign
BT	**B**alanced **T**runcation
EDA	**E**lectronic **D**esign **A**utomation
EKS	**E**xtended **K**rylov **S**ubspace
ADI	**A**lternating **D**irection **I**mplicit
LTI	**L**inear **T**ime **I**nvariant
SVD	**S**ingular **V**alue **D**ecomposition
GPU	**G**raphic **P**rocessor **U**nit
IC	**I**ntegrated **C**ircuit
SOI	**S**ilicon **o**n **I**nsulator
FDM	**F**inite **D**ifference **M**ethod
PCG	**P**reconditioned **C**onjugate **G**radient
FEM	**F**inite **E**lement **M**ethod
ADI	**A**lternating **D**irection **I**mplicit
RHS	**R**ight **H**and **S**ide
ROM	**R**educed **O**rder **R**eduction
NN	**N**eural **N**et
LUT	**L**ook **U**p **T**able
PDE	**P**artial **D**ifferential **E**quation
MNA	**M**odified **N**odal **A**nalysis
ODE	**O**rdinary **D**ifferential **E**quations

Dedicated to . . . wish I knew

Chapter 1

Introduction

1.1 Introduction

The ongoing miniaturization, below the 45-nm process technology, of modern IC devices like transistors, has continued unabated for the last 50 years, in strict accordance with the provisions of Moore's Law. This has led to extremely complex circuits (modern processors contain many billions of transistors and are easily the most complex human structures) and to a corresponding increase of the problems associated with the analysis and simulation of the physical operation of these circuits. In particular, the performance and reliable operation of integrated circuits are largely determined by several critical subsystems such as the power distribution network, multi-conductor interconnections, and the semiconductor substrate (through which undesired digital signals and noise propagation are transmitted by digital signals into analog sections). The electrical models of the above subsystems are very large, consisting of hundreds of millions or billions of electrical elements (mostly resistors R, capacitors C, and inductors L), and their simulation depends on solving systems with dimensions up to 10^{11} which is becoming a huge mathematical problem. Even if their individual solution is feasible, it is completely impossible to combine them with the rest of the integrated circuit and with many consecutive time-steps or frequencies. However, for the above subsystems it is often not necessary to fully simulate all internal state variables (node voltages and branch currents), but only to calculate the responses in the time or frequency domain for a small subset of output terminals (ports) for given excitations at some input ports. In such cases, the very large electrical model can be replaced by a much smaller model whose behavior at the input/output ports is close to the original model. This process is called Model Order Reduction (MOR).

Downgrading large-scale electrical models is of paramount importance for companies that are active in integrated circuits and development of CAD tools, as well as in applications in the field of electricity. This field has attracted a great deal of research interest in the last ten years, during which there has been a dramatic increase in the dimension of the aforementioned electrical models. As a result, compact modeling of passive RLC interconnect networks is a major research area nowadays due to increasing complexity of high-performance VLSI designs [1]. The dimension reduction at the model level, is mostly done by mathematical algorithms, which produce a much smaller model, usually by a sufficiently large factor, that reproduces the original model response at a prescribed level of accuracy. MOR techniques are generally classified into two classes, namely moment-matching (or Krylov-subspace) techniques and balancing-type (or Gramian-based) techniques.

The idea behind Krylov-subspace techniques is to project the original models into an orthonormal subspace, which is called Krylov subspace. Projection procedures try to preserve the moment information of the original transfer function. Moreover, in order to ensure the passivity of the reduced-order model they usually adopt a congruence transformation that can preserve the reduced-order model passivity if the original model matrices are also in a passive form. A successful moment-matching algorithm that is widely used and developed in

the field of circuit simulation is the PRIMA algorithm [2], which uses the Krylov subspace vectors to form the projector and build a congruence transformation, which leads to passive models with the matched moments. Projection-based methods, however, have several drawbacks, with the most prominent one, is that they are not efficient for circuits with many input and output terminals since reduction cost depends on the number of ports. Moreover, it can be shown that the number of poles of reduced-models has a direct relationship to the number of terminals. Secondly, moment-matching methods do not generally preserve structure properties like reciprocity of circuit equations. Finally, it is difficult to apply moment-matching methods to high frequency models where the model parameters are usually frequency dependent.

The other major class of model order reduction for circuit models is by means of system-theoretic Balanced Truncation (BT) methods [3], where the weak unobservable and uncontrollable state variables are eliminated in order to produce the reduced-order models. Moreover, in system-theoretic techniques, like BT the reduced-order model is not depended from the number of ports, and has very satisfactory and reliable bounds for the approximation error. The BT methods generally produce nearly optimal models but they are more computationally expensive than projection-based methods. This happens due to the solution of the two Lyapunov matrix equations that govern BT-type methods with $O(n^3)$ time complexity and $O(n^2)$ storage needs, which is an important drawback.

In this dissertation, we focus on the compact modeling of on-chip interconnects and thermal models that arise in VLSI designs by applying system-theoretic techniques along with the efficient computation of Lyapunov matrix equation in order to make this type of methods applicable to real-world circuit simulation problems. This class of problems generally, produces linear time invariant (LTI) systems because interconnect parasitics are usually modeled as linear RLC circuits, while thermal models have a direct representation in RC circuits which we will describe later in this thesis.

1.2 Contributions

In order to address the limitations of the existing system-theoretic techniques for Model Order Reduction in VLSI interconnect problems, the purpose of this dissertation is to introduce an efficient BT framework for the reduction of large-scale circuit and thermal models by adopting state-of-the-art matrix equation solution algorithms that can handle large-scale input models.

More specifically, the contributions are described below:

- Develop procedures for applying large-scale matrix equations solvers to the solution of standard, time-limited and frequency-limited Lyapunov equations (which are different than the standard Lyapunov equations). This was made possible by implementing efficient computational choices by keeping sparse system matrices into their original forms (both for circuit and thermal models). By the proposed approaches the distill mathematical techniques are transformed into a complete algorithmic procedure, and the potential of the system-theoretic BT approach is evaluated in actual VLSI problems.

- For circuit simulation problems, MOR techniques based on BT offer very good error estimates and can provide compact models with any desired accuracy over the whole range of frequencies (from DC to infinity). However, in most applications the circuit is only intended to operate at specific frequency windows, which means that the reduced-order model can become unnecessarily large to achieve approximation over all frequencies. In this dissertation, we present a frequency-limited approach which,

combined with efficient low-rank sparse implementations of the Extended Krylov Subspace (EKS) and the Alternating Direction Implicit (ADI) methods, can handle large input models and provably leads to reduced-order models that are either smaller or exhibit better accuracy than full-frequency BT.

- For efficient thermal simulation temperature is not required to be computed at every point of the IC but only at certain hotspots, in order to assess the circuit's compliance with thermal specifications. This makes the thermal analysis problem amenable to Model Order Reduction techniques. System-theoretic techniques like Balanced Truncation offer very reliable bounds for the approximation error, which can be used to control the order and accuracy of the reduced models during creation, at the expense of greater computational complexity to create them. To this end, we propose a computationally efficient low-rank Balanced Truncation algorithm based on the EKS method, which retains all the system-theoretic advantages in the reduction of model order for fast hotspot thermal simulation.

1.3 Outline

The next chapters of the dissertation are organized as follows. Firstly, the relevant background information regarding Model Order Reduction for dynamical systems is presented in Chapter 2. Then, in Chapter 3 we describe a methodology for reducing large-scale regular and singular electrical models in limited-frequency windows along with efficient computational choices, making the method applicable to large-scale interconnect problems. Chapter 4 describes the theory behind thermal analysis for VLSI circuits and an efficient method for thermal analysis of the hotpots that arise in modern VLSI designs. Finally, Chapter 5 concludes the dissertation.

Chapter 2

Model Order Reduction

2.1 MOR Overview for LTI Systems

In this dissertation, we focus on LTI systems which described by the following set of equations:

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{x}(t) \end{aligned} \tag{2.1}$$

Model Order Reduction (MOR) aims at generating a reduced-order model:

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

with $\tilde{\mathbf{A}} \in \mathbb{R}^{r\times r}$, $\tilde{\mathbf{B}} \in \mathbb{R}^{r\times p}$, $\tilde{\mathbf{L}} \in \mathbb{R}^{q\times r}$ as shown in Fig. 2.1, which both exhibits $r << n$ and constitutes a good approximation of (2.1), in that the output error is bounded as $||\tilde{\mathbf{y}}(t) - \mathbf{y}(t)||_2 < \varepsilon||\mathbf{u}(t)||_2$ for given input $\mathbf{u}(t)$ and given small ε. The bound in the output error can be equivalently written in the frequency domain as $||\tilde{\mathbf{y}}(s) - \mathbf{y}(s)||_2 < \varepsilon||\mathbf{u}(s)||_2$ via Plancherel's theorem [4]. If

$$\begin{aligned} \mathbf{H}(s) &= \mathbf{L}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \\ \tilde{\mathbf{H}}(s) &= \tilde{\mathbf{L}}(s\mathbf{I} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{B}} \end{aligned} \tag{2.3}$$

are the transfer functions of the original and the reduced-order model, then the output error in the frequency domain is:

$$\begin{aligned} ||\tilde{\mathbf{y}}(s) - \mathbf{y}(s)||_2 &= ||\tilde{\mathbf{H}}(s)\mathbf{u}(s) - \mathbf{H}(s)\mathbf{u}(s)||_2 \\ &\leq ||\tilde{\mathbf{H}}(s) - \mathbf{H}(s)||_\infty ||\mathbf{u}(s)||_2 \end{aligned} \tag{2.4}$$

where $||.||_\infty$ is the induced $\mathcal{L}_2$ matrix norm, or $\mathcal{H}_\infty$ norm of a rational transfer function. Therefore, the output error can be bounded by bounding the distance between the transfer functions as $||\tilde{\mathbf{H}}(s) - \mathbf{H}(s)||_\infty < \varepsilon$.

Due to the large dynamical systems that arise in nowadays VLSI interconnect problems the MOR algorithms should satisfy the following three requirements.

- Must have a reasonable accuracy in the reduced-order model.
- Must have a satisfactory reduction rate in the reduced-order model.
- Must have an efficient computational complexity for producing the reduced-order model.

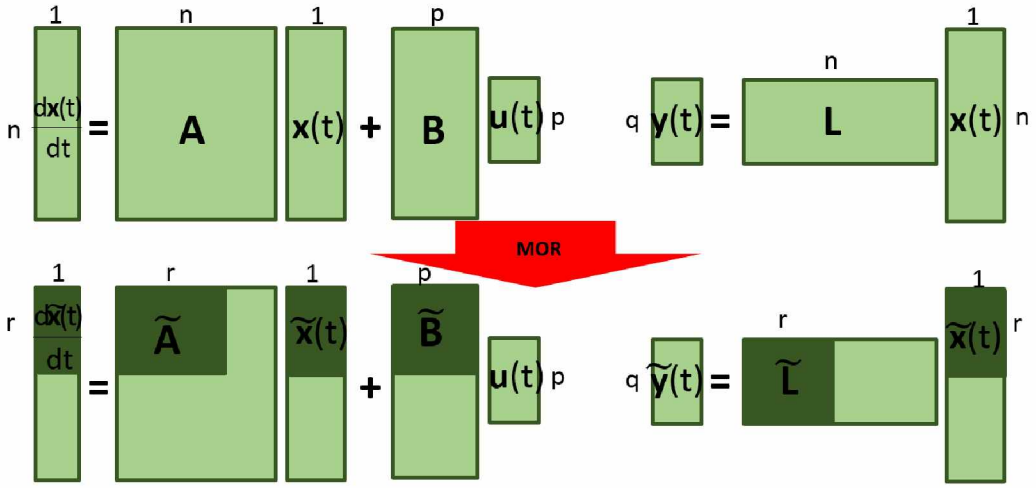


FIGURE 2.1: Model Order Reduction of LTI systems.

2.1.1 Moments

The transfer function of (2.3) is a function of s, and can be expanded into a moment expansion around $s = 0$ as follows:

$$\mathbf{H}(s) = \mathbf{M}_0 + \mathbf{M}_1 s + \mathbf{M}_2 s^2 + \mathbf{M}_3 s^3 \ldots \tag{2.5}$$

where the $\mathbf{M}_0$, $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$, ... are the moments of the transfer function. Specifically, in circuit simulation problems the $\mathbf{M}_0$ moment is the DC solution of the linear system. This means that the inductors of the circuit are considered as short circuits, and the capacitors as open circuits. Moreover, the $\mathbf{M}_1$ moment is the Elmore delay of the linear model, which calculates the time for a signal at the input port to reach the output port. The Elmore formula can be defined as follows:

$$t_{ed} = \int_0^\infty t\mathbf{h}(t)dt \tag{2.6}$$

where $\mathbf{h}(t)$ is the impulse response function. By applying the Laplace transformation in the transfer function of (2.3) and expanding the exponential factor with a Taylor series, it can be shown that the Elmore delay corresponds to the $\mathbf{M}_1$ moment of the transfer function.

2.1.2 Stability

The eigenvalues and the poles of a system are related to the stability of the system. The basic property of a stable system is that ensures that its output signal is limited in the time domain. Recalling the linear model of (2.1), the system is stable if and only if, it holds that for all the eigenvalues λ_j, $Re(\lambda_j) \leq 0$ and for all eigenvalues that $Re(\lambda_j) = 0$, λ_j is simple. In this case the system matrix $\mathbf{A}$ is proved to be stable.

Stability of the system is strongly associated with several properties. The most important property is that if $\mathbf{A}$ is stable the inverse and the transpose matrix, $\mathbf{A}^{-1}$ and $\mathbf{A}^T$ respectively, are also stable.

2.1.3 Passivity

The stability of physical structures is usually a common property. However, stability is not always strong enough for VLSI applications. A stable structure can become unstable if non-linear components are connected to it. Therefore passivity is a stronger property than stability

and it shows, that a passive model is incapable of generating energy. Generally for linear systems that are stable and passive, it would be preferable to apply a reduction process in order to preserve this properties after the reduction.

In order to define the passivity property intuitively, we first consider a linear circuit model with n-ports, then the passivity is defined with the following way. If the power absorbed by this n-port circuit is greater or equal to zero for all frequencies s such that $Re(s) \geq 0$, then this model is passive.

Following this example, we can now define the passivity based on this argument. As a result, the transfer function $\mathbf{H}(s)$ of a linear system is passive if and only if

$$\mathbf{H}(s) + \mathbf{H}^*(s) \geq 0 \tag{2.7}$$

with $Re(s) \geq 0$ for all s.

2.2 Moment Matching Methods

One of the most important and successful MOR methods for linear systems is based on Krylov subspaces. They are usually called moment-matching methods and they are very efficient in circuit simulation problems. Methods based on moment matching are formulated in order to have for a direct application to the linear model of (2.1).

By applying the Laplace transformation to (2.1), we obtain s domain equations as:

$$\begin{aligned} s\mathbf{X}(s) - \mathbf{X}(0) &= \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s) \\ \mathbf{Y}(s) &= \mathbf{L}\mathbf{X}(s) \end{aligned} \tag{2.8}$$

Assuming that $\mathbf{X}(0) = 0$ and an impulse response in applied in $\mathbf{U}(s)$ (i.e. $\mathbf{U}(s) = 1$) then the above system of equations is:

$$\begin{aligned} (s\mathbf{I} - \mathbf{A})\mathbf{X}(s) &= \mathbf{B} \\ \mathbf{Y}(s) &= \mathbf{L}\mathbf{X}(s) \end{aligned} \tag{2.9}$$

and by expanding the Taylor series at $s = 0$:

$$(s\mathbf{I} - \mathbf{A})(\mathbf{x}_0 + \mathbf{x}_1 s + \mathbf{x}_2 s^2 + \dots) = \mathbf{B} \tag{2.10}$$

Finally, we can obtain a moment computation formula as follows:

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{A}^{-1}\mathbf{B}, & \mathbf{m}_0 &= \mathbf{L}\mathbf{x}_0 \\ \mathbf{x}_1 &= \mathbf{A}^{-1}\mathbf{x}_0, & \mathbf{m}_1 &= \mathbf{L}\mathbf{x}_1 \\ \mathbf{x}_2 &= \mathbf{A}^{-1}\mathbf{x}_1, & \mathbf{m}_2 &= \mathbf{L}\mathbf{x}_2 \\ & & & \dots \end{aligned} \tag{2.11}$$

while generally the i-th moment can be computed as

$$\mathbf{m}_i = \mathbf{L}(\mathbf{A}^{-1})^i\mathbf{A}^{-1}\mathbf{B} \tag{2.12}$$

For moment-matching reduction techniques the goal is the derivation of a reduced-order model where some moments $\tilde{\mathbf{m}}_i$ of the reduced-order transfer function $\tilde{\mathbf{H}}(s)$ match some moments of the original transfer function $\mathbf{H}(s)$.

Let us now denote the two projection matrices onto a lower dimensional subspace, as $\mathbf{W} \in \mathbb{R}^{n\times r}$ and $\mathbf{V} \in \mathbb{R}^{r\times n}$ respectively. This matrices can be computed from the associated moment vectors using one or more expansion points. For the ease of representation we

assume that $s = 0$, then the matrices $\mathbf{W}$ and $\mathbf{V}$ are defined as follows:

$$\begin{aligned} range(\mathbf{W}) &= span\{\mathbf{A}^{-1}\mathbf{B}, (\mathbf{A}^{-1})^2\mathbf{B}, \ldots, (\mathbf{A}^{-1})^r\mathbf{B}\} \\ range(\mathbf{V}) &= span\{\mathbf{L}, \mathbf{A}^{-T}\mathbf{L}, (\mathbf{A}^{-T})^2\mathbf{B}, \ldots, (\mathbf{A}^{-T})^r\mathbf{B}\} \end{aligned} \tag{2.13}$$

The computed reduced-order model matches the first $2r$ moments and is obtained by the following matrices

$$\tilde{\mathbf{A}} = \mathbf{W}^T\mathbf{A}\mathbf{V}, \quad \tilde{\mathbf{B}} = \mathbf{W}^T\mathbf{B}, \quad \tilde{\mathbf{L}} = \mathbf{L}\mathbf{V} \tag{2.14}$$

and provides a good approximation around 0.

The first approach on moment-matching MOR for circuit simulation problems was the Asymptotic Waveform Evaluation method (AWE) presented in [5]. It is an efficient frequency-domain analysis approach, however, the method was suffered due to the explicit moments computation, causing numerical instability. In order to overcome the problem associated with the AWE method, a more recent research work led to a numerically robust method of Pade via Lanczos (PVL) [6] where the Lanczos process, which has better numerical stability for computing the eigenvalues of a matrix, was deployed to find the Krylov subspaces of the corresponding matrices. This method computes orthogonal bases of $\mathbf{V}$ and $\mathbf{W}$ implicitly, and also the moments are implicitly matched by the reduced model. This method has been succeeded from the most successful moment-matching reduction method for passive reduced-order interconnect macromodeling (PRIMA) [2] and the block structure preservation alternates known as SPRIM [7] and BSMOR [8].

Because the field of moment matching is mature, recent methods try to address the MOR problem by attacking in three domains, sparsity, terminal-merging, and efficient frequency selection points. Firstly, since the reduced-order models are usually dense, a recent approach tried to exploit sparsity preservation techniques [9]. Moreover, since the dimension of the reduced model has a direct relationship with the input and output number of ports, authors in [10,11] compute a reduced-order model for a subset of terminals, while the authors in [12] firstly partition the circuit in several blocks, and reduce each one separately. Finally, for matching the transfer function over a larger frequency range, rational methods including multiple expansion points such as [13] were developed, as well as efficient frequency hopping algorithms for choosing the expansion points [14]. In general, moment-matching techniques suffer from several drawbacks, the most prominent of which is that they do not offer any a-priori estimation for the approximation error. This can result in reduced models that are not very accurate or of sufficient small order. Error-bounds are very important in VLSI interconnect modeling problems and we need to employ system-theoretic techniques in order to address this problem.

2.3 Balanced Truncation Methods

The other important class of MOR methods is the Balanced Truncation type methods which generally deliver better reduced models than Krylov subspace techniques, by making an extra effort in choosing the projection subspaces based on the Controllability and Observability of the Linear Time Invariant (LTI) system. The controllability and observability Gramian matrices are:

$$\begin{aligned} \mathbf{P} &= \int_0^\infty exp(\mathbf{A}t)\mathbf{B}\mathbf{B}^T exp(\mathbf{A}t)^T dt \\ \mathbf{Q} &= \int_0^\infty exp(\mathbf{A}t)^T\mathbf{L}^T\mathbf{L} exp(\mathbf{A}t) dt \end{aligned} \tag{2.15}$$

which are equivalently derived by the solution of the Lyapunov matrix equations [15]:

$$\begin{aligned} \mathbf{AP} + \mathbf{PA}^T &= -\mathbf{BB}^T \\ \mathbf{A}^T\mathbf{Q} + \mathbf{QA} &= -\mathbf{L}^T\mathbf{L} \end{aligned} \tag{2.16}$$

The controllability Gramian $\mathbf{P}$ characterizes the input-to-state behavior, i.e. the degree to which the states are controllable (reachable) by the inputs, while the observability Gramian $\mathbf{Q}$ characterizes the state-to-output behavior, i.e. the degree to which the states are observable at the outputs. A reduced-order model can, in principle, be obtained by eliminating (truncating) the states that are difficult to reach or observe. However, in the original state-space coordinates there might be states that are difficult to reach but easy to observe, and vice versa. The process of "balancing" is to transform the state vector into a new coordinate system where for every state the degree of difficulty is the same for both reaching and observing it. There exists such a transformation $\mathbf{Tx}(t)$, which leads to a new model:

$$\begin{aligned} \frac{d(\mathbf{Tx}(t))}{dt} &= \mathbf{TAT}^{-1}(\mathbf{Tx}(t)) + \mathbf{TBu}(t) \\ \mathbf{y}(t) &= \mathbf{LT}^{-1}(\mathbf{Tx}(t)) \end{aligned} \tag{2.17}$$

(thus preserving the transfer function $\mathbf{H}(s)$) and makes [16]:

$$\mathbf{P} = \mathbf{Q} = diag(\sigma_1, \sigma_2, ..., \sigma_n) \tag{2.18}$$

where $\sigma_i, i = 1, \ldots, n$ are known as the Hankel singular values (HSVs) of the model and are equal to the square roots of the eigenvalues of the product $\mathbf{PQ}$ (in any coordinate system of state-space), i.e. $\sigma_i = \sqrt{\lambda_i(\mathbf{PQ})}, i = 1, \ldots, n$. In the balanced model (2.17) the states that are easier to reach and observe correspond to the largest HSVs, and if r of them are kept (truncating the $n - r$ states corresponding to the smallest HSVs) it can be shown that the distance between the original and the reduced-order transfer functions is bounded as [15]:

$$||\mathbf{H}(s) - \tilde{\mathbf{H}}(s)||_\infty \leq 2(\sigma_{r+1} + \sigma_{r+2} + ... + \sigma_n) \tag{2.19}$$

The latter is an "a-priori" criterion for selecting the order of the reduced model for a desired output error tolerance ε, and is a significant advantage of BT-type methods for MOR. The main steps of BT are summarized in the following Algorithm 1 [17]:

Algorithm 1 MOR by Balanced Truncation

1: Solve the Lyapunov equations (2.16) to obtain the Gramian matrices $\mathbf{P}$ and $\mathbf{Q}$
2: Compute the eigenvalue decomposition of $\mathbf{PQ}$, or equivalently the singular value decomposition (SVD) of the product of the Cholesky factors $\mathbf{P} = \mathbf{Z}_P\mathbf{Z}_P^T$ and $\mathbf{Q} = \mathbf{Z}_Q\mathbf{Z}_Q^T$, i.e. $\mathbf{Z}_P^T\mathbf{Z}_Q = \mathbf{U\Sigma V}$
3: Compute the truncated part of the balancing transformations $\mathbf{T}_{(r\times n)} = \mathbf{\Sigma}_{(r\times r)}^{-1/2}\mathbf{V}_{(r\times n)}\mathbf{Z}_Q^T$ and $\mathbf{T}_{(n\times r)}^{-1} = \mathbf{Z}_P\mathbf{U}_{(n\times r)}\mathbf{\Sigma}_{(r\times r)}^{-1/2}$, and the corresponding reduced-order model matrices as $\tilde{\mathbf{A}} = \mathbf{T}_{(r\times n)}\mathbf{A}\mathbf{T}_{(n\times r)}^{-1}$, $\tilde{\mathbf{B}} = \mathbf{T}_{(r\times n)}\mathbf{B}$, $\tilde{\mathbf{L}} = \mathbf{L}\mathbf{T}_{(n\times r)}^{-1}$

BT-type algorithms provide generally better approximations, since they have explicit error bounds. While this class of methods has not the same maturity like moment-matching algorithms, in the last decade have been adopted in circuit simulation problems. Recent works in [18–20] try to address the passivity preserving problem for reduced-order interconnect problems, by computing passive models through balancing and truncation methods. Moreover, similar to moment matching several approaches try to find a better approximation

in specific frequency ranges [21–23], where circuits usually work. However, recent trends try to address the computation of the system Gramians [24] which is the first step to perform the BT algorithm. For large scale problems of order 10^3 and larger, the Gramians can not be determined exactly, but they are only approximated. In the next chapter of this dissertation, we provide an extensive overview of computing reduced-order models with system-theoretic techniques and we describe a comprehensive methodology for efficient matrix solution algorithms.

Chapter 3

Frequency-Limited Reduction of Regular and Singular Circuit Models via Low-Rank Model Order Reduction

3.1 Introduction

Computer simulation of modern IC subsystems, like power grids, interconnect structures and substrate regions, possesses a fundamental role in the EDA industry. The electric models of the aforementioned systems are enormous, and their functional simulation requires solving systems of equations with dimension that can reach several millions or even billions. Model Order Reduction (MOR) techniques provide attractive ways to reduce these highly complex models, replacing them by models with much smaller internal dimensions whose behavior at the input/output ports approximates that of the original models.

As we stated previously, MOR methods are divided in two main categories. Moment matching techniques [2] are well established due to their computational efficiency to produce the reduced-order models for circuit simulation problems. Their drawback is that the size of the reduced models is based on an ad-hoc choice of the number of matching moments, since they do not provide an "a priori" metric for the accuracy of the reduced models. On the other hand, system theoretic techniques like Balanced Truncation (BT) [3] have very satisfactory and reliable bounds for the approximation error. However, BT techniques require the solution of Lyapunov matrix equations which are very expensive computationally, and also involve the storage of dense matrices even if the system matrices are sparse. To alleviate the computational cost and storage needs, low-rank solution methods such as the Extended Krylov Subspace (EKS) and Alternating Direction Implicit have been developed [25, 26].

The majority of BT-type methods attempt to approximate the original model over the whole frequency range (from DC to infinity). In most practical applications, however, we are only interested in a specific finite frequency range. Frequency-weighted BT methods have been proposed in the past [27, 28], where a user-specified frequency weighting function is introduced so as to obtain solutions of Lyapunov matrix equations that improve accuracy according to this particular function. The problem is that the specification of the weighting-function is not straightforward [29], and the user would rather give only the intended frequency range as input to the BT method.

A different *frequency-limited* BT method has been proposed in the past in [30] (and recently rediscovered and analyzed in detail in [31]), in which only a frequency range needs to be specified instead of a vague frequency-weighted function. The purpose of this dissertation is to introduce the frequency-limited BT framework for the reduction of large circuit models, with its specific goals and contributions being to: (i) develop procedures for applying both the

EKS and the ADI method to the solution of frequency-limited Lyapunov equations (which are different than the standard Lyapunov equations), (ii) implement efficient computational choices by keeping sparse system matrices into their original forms (both for regular and singular circuit models), (iii) distill mathematical techniques into a complete algorithmic procedure, (iv) evaluate the potential of the frequency-limited BT approach in actual circuit benchmarks. In particular, experimental results demonstrate that frequency-limited BT can produce reduced-order models with either smaller size or superior accuracy compared to standard BT in a specific frequency range.

The rest of the chapter, is organized as follows. Section 3.2 describes previous work on existing MOR techniques developed for circuit simulation problems. Section 3.3 presents the theoretical background of BT for the reduction of regular and singular circuit models, and the low-rank approximate solution of Lyapunov equations along with the frequency-limited BT methodology. Section 3.4 presents our main contributions on the application of low-rank EKS to frequency-limited Lyapunov equations, as well as its efficient execution by sparse matrix manipulations (both for regular and singular circuit models), while Section 3.5 presents the application of the ADI method to frequency-limited Lyapunov equations. Sections 3.6 and 3.7 describe the proposed complete procedures for the EKS and the ADI method respectively. Finally, Section 3.8 presents our experimental results.

3.2 Related Work

In this section, we briefly describe some previous works in the area of MOR techniques developed for circuit simulation problems. As mentioned before, mainly model order reduction methods have so far relied on moment matching and system theoretic techniques.

The moment matching techniques like [2] and the structure preserving alternate [7] perform moment matching by projecting the original system onto a Krylov subspace. Furthermore, the frequency weighting approach has been incorporated in these methods by multi-point moment matching approaches [32], in order to achieve better accuracy in a specific frequency range. For this kind of methods, the cost associated with model derivation is proportional to the number of ports, and the projection matrix can become dense and very large with the increasing number of ports. To address this challenge, research works introduce decentralized techniques like [33] or terminal reduction techniques [34, 35] before applying a standard moment matching method to a multi-port system. These algorithms exploit correlations between different ports in order to merge them together. However, this is not always applicable to practical circuits and is usually an error-prone process. Despite the successful application of these methods in circuit simulation problems they do not yet provide any error control for the accuracy of the reduced models.

In contrast to moment matching techniques, system theoretic approaches like Balanced Truncation type methods [3,36] are capable of providing a global error bound, without having to compute the reduced-order models first. The adaptive choice of the order of the approximate models based on a predefined error bound has led research works to seek reduce models in frequency windows that circuits are usually work in order to produce smaller models or models that exhibit better accuracy. Based on this fact, several frequency weighted BT methods are developed in order to exploit this attribute [27, 28, 37, 38]. While these methods provide better approximation the need of input and output weights which are usually not explicitly specified, hinders the applicability of these methods.

The approach in [39] bears a resemblance to the proposed method since a frequency-limited Gramian approach is used, where only the end frequencies are required. However, this method is inspired by the modified frequency-limited Gramians [62] which they usually

do not exhibit fast eigenvalue decay [31] and they induce a significant computational cost making this method difficult to apply in large-scale circuit models.

Clearly, the potential of the frequency-limited Gramians that introduced in [30] has not yet been explored in the context of circuit simulation. The proposed methodology, alleviate the computational costs by applying state-of-the-art sparse numerical techniques in order to compute the frequency-limited Gramians making this method amenable to large circuit models. Moreover, the singular descriptor models are treated with sparse matrix operations, without introducing significant computational cost to the proposed methodology.

3.3 Background

3.3.1 MOR by Balanced Truncation for Circuit Simulation Problems

Consider the Modified Nodal Analysis (MNA) description of an n-node, m-branch (inductive), p-input, and q-output RLC circuit in the time domain:

$$\begin{aligned} \begin{pmatrix} \mathbf{G} & \mathbf{W} \\ -\mathbf{W}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{i}(t) \end{pmatrix} + \begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{v}}(t) \\ \dot{\mathbf{i}}(t) \end{pmatrix} &= \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{0} \end{pmatrix} \mathbf{u}(t) \\ \mathbf{y}(t) &= \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{i}(t) \end{pmatrix} + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{3.1}$$

where $\mathbf{G} \in \mathbb{R}^{n\times n}$ (node conductance matrix), $\mathbf{C} \in \mathbb{R}^{n\times n}$ (node capacitance matrix), $\mathbf{M} \in \mathbb{R}^{m\times m}$ (branch inductance matrix), $\mathbf{W} \in \mathbb{R}^{n\times m}$ (node-to-branch incidence matrix), $\mathbf{v} \in \mathbb{R}^n$ (vector of node voltages), $\mathbf{i} \in \mathbb{R}^m$ (vector of inductive branch currents), $\mathbf{u} \in \mathbb{R}^p$ (vector of input excitations from current sources), $\mathbf{B}_1 \in \mathbb{R}^{n\times p}$ (input-to-node connectivity matrix), $\mathbf{y} \in \mathbb{R}^q$ (vector of output measurements), $\mathbf{L}_1 \in \mathbb{R}^{q\times n}$ (node-to-output connectivity matrix), $\mathbf{D} \in \mathbb{R}^{q\times p}$ (input-to-output connectivity matrix). Without loss of generality, we assume in the above that any voltage sources have been transformed to Norton-equivalent current sources, and that all outputs are obtained at the nodes as node voltages. Further, we are denoting $\dot{\mathbf{v}}(t) \equiv \frac{d\mathbf{v}(t)}{dt}$ and $\dot{\mathbf{i}}(t) \equiv \frac{d\mathbf{i}(t)}{dt}$.

If we now denote the model *order* as $N \equiv n+m$ and the *state* vector as $\mathbf{x}(t) \equiv \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{i}(t) \end{pmatrix}$, and also:

$$\mathbf{A} \equiv -\begin{pmatrix} \mathbf{G} & \mathbf{W} \\ -\mathbf{W}^T & \mathbf{0} \end{pmatrix}, \quad \mathbf{E} \equiv \begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix},$$
$$\mathbf{B} \equiv \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{L} \equiv \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} \end{pmatrix}$$

then the expression (3.1) can be written in the following generalized state-space form, or so-called *descriptor* form:

$$\begin{aligned} \mathbf{E}\frac{d\mathbf{x}(t)}{dt} &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{3.2}$$

The objective of MOR is to produce a reduced-order model:

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

where $\tilde{\mathbf{A}}, \tilde{\mathbf{E}} \in \mathbb{R}^{r\times r}$, $\tilde{\mathbf{B}} \in \mathbb{R}^{r\times p}$, $\tilde{\mathbf{L}} \in \mathbb{R}^{q\times r}$, and in which both the order $r << N$ and the output error is bounded as $||\tilde{\mathbf{y}}(t) - \mathbf{y}(t)||_2 < \varepsilon||\mathbf{u}(t)||_2$ for given input $\mathbf{u}(t)$ and given small

ε. The bound in the output error can be equivalently written in the frequency domain, since in circuit simulation problems the interest is usually in the approximation of the frequency response, as $||\tilde{\mathbf{y}}(s)-\mathbf{y}(s)||_2 < \varepsilon||\mathbf{u}(s)||_2$ via Plancherel's theorem [4]. If

$$\mathbf{H}(s) = \mathbf{L}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}+\mathbf{D}$$
$$\tilde{\mathbf{H}}(s) = \tilde{\mathbf{L}}(s\tilde{\mathbf{E}}-\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{B}}+\mathbf{D}$$

are the transfer functions of the original and the reduced-order model, then the output error in the frequency domain is:

$$\begin{aligned} ||\tilde{\mathbf{y}}(s)-\mathbf{y}(s)||_2 &= ||\tilde{\mathbf{H}}(s)\mathbf{u}(s)-\mathbf{H}(s)\mathbf{u}(s)||_2 \\ &\leq ||\tilde{\mathbf{H}}(s)-\mathbf{H}(s)||_\infty ||\mathbf{u}(s)||_2 \end{aligned} \tag{3.4}$$

where $||.||_\infty$ is the induced $\mathcal{L}_2$ matrix norm, or $\mathcal{H}_\infty$ norm of a rational transfer function. Therefore, in order to bound the output error, we need to bound the distance between the transfer functions $||\tilde{\mathbf{H}}(s)-\mathbf{H}(s)||_\infty < \varepsilon$.

Balanced Truncation (BT) and related methods for MOR make use of the controllability and observability Gramian matrices:

$$\begin{aligned} \mathbf{P} &= \int_0^\infty exp(\mathbf{E}^{-1}\mathbf{A}t)\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T}exp(\mathbf{E}^{-1}\mathbf{A}t)^T dt \\ \mathbf{Q} &= \int_0^\infty exp(\mathbf{E}^{-1}\mathbf{A}t)^T\mathbf{L}^T\mathbf{L}exp(\mathbf{E}^{-1}\mathbf{A}t) dt \end{aligned} \tag{3.5}$$

which are equivalently derived by the solution of the Lyapunov matrix equations [62]:

$$\begin{aligned} (\mathbf{E}^{-1}\mathbf{A})\mathbf{P}+\mathbf{P}(\mathbf{E}^{-1}\mathbf{A})^T &= -(\mathbf{E}^{-1}\mathbf{B})(\mathbf{E}^{-1}\mathbf{B})^T \\ (\mathbf{E}^{-1}\mathbf{A})^T\mathbf{Q}+\mathbf{Q}(\mathbf{E}^{-1}\mathbf{A}) &= -\mathbf{L}^T\mathbf{L} \end{aligned} \tag{3.6}$$

where we have assumed that the matrix $\mathbf{E}$ is nonsingular (we will present a treatment for the case of singular $\mathbf{E}$ in the next sub-section).

The controllability Gramian $\mathbf{P}$ characterizes the input-to-state behavior, i.e. the degree to which the states are controllable (reachable) by the inputs, while the observability Gramian $\mathbf{Q}$ characterizes the state-to-output behavior, i.e. the degree to which the states are observable at the outputs. A reduced-order model can, in principle, be obtained by eliminating (truncating) the states that are difficult to reach or observe. However, in the original state-space coordinates there are states that are difficult to reach but easy to observe, and vice versa. The process of "balancing" is to transform the state vector into a new coordinate system where for every state the degree of difficulty is the same for both reaching and observing it. There exists such a transformation $\mathbf{T}\mathbf{x}(t)$, which leads to a new model:

$$\begin{aligned} \mathbf{T}\mathbf{E}\mathbf{T}^{-1}\frac{d(\mathbf{T}\mathbf{x}(t))}{dt} &= \mathbf{T}\mathbf{A}\mathbf{T}^{-1}(\mathbf{T}\mathbf{x}(t))+\mathbf{T}\mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{T}^{-1}(\mathbf{T}\mathbf{x}(t))+\mathbf{D}\mathbf{u}(t) \end{aligned} \tag{3.7}$$

(thus preserving the transfer function H(s)) and makes [62]:

$$\mathbf{P} = \mathbf{Q} = diag(\sigma_1,\sigma_2,....,\sigma_N) \tag{3.8}$$

where $\sigma_i, i = 1,\ldots,N$ are known as the Hankel singular values (HSVs) of the model and are equal to the square roots of the eigenvalues of product $\mathbf{PQ}$ (in any coordinate system of state-space), i.e. $\sigma_i = \sqrt{\lambda_i(\mathbf{PQ})}, i = 1,\ldots,N$. In the balanced model (8) the states that are easier to reach and observe correspond to the largest HSVs, and if r of them are kept (truncating the $N-r$ states corresponding to the smallest HSVs) it can be shown that the

distance between the original and the reduced-order transfer functions is bounded as:

$$||\mathbf{H}(s) - \tilde{\mathbf{H}}(s)||_\infty \leq 2(\sigma_{r+1} + \sigma_{r+2} + ... + \sigma_N) \tag{3.9}$$

The latter is an "a-priori" criterion for selecting the order of the reduced model for a desired output error tolerance ε, and is a significant advantage of BT-type methods for MOR. The main steps of BT are summarized in Algorithm 2.

Algorithm 2 MOR by Balanced Truncation for descriptor systems

1: Solve the Lyapunov equations (3.6) to obtain the Gramian matrices $\mathbf{P}$ and $\mathbf{Q}$ in low-rank format as described in Sections 3.4 and 3.5
2: Compute the eigenvalue decomposition of $\mathbf{PQ}$, or equivalently the singular value decomposition (SVD) of the product of the Cholesky factors $\mathbf{P} = \mathbf{Z}_P\mathbf{Z}_P^T$ and $\mathbf{Q} = \mathbf{Z}_Q\mathbf{Z}_Q^T$, i.e. $\mathbf{Z}_P^T\mathbf{Z}_Q = \mathbf{U\Sigma V}$
3: Compute the truncated part of the balancing transformations $\mathbf{T}_{(r\times N)} = \mathbf{\Sigma}_{(r\times r)}^{-1/2}\mathbf{V}_{(r\times N)}\mathbf{Z}_Q^T$ and $\mathbf{T}_{(N\times r)}^{-1} = \mathbf{Z}_P\mathbf{U}_{(N\times r)}\mathbf{\Sigma}_{(r\times r)}^{-1/2}$, and the corresponding reduced-order model matrices as $\tilde{\mathbf{E}} = \mathbf{T}_{(r\times N)}\mathbf{E}\mathbf{T}_{(N\times r)}^{-1}$, $\tilde{\mathbf{A}} = \mathbf{T}_{(r\times N)}\mathbf{A}\mathbf{T}_{(N\times r)}^{-1}$, $\tilde{\mathbf{B}} = \mathbf{T}_{(r\times N)}\mathbf{B}$, $\tilde{\mathbf{L}} = \mathbf{L}\mathbf{T}_{(r\times N)}^{-1}$

The main drawback of BT is the significant computational and memory cost for deriving the reduced model, which seriously hinders the applicability of BT for the reduction of large-scale models (with N more than a few thousand states or so). That is because the solution of Lyapunov equations, the Cholesky factorization and the SVD are all computationally expensive tasks of complexity $O(N^3)$, and also involve dense matrices since the Gramians $\mathbf{P}, \mathbf{Q}$ are dense even if the system matrices $\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{L}$ are sparse.
However, it is almost always the practical case that the number of inputs and outputs is much smaller than the number of states, i.e. $p, q << N$. This means that the products $\mathbf{BB}^T$ and $\mathbf{L}^T\mathbf{L}$ will have low numerical rank compared to N, and this will also hold for the corresponding Gramians [40], allowing their own approximation by low-rank products instead of full Cholesky factorizations, i.e. $\mathbf{P} \approx \mathbf{Z}_P\mathbf{Z}_P^T$ and $\mathbf{Q} \approx \mathbf{Z}_Q\mathbf{Z}_Q^T$ with $\mathbf{Z}_P, \mathbf{Z}_Q \in \mathbb{R}^{N\times k}$ $(k << N)$. This greatly reduces the memory requirements, as well as the complexity of the factorization and SVD which are now of size k instead of full N, leaving the solution of Lyapunov equations as the main computational task of low-rank BT.

Two recent classes of algorithms that have been developed for directly solving the Lyapunov equations in low-rank factorized form are the Alternating Direction Implicit (ADI) [26] and the projection-type or Krylov-subspace methods [41]. The ADI method exhibits fast convergence but requires the input of a number of shift parameters, whose choice greatly affects convergence but until recently relied on unclear heuristics and was very problem-dependent. On the other hand, projection-type methods do not depend on the selection of specific parameters and their algorithmic implementation is more straightforward and well-studied, having been successfully used for several years for the solution of conventional linear systems of equations. However, they generally have not been competitive with ADI methods, until the recent development of the extended Krylov subspace (EKS) method [25] which employs two complementary subspaces to radically speed up convergence [42]. In the next Sections we adopt both the EKS and ADI methods for the low-rank solution of Lyapunov equations arising in the reduction by BT of large-scale circuit models and provide details for efficient application in both cases.

3.3.2 Handling of Singular Descriptor Models

In certain circuit simulation problems the matrix $\mathbf{E}$ might be singular. A method for dealing with such models is to compute spectral projections onto the left and right deflating subspaces corresponding to the finite eigenvalues of the model [43], which is computationally prohibitive for large-scale systems. However, in circuit problems singular descriptor models typically result when there are some nodes, say n_2, where no capacitance is connected, leading to corresponding all-zero rows and columns in the submatrix $\mathbf{C}$ (the submatrix $\mathbf{M}$ of inductive branches is always nonsingular if the circuit contains no voltage sources). If the n_2 nodes with no capacitance connection are enumerated last, and the remaining $n_1 = n - n_2$ nodes first, then (3.1) can pe partitioned as follows:

$$\begin{aligned}
&\begin{pmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} & \mathbf{W}_1 \\ \mathbf{G}_{12}^T & \mathbf{G}_{22} & \mathbf{W}_2 \\ -\mathbf{W}_1^T & -\mathbf{W}_2^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \\ \mathbf{i}(t) \end{pmatrix} + \\
&\begin{pmatrix} \mathbf{C}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{v}}_1(t) \\ \dot{\mathbf{v}}_2(t) \\ \dot{\mathbf{i}}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{0} \end{pmatrix} \mathbf{u}(t) \\
&\mathbf{y}(t) = \begin{pmatrix} \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \\ \mathbf{i}(t) \end{pmatrix} + \mathbf{D}\mathbf{u}(t)
\end{aligned} \tag{3.10}$$

where $\mathbf{G}_{11} \in \mathbb{R}^{n_1 \times n_1}$, $\mathbf{G}_{12} \in \mathbb{R}^{n_1 \times n_2}$, $\mathbf{G}_{22} \in \mathbb{R}^{n_2 \times n_2}$, $\mathbf{W}_1 \in \mathbb{R}^{n_1 \times m}$, $\mathbf{W}_2 \in \mathbb{R}^{n_2 \times m}$, $\mathbf{C}_1 \in \mathbb{R}^{n_1 \times n_1}$, $\mathbf{v}_1 \in \mathbb{R}^{n_1}$, $\mathbf{v}_2 \in \mathbb{R}^{n_1}$, $\mathbf{B}_1 \in \mathbb{R}^{n_1 \times p}$, $\mathbf{B}_2 \in \mathbb{R}^{n_2 \times p}$, $\mathbf{L}_1 \in \mathbb{R}^{q \times n_1}$, $\mathbf{L}_2 \in \mathbb{R}^{q \times n_2}$. Assuming now that the submatrix $\mathbf{G}_{22}$ is nonsingular (a sufficient condition for this is at least one resistive connection to ground at the n_2 non-capacitive nodes), the second row of (3.10) can be solved for $\mathbf{v}_2(t)$ as follows:

$$\mathbf{v}_2(t) = \mathbf{G}_{22}^{-1}\mathbf{B}_2\mathbf{u}(t) - \mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T\mathbf{v}_1(t) - \mathbf{G}_{22}^{-1}\mathbf{W}_2\mathbf{i}(t) \tag{3.11}$$

The above can be substituted to the first and third row of (3.10), as well as the output part of (3.10), to give:

$$\begin{aligned}
(\mathbf{G}_{11} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T)\mathbf{v}_1(t) + (\mathbf{W}_1 - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{W}_2)\mathbf{i}(t) \\
+\mathbf{C}_1\dot{\mathbf{v}}_1(t) = (\mathbf{B}_1 - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{B}_2)\mathbf{u}(t)
\end{aligned}$$

$$\begin{aligned}
(\mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T - \mathbf{W}_1^T)\mathbf{v}_1(t) + \mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{W}_2\mathbf{i}(t) + \mathbf{M}\dot{\mathbf{i}}(t) \\
= \mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{B}_2\mathbf{u}(t)
\end{aligned}$$

$$\begin{aligned}
\mathbf{y}(t) = (\mathbf{L}_1 - \mathbf{L}_2\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T)\mathbf{v}_1(t) - \mathbf{L}_2\mathbf{G}_{22}^{-1}\mathbf{W}_2\mathbf{i}(t) \\
+(\mathbf{L}_2\mathbf{G}_{22}^{-1}\mathbf{B}_2 + \mathbf{D})\mathbf{u}(t)
\end{aligned}$$

This can be put together in the following descriptor form:

$$\begin{aligned}
\begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{v}}_1(t) \\ \dot{\mathbf{i}}(t) \end{pmatrix} = \\
- \begin{pmatrix} \mathbf{G_{11}} - \mathbf{G_{12}}\mathbf{G_{22}^{-1}}\mathbf{G_{12}^{\mathrm{T}}} & \mathbf{W}_1 - \mathbf{G_{12}}\mathbf{G_{22}^{-1}}\mathbf{W}_2 \\ \mathbf{W}_2^T\mathbf{G_{22}^{-1}}\mathbf{G}_{12}^T - \mathbf{W}_1^T & \mathbf{W}_2^T\mathbf{G_{22}^{-1}}\mathbf{W}_2 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{i}(t) \end{pmatrix} \\
+ \begin{pmatrix} \mathbf{B}_1 - \mathbf{G_{12}}\mathbf{G_{22}^{-1}}\mathbf{B}_2 \\ \mathbf{W}_2^T\mathbf{G_{22}^{-1}}\mathbf{B}_2 \end{pmatrix} \mathbf{u}(t) \\
\mathbf{y}(t) = \begin{pmatrix} \mathbf{L}_1 - \mathbf{L}_2\mathbf{G_{22}^{-1}}\mathbf{G}_{12}^T & \mathbf{L}_2\mathbf{G_{22}^{-1}}\mathbf{W}_2 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{i}(t) \end{pmatrix} \\
+ (\mathbf{L}_2\mathbf{G_{22}^{-1}}\mathbf{B}_2 + \mathbf{D})\mathbf{u}(t)
\end{aligned} \tag{3.12}$$

The above is a nonsingular (or regular) state-space model which can be reduced normally by the BT method of Algorithm 2.

3.3.3 Balanced Truncation in Limited Frequency Intervals

By employing Parseval's theorem in (3.5) we can obtain expressions for the Gramians in the frequency domain:

$$\begin{aligned}
\mathbf{P} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T}(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-H}d\omega \\
\mathbf{Q} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-H}\mathbf{L}^T\mathbf{L}(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}d\omega
\end{aligned} \tag{3.13}$$

where $(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}$ is the Fourier transform of $exp(\mathbf{E}^{-1}\mathbf{A}t)$.
If now the frequency interval is not taken as the entire $(-\infty, \infty)$ but is restricted to a certain $[-\omega_2, -\omega_1] \cup [\omega_1, \omega_2]$, we obtain the frequency-limited Gramians:

$$\begin{aligned}
\mathbf{P}_\omega = \frac{1}{2\pi}\Big(&\int_{-\omega_2}^{-\omega_1} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T}(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-H}d\omega \\
&+ \int_{\omega_1}^{\omega_2} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T}(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-H}d\omega\Big) \\
\mathbf{Q}_\omega = \frac{1}{2\pi}\Big(&\int_{-\omega_2}^{-\omega_1} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-H}\mathbf{L}^T\mathbf{L}(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}d\omega \\
&+ \int_{\omega_1}^{\omega_2} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-H}\mathbf{L}^T\mathbf{L}(i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}d\omega\Big)
\end{aligned} \tag{3.14}$$

Equivalently, $\mathbf{P}_\omega$ and $\mathbf{Q}_\omega$ can be derived by the solution of the following modified Lyapunov equations [30, 31]:

$$\begin{aligned}
(\mathbf{E}^{-1}\mathbf{A})\mathbf{P}_\omega + \mathbf{P}_\omega(\mathbf{E}^{-1}\mathbf{A})^T \\
= -(\mathbf{F}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T} + (\mathbf{F}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T})^T) \\
(\mathbf{E}^{-1}\mathbf{A})^T\mathbf{Q}_\omega + \mathbf{Q}_\omega(\mathbf{E}^{-1}\mathbf{A}) \\
= -((\mathbf{L}^T\mathbf{L}\mathbf{F})^T + \mathbf{L}^T\mathbf{L}\mathbf{F})
\end{aligned} \tag{3.15}$$

where

$$\mathbf{F} = \frac{1}{2\pi}\Big(\int_{-\omega_2}^{-\omega_1} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}d\omega + \int_{\omega_1}^{\omega_2} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1}d\omega\Big) \tag{3.16}$$

The above matrix integral can be evaluated as:

$$\begin{aligned} \mathbf{F} &= \frac{1}{\pi} Re\left(\int_{\omega_1}^{\omega_2} (i\omega\mathbf{I} - \mathbf{E}^{-1}\mathbf{A})^{-1} d\omega\right) \\ &= Re\left(\frac{i}{\pi} ln\left((\mathbf{E}^{-1}\mathbf{A} + i\omega_1\mathbf{I})^{-1}(\mathbf{E}^{-1}\mathbf{A} + i\omega_2\mathbf{I})\right)\right) \\ &= Re\left(\frac{i}{\pi} ln\left((\mathbf{A} + i\omega_1\mathbf{E})^{-1}(\mathbf{A} + i\omega_2\mathbf{E})\right)\right) \end{aligned} \tag{3.17}$$

where the matrix logarithm $ln(\mathbf{Y})$ is the inverse of the matrix exponential, i.e. a matrix $\mathbf{X}$ such that $exp(\mathbf{X}) = \mathbf{Y}$.

The frequency-limited Gramians characterize the controllability and observability of the model in the selected frequency range, and the process of balancing and truncation will eliminate states that are difficult to reach and observe inside this frequency range. This means that more states can be eliminated for a given tolerance in (3.9) (e.g. states that are easy to reach/observe in other frequencies and which would not have been eliminated otherwise), leading to lower order r in the reduced model, or alternatively to lower error in the frequency range for a given order r.

3.4 Low-rank EKS method for frequency-limited Lyapunov equations

3.4.1 Extended Krylov Subspace Method for Solving Lyapunov Equations

The essence of low-rank projection-type methods for solving the large-scale Lyapunov equations (3.6) is to iteratively project them onto a lower-dimensional subspace, and then solve the resulting small-scale equations to obtain the low-rank approximate solutions of (3.6). The dimension of the projection subspace is increased in every iteration until convergence is achieved. More specifically, if $\mathbf{K} \in \mathbb{R}^{N\times k}$ ($k << N$) is a projection matrix whose columns span the k-dimensional Krylov subspace:

$$\mathcal{K}_k(\mathbf{A}_E, \mathbf{B}_E) = span\{\mathbf{B}_E, \mathbf{A}_E\mathbf{B}_E, \mathbf{A}_E^2\mathbf{B}_E, \ldots, \mathbf{A}_E^{k-1}\mathbf{B}_E\}$$

where

$$\mathbf{A}_E \equiv \mathbf{E}^{-1}\mathbf{A}, \quad \mathbf{B}_E \equiv \mathbf{E}^{-1}\mathbf{B}$$

then the projected Lyapunov equation (for the controllability Gramian $\mathbf{P}$) onto $\mathcal{K}_k(\mathbf{A}_E, \mathbf{B}_E)$ is:

$$(\mathbf{K}^T\mathbf{A}_E\mathbf{K})\mathbf{X} + \mathbf{X}(\mathbf{K}^T\mathbf{A}_E\mathbf{K})^T = -\mathbf{K}^T\mathbf{B}_E\mathbf{B}_E^T\mathbf{K} \tag{3.18}$$

(the same things hold for the observability Gramian $\mathbf{Q}$ with $\mathbf{A}_E^T$, $\mathbf{L}^T$ in place of $\mathbf{A}_E$, $\mathbf{B}_E$). The solution $\mathbf{X} \in \mathbb{R}^{k\times k}$ of (3.18) can be back-projected to the N-dimensional space to give an approximate solution $\mathbf{P} = \mathbf{K}\mathbf{X}\mathbf{K}^T$ for the original large-scale equation (3.6), and a low-rank factor $\mathbf{Z} \in \mathbb{R}^{N\times k}$ of $\mathbf{P}$ can be obtained as $\mathbf{Z} = \mathbf{K}\mathbf{S}$ where $\mathbf{X} = \mathbf{S}\mathbf{S}^T$ is the Cholesky factorization of $\mathbf{X}$.

The projection process is independent of the subspace selection, but its effectiveness is critically dependent on the chosen subspace and can sometimes take many iterations of subspace updating before converging to the final solution. The convergence can be accelerated by enriching the standard Krylov subspace $\mathcal{K}_k(\mathbf{A}_E, \mathbf{B}_E)$ with information from the subspace $\mathcal{K}_k(\mathbf{A}_E^{-1}, \mathbf{B}_E)$ corresponding to the inverse matrix $\mathbf{A}_E^{-1}$, leading to the extended Krylov subspace:

$$\mathcal{K}_k^E(\mathbf{A}_E, \mathbf{B}_E) = \mathcal{K}_k(\mathbf{A}_E, \mathbf{B}_E) + \mathcal{K}_k(\mathbf{A}_E^{-1}, \mathbf{B}_E) =$$

$$span\{\mathbf{B}_E, \mathbf{A}_E^{-1}\mathbf{B}_E, \mathbf{A}_E\mathbf{B}_E, \mathbf{A}_E^{-2}\mathbf{B}_E, \mathbf{A}_E^{2}\mathbf{B}_E, \ldots, \mathbf{A}_E^{-(k-1)}\mathbf{B}_E, \mathbf{A}_E^{k-1}\mathbf{B}_E\} \tag{3.19}$$

The extended Krylov subspace (EKS) method starts by the pair $\{\mathbf{B}_E, \mathbf{A}_E^{-1}\mathbf{B}_E\}$ and generates a sequence of extended subspaces $\mathcal{K}_k^E(\mathbf{A}_E, \mathbf{B}_E)$ of increasing dimensions, solving the projected Lyapunov equation (3.18) in each iteration, until a sufficiently accurate approximation of the solution of (3.6) is obtained. The complete EKS method is given in Algorithm 3.

Algorithm 3 Extended Krylov Subspace method for low-rank solution of Lyapunov equations that arise in descriptor systems

Input: $\mathbf{A}_E \equiv \mathbf{E}^{-1}\mathbf{A}, \mathbf{B}_E \equiv \mathbf{E}^{-1}\mathbf{B}$ (or $\mathbf{A}_E^T, \mathbf{L}^T$)
Output: $\mathbf{Z}$ such that $\mathbf{P} \approx \mathbf{Z}\mathbf{Z}^T$,

1: $j = 1$; $p = size\_col(\mathbf{B}_E)$;
2: $\mathbf{K}^{(j)} = Orth([\mathbf{B}_E, \mathbf{A}_E^{-1}\mathbf{B}_E])$
3: **repeat**
4: $\quad \mathbf{M} = \mathbf{K}^{(j)T}\mathbf{A}_E\mathbf{K}^{(j)}$; $\quad \mathbf{R} = \mathbf{K}^{(j)T}\mathbf{B}_E$
5: $\quad$ Solve $\mathbf{M}\mathbf{X} + \mathbf{X}\mathbf{M}^T = -\mathbf{R}\mathbf{R}^T$ for $\mathbf{X} \in \mathcal{R}^{2pj \times 2pj}$
6: $\quad k_1 = 2p(j-1)$; $k_2 = k_1 + p$; $k_3 = 2pj$
7: $\quad \mathbf{K}_1 = [\mathbf{A}_E\mathbf{K}^{(j)}(:, k_1+1 : k_2), \mathbf{A}_E^{-1}\mathbf{K}^{(j)}(:, k_2+1 : k_3)]$
8: $\quad \mathbf{K}_2 = Orth(\mathbf{K}_1)$ $\quad$ w.r.t $\quad \mathbf{K}^{(j)}$
9: $\quad \mathbf{K}_3 = Orth(\mathbf{K}_2)$
10: $\quad \mathbf{K}^{(j+1)} = [\mathbf{K}^{(j)}, \mathbf{K}_3]$
11: $\quad j = j + 1$
12: **until** convergence
13: $\mathbf{S} = Chol(\mathbf{X})$
14: $\mathbf{Z} = \mathbf{K}^{(j)}\mathbf{S}$

Some details on the efficient implementation of the EKS method of Algorithm 3 are as follows:

Sparse matrix inputs

The inputs to Algorithm 3 are not actually $\mathbf{A}_E \equiv \mathbf{E}^{-1}\mathbf{A}$ or $\mathbf{A}_E^T \equiv (\mathbf{E}^{-1}\mathbf{A})^T$ but the sparse system matrices $\mathbf{A}$, $\mathbf{E}$ or $\mathbf{A}^T$, $\mathbf{E}^T$, since the (generally dense) inverse matrices are only needed in products with p vectors (initially and in step 2) and $2pj$ vectors (in steps 4 and 7 at every iteration, where the iteration count j is typically very small and thus $2pj << N$). These can be implemented as sparse linear solves $\mathbf{E}\mathbf{Y} = \mathbf{R}$ and $\mathbf{A}\mathbf{Y} = \mathbf{R}$ (or $\mathbf{E}^T\mathbf{Y} = \mathbf{R}$, $\mathbf{A}^T\mathbf{Y} = \mathbf{R}$) by any sparse direct or iterative algorithm like [44] or [45].

Orthogonalization in steps 2 and 9

A modified Gram-Schmidt procedure [46] is employed to implement the corresponding ($Orth()$) procedures.

Solution of small-scale Lyapunov equation in step 5

A direct Schur decomposition method [47] can be employed for the solution of the small-scale ($2pj \times 2pj$) Lyapunov equation in each iteration of Algorithm 3.

Orthogonalization in step 8

In order to perform orthogonalization with respect to matrix $\mathbf{K}^{(j)}$ we employ the following Gram-Schmidt procedure [46]:

for $k_1 = 1, \ldots, j$ **do**
  $k_2 = (k_1 - 1) * 2p$; $k_3 = k_1 * 2p$;
  $\mathbf{T} = \mathbf{K}^{(j)T}((:, k_2 + 1 : k_3))\mathbf{K}_1$
  $\mathbf{K}_2 = \mathbf{K}_1 - \mathbf{K}^{(j)}(:, k_2 + 1 : k_3)\mathbf{T}$
end for

Convergence criterion

An appropriate stopping criterion is the residual of (3.6) with the approximate solution $\mathbf{P} = \mathbf{K}\mathbf{X}\mathbf{K}^T$ to reach a certain threshold in magnitude, i.e.

$$\frac{||\mathbf{A}_E\mathbf{K}^{(j)}\mathbf{X}\mathbf{K}^{(j)T} + \mathbf{K}^{(j)}\mathbf{X}\mathbf{K}^{(j)T}\mathbf{A}_E + \mathbf{B}_E\mathbf{B}_E^T||}{||\mathbf{B}_E\mathbf{B}_E^T||} \leq tol \tag{3.20}$$

A tolerance of $tol = 10^{-10}$ is typically sufficient in practice to acquire a good approximation of the solution.

Lower rank solution

The solution $\mathbf{Z}$ obtained after the termination of Algorithm 3 has rank $2pj$, where j is the final iteration count. This can be reduced even further by employing in step 13 the eigendecomposition $\mathbf{X} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^T$ instead of the Cholesky factorization $\mathbf{X} = \mathbf{S}\mathbf{S}^T$, for the solution $\mathbf{X}$ of the final projected Lyapunov equation. By keeping only the k eigenvalues above a certain threshold (a fair choice of threshold is 10^{-12}), along with the corresponding eigenvectors, a factor $\mathbf{Z}$ of $\mathbf{P}$ with lower rank $k < 2pj$ can be obtained as $\mathbf{Z} = \mathbf{K}\mathbf{W}_{(2pj\times k)}\mathbf{\Lambda}^{\frac{1}{2}}_{(k\times k)}$.

3.4.2 Application of EKS Method to Frequency-Limited Lyapunov Equations

The frequency-limited Lyapunov equations (3.15) differ from the standard Lyapunov equations (3.6) in the right-hand-sides (RHS) which do not have the forms $-\mathbf{B}_E\mathbf{B}_E^T$ (with $\mathbf{B}_E \equiv \mathbf{E}^{-1}\mathbf{B}$) or $-\mathbf{L}^T\mathbf{L}$. The EKS algorithm presented in [31] modified the step 5 of Algorithm 3 to solve a small-scale frequency-limited Lyapunov equation, which is a projected version of the large-scale equations (3.15) onto a lower-dimensional subspace. However, the projection subspace is still the extended Krylov subspace (3.19) related to the RHS of the standard Lyapunov equations, which can render the projection procedure ineffective. Here, we show that it is possible to write the RHS of (3.15) in a factorized form similar to $-\mathbf{B}_E\mathbf{B}_E^T$ or $-\mathbf{L}^T\mathbf{L}$, so that Algorithm 3 can be invoked with the corresponding factor and create the proper Krylov subspace. Specifically, observe that the RHS of (3.15) can be written as:

$$\begin{aligned}
-(\mathbf{F}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T} + (\mathbf{F}\mathbf{E}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{E}^{-T})^T) &= \\
-\begin{pmatrix}\mathbf{E}^{-1}\mathbf{B} & \mathbf{F}\mathbf{E}^{-1}\mathbf{B}\end{pmatrix}\begin{pmatrix}\mathbf{0}_p & \mathbf{I}_p \\ \mathbf{I}_p & \mathbf{0}_p\end{pmatrix}\begin{pmatrix}(\mathbf{E}^{-1}\mathbf{B})^T \\ (\mathbf{F}\mathbf{E}^{-1}\mathbf{B})^T\end{pmatrix} & \\
-((\mathbf{L}^T\mathbf{L}\mathbf{F})^T + \mathbf{L}^T\mathbf{L}\mathbf{F}) &= \\
-\begin{pmatrix}\mathbf{L}^T & (\mathbf{L}\mathbf{F})^T\end{pmatrix}\begin{pmatrix}\mathbf{0}_q & \mathbf{I}_q \\ \mathbf{I}_q & \mathbf{0}_q\end{pmatrix}\begin{pmatrix}\mathbf{L} \\ \mathbf{L}\mathbf{F}\end{pmatrix} &
\end{aligned}$$

(where $\mathbf{I}_p$, $\mathbf{I}_q$,$\mathbf{0}_p$, $\mathbf{0}_q$ are the $p \times p$ and $q \times q$ identity and zero matrices respectively), i.e. in factorized $-\mathbf{B}_\omega \mathbf{J} \mathbf{B}_\omega^T$ and $-\mathbf{L}_\omega^T \mathbf{J} \mathbf{L}_\omega$ forms with

$$\begin{aligned} \mathbf{B}_\omega &\equiv \begin{pmatrix} \mathbf{E}^{-1}\mathbf{B} & \mathbf{F}\mathbf{E}^{-1}\mathbf{B} \end{pmatrix} \\ \mathbf{L}_\omega^T &\equiv \begin{pmatrix} \mathbf{L}^T & (\mathbf{L}\mathbf{F})^T \end{pmatrix} \\ \mathbf{J} &\equiv \begin{pmatrix} \mathbf{0}_p & \mathbf{I}_p \\ \mathbf{I}_p & \mathbf{0}_p \end{pmatrix} \end{aligned} \tag{3.21}$$

By entering $\mathbf{B}_\omega \in \mathbb{R}^{n \times 2p}$ or $\mathbf{L}_\omega^T \in \mathbb{R}^{n \times 2q}$ instead of $\mathbf{B}_E \in \mathbb{R}^{n \times p}$ or $\mathbf{L}^T \in \mathbb{R}^{n \times q}$ in Algorithm 3, the proper Krylov subspaces related to the actual RHS of (3.15) are created, and the only modification required is the RHS of the small-scale equation in step 5 to become $-\mathbf{R}\mathbf{J}\mathbf{R}^T$. Note that the Krylov subspace spanned by the columns of projection matrix $\mathbf{K}$ is not affected by the presence of matrix $\mathbf{J}$, since $\mathbf{J}$ is symmetric and thus diagonal up to an orthogonal similarity transformation [48,49] - specifically $\begin{pmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix}$.

To construct the $\mathbf{B}_\omega$ and $\mathbf{L}_\omega^T$ inputs of Algorithm 3 we need to compute the matrix logarithm (3.17), which is generally an intensive computational procedure. However, the matrix logarithms are not explicitly needed in (3.21), but only their effects in pre-multiplying p vectors ($ln(\mathbf{Y})\mathbf{E}^{-1}\mathbf{B}$) or post-multiplying q vectors ($\mathbf{L}ln(\mathbf{Y}) = (ln(\mathbf{Y})^H \mathbf{L}^T)^H$) are required. These can be efficiently computed by a similar EKS algorithm for evaluating matrix functions [51], which iteratively creates an extended Krylov subspace related to the matrix input and the multiplied vectors:

$$\begin{aligned} \mathcal{K}_k^E(\mathbf{Y},\mathbf{B}) = \mathcal{K}_k(\mathbf{Y},\mathbf{B}) + \mathcal{K}_k(\mathbf{Y}^{-1},\mathbf{B}) = \\ span\{\mathbf{B}, \mathbf{Y}^{-1}\mathbf{B}, \mathbf{Y}\mathbf{B}, \mathbf{Y}^{-2}\mathbf{B}, \mathbf{Y}^2\mathbf{B}, \ldots, \mathbf{Y}^{-(k-1)}\mathbf{B}, \mathbf{Y}^{k-1}\mathbf{B}\} = \\ span\{\mathbf{B}, (\mathbf{A}+i\omega_2\mathbf{E})^{-1}(\mathbf{A}+i\omega_1\mathbf{E})\mathbf{B}, \ldots, \\ (\mathbf{A}+i\omega_1\mathbf{E})^{-k+1}(\mathbf{A}+i\omega_2\mathbf{E})^{k-1}\mathbf{B}\} \end{aligned}$$

and whose implementation is shown in Algorithm 4.
Note that in steps 4 and 6 the operations $\mathbf{Y}\mathbf{K}^{(j)}$ and $\mathbf{Y}^{-1}\mathbf{K}^{(j)}$ can be written as:

$$\begin{aligned} \mathbf{Y}\mathbf{K}^{(j)} = (\mathbf{A}+i\omega_1\mathbf{E})^{-1}(\mathbf{A}+i\omega_2\mathbf{E})\mathbf{K}^{(j)} = \\ (\mathbf{A}+i\omega_1\mathbf{E})^{-1}(\mathbf{A}+i\omega_1\mathbf{E}-i\omega_1\mathbf{E}+i\omega_2\mathbf{E})\mathbf{K}^{(j)} = \\ \mathbf{K}^{(j)} + i(\omega_2-\omega_1)(\mathbf{A}+i\omega_1\mathbf{E})^{-1}\mathbf{E}\mathbf{K}^{(j)} \end{aligned}$$

$$\begin{aligned} \mathbf{Y}^{-1}\mathbf{K}^{(j)} = (\mathbf{A}+i\omega_2\mathbf{E})^{-1}(\mathbf{A}+i\omega_1\mathbf{E})\mathbf{K}^{(j)} = \\ \mathbf{K}^{(j)} + i(\omega_1-\omega_2)(\mathbf{A}+i\omega_2\mathbf{E})^{-1}\mathbf{E}\mathbf{K}^{(j)} \end{aligned}$$

which involve a sparse product of $2pj$ vectors with $\mathbf{E}$ followed by a complex sparse linear solve with $\mathbf{A}+i\omega_1\mathbf{E}$ or $\mathbf{A}+i\omega_2\mathbf{E}$. Also, a standard inverse scaling and squaring algorithm [47] is used for the small-scale computation of matrix logarithm times p vectors in step 5.

3.4.3 Sparse Implementation for Singular Descriptor Models

For the reduction in limited frequency intervals of the model (3.12) that results from the regularization of a singular descriptor model, the execution of Algorithms 3 and 4 is computationally inefficient because the inversion of $\mathbf{G}_{22}$ renders the matrices dense and hinders

Algorithm 4 Extended Krylov subspace method for matrix logarithm multiplying p vectors $\mathbf{B}$

Input: $\mathbf{Y}, \mathbf{B}$, Convergence tolerance ϵ
Output: $\mathbf{v} = ln(\mathbf{Y})\mathbf{B}$

1: $j = 1$;
2: $\mathbf{K}^{(j)} = Orth([\mathbf{B}, \mathbf{Y}^{-1}\mathbf{B}])$;
3: **repeat**
4: $\quad \mathbf{X} = \mathbf{K}^{(j)T}\mathbf{Y}\mathbf{K}^{(j)}; \quad \mathbf{R} = \mathbf{K}^{(j)T}\mathbf{B}$
5: $\quad$ Compute $\mathbf{y} = ln(\mathbf{X})\mathbf{R}$
6: $\quad k_1 = 2p(j-1); k_2 = k_1 + p; k_3 = 2pj$
7: $\quad \mathbf{K}_1 = [\mathbf{Y}\mathbf{K}^{(j)}(:, k_1+1 : k_2); \mathbf{Y}^{-1}\mathbf{K}^{(j)}(:, k_2+1 : k_3)]$
8: $\quad \mathbf{K}_2 = Orth(\mathbf{K}_1) \quad$ w.r.t $\quad \mathbf{K}^{(j)}$
9: $\quad \mathbf{K}_3 = Orth(\mathbf{K}_2)$
10: $\quad \mathbf{K}^{(j+1)} = [\mathbf{K}^{(j)}, \mathbf{K}_3]$
11: $\quad j = j + 1$
12: **until** $\frac{||\mathbf{y}^{(j)} - \mathbf{y}^{(j-1)}||}{||\mathbf{y}^{(j)}||} < \epsilon$
13: $\mathbf{v} = \mathbf{K}^{(j)}\mathbf{y}$

the solution procedure. In this section we present efficient ways to implement the EKS algorithms by keeping the system matrices in their original sparse forms.

Construction of RHS

The input-to-state and state-to-output connectivity matrices

$$\mathbf{B} \equiv \begin{pmatrix} \mathbf{B}_1 - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{B}_2 \\ \mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{B}_2 \end{pmatrix}, \quad \mathbf{L}^T \equiv \begin{pmatrix} \mathbf{L}_1^T - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{L}_2^T \\ \mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{L}_2^T \end{pmatrix} \tag{3.22}$$

are constructed explicitly to compute the $\mathbf{B}_\omega$ and $\mathbf{L}_\omega^T$ inputs of Algorithm 3 from (3.21), where the products $\mathbf{G}_{22}^{-1}\mathbf{B}_2$ and $\mathbf{G}_{22}^{-1}\mathbf{L}_2^T$ are computed by p and q sparse linear solves respectively.

Sparse linear system solutions

The system matrix

$$\mathbf{A} \equiv -\begin{pmatrix} \mathbf{G}_{11} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^{\mathbf{T}} & \mathbf{W}_1 - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{W}_2 \\ \mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T - \mathbf{W}_1^T & \mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{W}_2 \end{pmatrix} \tag{3.23}$$

of model (3.12) is rendered dense due to the inversion of $\mathbf{G}_{22}$. The linear system solutions with $\mathbf{A}$ (or $\mathbf{A}^T$), $\mathbf{E}$, and $\mathbf{A} + i\omega\mathbf{E}$ in steps 2, 4, 7 of Algorithms 3 and 4 can be handled by partitioning the RHS of these systems conformally to $\mathbf{A}$, i.e. $\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix}$ with $\mathbf{R}_1 \in \mathbb{R}^{n_1 \times 2pj}$, $\mathbf{R}_2 \in \mathbb{R}^{m \times 2pj}$ (or $\mathbf{R}_1 \in \mathbb{R}^{n_1 \times 2qj}$, $\mathbf{R}_2 \in \mathbb{R}^{m \times 2qj}$), and implementing their solution efficiently

by keeping sub-blocks in their original sparse form as follows:

$$\begin{gathered}
\mathbf{AX}=\mathbf{R} \Longrightarrow \\
\begin{pmatrix} -\mathbf{G}_{11} & -\mathbf{W}_1 & -\mathbf{G}_{12} \\ \mathbf{W}_1^T & \mathbf{0} & \mathbf{W}_2^T \\ -\mathbf{G}_{12}^T & -\mathbf{W}_2 & -\mathbf{G}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{0} \end{pmatrix} \\
\mathbf{EX}=\mathbf{R} \Longrightarrow \\
\begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ -\mathbf{0} & \mathbf{M} \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix} \\
(\mathbf{A}+i\omega\mathbf{E})\mathbf{X}=\mathbf{R} \Longrightarrow \\
\begin{pmatrix} -\mathbf{G}_{11}+i\omega\mathbf{C}_1 & -\mathbf{W}_1 & -\mathbf{G}_{12} \\ \mathbf{W}_1^T & i\omega\mathbf{M} & \mathbf{W}_2^T \\ -\mathbf{G}_{12}^T & -\mathbf{W}_2 & -\mathbf{G}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{0} \end{pmatrix}
\end{gathered} \tag{3.24}$$

where $\mathbf{Y} \in \mathbb{R}^{n_2 \times 2pj}$ (or $\mathbf{Y} \in \mathbb{R}^{n_2 \times 2qj}$) is a temporary sub-matrix.
A comparable approach for the reduction of power systems via the low-rank ADI method was presented in [52], but is more complicated computationally and is applicable to RC-only circuits.

Sparse matrix-vector products

The matrix-vector products with $\mathbf{K}^{(j)}$ in steps 4 and 7 of Algorithm 3 can be implemented efficiently by observing that:

$$\begin{aligned}
\mathbf{A} &= \begin{pmatrix} -\mathbf{G}_{11} & -\mathbf{W}_1 \\ \mathbf{W}_1^T & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T & \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{W}_2 \\ -\mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^T & -\mathbf{W}_2^T\mathbf{G}_{22}^{-1}\mathbf{W}_2 \end{pmatrix} \\
&= \begin{pmatrix} -\mathbf{G}_{11} & -\mathbf{W}_1 \\ \mathbf{W}_1^T & \mathbf{0} \end{pmatrix} + \begin{pmatrix} -\mathbf{G}_{12} \\ \mathbf{W}_2^T \end{pmatrix} \mathbf{G}_{22}^{-1} \begin{pmatrix} -\mathbf{G}_{12}^T & -\mathbf{W}_2 \end{pmatrix}
\end{aligned} \tag{3.25}$$

Thus the product $\mathbf{AK}^{(j)}$ with the $2pj$ vectors $\mathbf{K}^{(j)}$ can be carried out by a sparse solve $\mathbf{G}_{22}\mathbf{X} = \begin{pmatrix} -\mathbf{G}_{12}^T & -\mathbf{W}_2 \end{pmatrix} \mathbf{K}^{(j)}$, followed by a sum of products $\begin{pmatrix} -\mathbf{G}_{11} & -\mathbf{W}_1 \\ \mathbf{W}_1^T & \mathbf{0} \end{pmatrix} \mathbf{K}^{(j)} + \begin{pmatrix} -\mathbf{G}_{12} \\ \mathbf{W}_2^T \end{pmatrix} \mathbf{X}$.

Construction of system matrix

The dense system matrix (3.23) to be reduced needs sparse solves in the submatrix $\mathbf{G}_{22}$ for its construction. Since it is usually $n_2 << n_1$, it is better to compute first the left-solves $\mathbf{G}_{12}\mathbf{G}_{22}^{-1}$ and $\mathbf{W}_2^T\mathbf{G}_{22}^{-1}$, followed by products with $\mathbf{G}_{12}^T$ and $\mathbf{W}_2$. The left-solves can be performed as $\mathbf{G}_{22}\mathbf{X} = \mathbf{G}_{12}$ and $\mathbf{G}_{22}\mathbf{X} = \mathbf{W}_2^T$ where $\mathbf{X}$ contains the *rows* of each left-solve.

3.5 Modified ADI method for solving frequency-limited Lyapunov equations

Besides the Lyapunov equation of (3.6) one can work on the following *generalized* Lyapunov matrix equations [62]:

$$\begin{aligned} \mathbf{A}\mathbf{P}\mathbf{E}^T + \mathbf{E}\mathbf{P}\mathbf{A}^T &= -\mathbf{B}\mathbf{B}^T, \\ \mathbf{A}^T\mathbf{Q}'\mathbf{E} + \mathbf{E}^T\mathbf{Q}'\mathbf{A} &= -\mathbf{L}^T\mathbf{L} \end{aligned} \tag{3.26}$$

where a post-processing step is needed to obtain the observability Gramian as

$$\mathbf{Q} = \mathbf{E}^T\mathbf{Q}'\mathbf{E}$$

Equivalently to the *generalized* Lyapunov equations, $\mathbf{P}_\omega$ and $\mathbf{Q}_\omega$ can be derived by the solution of the following modified *generalized* Lyapunov equations [30, 31]:

$$\begin{aligned} \mathbf{A}\mathbf{P}_\omega\mathbf{E}^T + \mathbf{E}\mathbf{P}_\omega\mathbf{A}^T &= -(\mathbf{E}\mathbf{F}\mathbf{B}\mathbf{B}^T + (\mathbf{E}\mathbf{F}\mathbf{B}\mathbf{B}^T)^T) \\ \mathbf{A}^T\mathbf{Q}'_\omega\mathbf{E} + \mathbf{E}^T\mathbf{Q}'_\omega\mathbf{A} &= -((\mathbf{L}^T\mathbf{L}\mathbf{F}\mathbf{E})^T + \mathbf{L}^T\mathbf{L}\mathbf{F}\mathbf{E}) \end{aligned} \tag{3.27}$$

where a similar post-processing step is needed to obtain the observability Gramian as

$$\mathbf{Q}_\omega = \mathbf{E}^T\mathbf{Q}'_\omega\mathbf{E}$$

For the modified *generalized* Lyapunov equations (3.27) we have to deal again with the different RHS, which is in the forms $-(\mathbf{B}_\omega\mathbf{B}^T + \mathbf{B}\mathbf{B}_\omega^T)$ and $-(\mathbf{C}_\omega^T\mathbf{C} + \mathbf{L}^T\mathbf{L}_\omega)$ (where $\mathbf{B}_\omega \equiv \mathbf{E}\mathbf{F}\mathbf{B}$ and $\mathbf{L}_\omega \equiv \mathbf{L}\mathbf{F}\mathbf{E}$) instead of the standard forms $-\mathbf{B}\mathbf{B}^T$ and $-\mathbf{L}^T\mathbf{L}$ of (3.6). This requires the modification of the standard ADI method presented in [26], which we describe in this section.

Observe that the RHS of the frequency-limited *generalized* Lyapunov equations can be similarly written as:

$$\begin{aligned} -(\mathbf{B}_\omega\mathbf{B}^T + \mathbf{B}\mathbf{B}_\omega^T) &= -\begin{pmatrix}\mathbf{B} & \mathbf{B}_\omega\end{pmatrix}\begin{pmatrix}\mathbf{0}_p & \mathbf{I}_p \\ \mathbf{I}_p & \mathbf{0}_p\end{pmatrix}\begin{pmatrix}\mathbf{B}^T \\ \mathbf{B}_\omega^T\end{pmatrix} \\ -(\mathbf{L}_\omega^T\mathbf{L} + \mathbf{L}^T\mathbf{L}_\omega) &= -\begin{pmatrix}\mathbf{L}^T & \mathbf{L}_\omega^T\end{pmatrix}\begin{pmatrix}\mathbf{0}_q & \mathbf{I}_q \\ \mathbf{I}_q & \mathbf{0}_q\end{pmatrix}\begin{pmatrix}\mathbf{L} \\ \mathbf{L}_\omega\end{pmatrix} \end{aligned} \tag{3.28}$$

(where $\mathbf{I}_p$ and $\mathbf{I}_q$ are the $p \times p$ and $q \times q$ identity matrices and $\mathbf{0}_p$, $\mathbf{0}_q$ are the corresponding zero matrices).

This permits the use of an $\mathbf{L}\mathbf{D}\mathbf{L}^T$ variant of ADI proposed in [48], which expects the RHS to be in the form of $-\mathbf{S}\mathbf{R}\mathbf{S}^T$ and gives the solution of Lyapunov equations in the form $\mathbf{L}\mathbf{D}\mathbf{L}^T$, where $\mathbf{L}$ is the low rank factor and $\mathbf{D}$ is a block diagonal matrix. This variant of ADI is given in Algorithm 3 and we can use it to solve (3.15) by entering respectively:

$$\begin{aligned} \mathbf{S} \equiv \begin{pmatrix}\mathbf{B} & \mathbf{B}_\omega\end{pmatrix} = \begin{pmatrix}\mathbf{B} & \mathbf{E}\mathbf{F}\mathbf{B}\end{pmatrix}, \quad \mathbf{R} \equiv \begin{pmatrix}\mathbf{0}_p & \mathbf{I}_p \\ \mathbf{I}_p & \mathbf{0}_p\end{pmatrix} \\ \mathbf{S} \equiv \begin{pmatrix}\mathbf{L}^T & \mathbf{L}_\omega^T\end{pmatrix} = \begin{pmatrix}\mathbf{L}^T & (\mathbf{L}\mathbf{F}\mathbf{E})^T\end{pmatrix}, \quad \mathbf{R} \equiv \begin{pmatrix}\mathbf{0}_q & \mathbf{I}_q \\ \mathbf{I}_q & \mathbf{0}_q\end{pmatrix} \end{aligned} \tag{3.29}$$

Since the output of Algorithm 5 is obtained in the form of $\mathbf{L}\mathbf{D}\mathbf{L}^T$, we must transform it in the form of $\mathbf{Z}\mathbf{Z}^T$ to be able to use it in the BT Algorithm 2. To do that, observe that with $\mathbf{I} \equiv \mathbf{I}_p$ or $\mathbf{I}_q$ the matrix $\mathbf{R}$ of (3.29) can be written as:

$$2\mathbf{R} = 2\begin{pmatrix}\mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0}\end{pmatrix} = \begin{pmatrix}\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{I} \\ -\mathbf{I} & \mathbf{I}\end{pmatrix}$$

and thus the block diagonal matrix $\mathbf{D} = -2\,blkdiag(Re(\mu_1)\mathbf{R}, \ldots, Re(\mu_j)\mathbf{R}))$ can be written in factorized form as $\mathbf{D} = \mathbf{M}\mathbf{M}^T$ where:

$$\mathbf{M} = \begin{pmatrix} \sqrt{-Re(\mu_1)}\begin{pmatrix}\mathbf{I} & -\mathbf{I}\\ \mathbf{I} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & i\mathbf{I}\end{pmatrix} & & \mathbf{0} & \\ & \ddots & & \\ \mathbf{0} & & \sqrt{-Re(\mu_j)}\begin{pmatrix}\mathbf{I} & -\mathbf{I}\\ \mathbf{I} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & i\mathbf{I}\end{pmatrix} \end{pmatrix}$$
$$\mathbf{M}^T = \begin{pmatrix} \sqrt{-Re(\mu_1)}\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & i\mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{I}\\ -\mathbf{I} & \mathbf{I}\end{pmatrix} & & \mathbf{0} & \\ & \ddots & & \\ \mathbf{0} & & \sqrt{-Re(\mu_j)}\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & i\mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{I}\\ -\mathbf{I} & \mathbf{I}\end{pmatrix} \end{pmatrix} \tag{3.30}$$

(with i being the imaginary unit). In this way, we can write the solution of the frequency-limited Lyapunov equations (3.15) in the form $\mathbf{LDL}^T = \mathbf{LM}(\mathbf{LM})^T$, and employ them normally in Algorithm 2.

The ADI method depends crucially on the choice of shift parameters μ_j (complex numbers in general), for which an efficient adaptive strategy has been proposed in [50]. The ADI algorithm also involves the solution of p (or q) sparse linear systems in the matrix $(\mathbf{A} + \mu_j\mathbf{E})$ in each iteration j, which can be solved by any sparse direct or iterative method.

Algorithm 5 ADI method for solving Lyapunov equations of descriptor systems in low-rank $\mathbf{LDL}^T$ form

Input: $\mathbf{E}, \mathbf{A}$ (or $\mathbf{E}^T, \mathbf{A}^T$), $\mathbf{S}$, $\mathbf{R}$, ADI-shifts $\mu_1, \mu_2, \ldots$, Convergence tolerance ϵ
Output: $\mathbf{L}$, $\mathbf{D}$ such that $\mathbf{P} \approx \mathbf{LDL}^T$ (or $\mathbf{Q} \approx \mathbf{LDL}^T$)

$\mathbf{L} = []$; $\mathbf{K}^{(0)} = \mathbf{S}$; $j = 1$
while $||\mathbf{K}^{(j-1)}\mathbf{R}\mathbf{K}^{(j-1)T}|| \geq \epsilon||\mathbf{S}\mathbf{R}\mathbf{S}^T||$ **do**
  if $j = 1$ **then**
    Solve $(\mathbf{A} + \mu_j\mathbf{E})\mathbf{X} = \mathbf{K}^{(j-1)}$ for $\mathbf{X}$
  else
    Solve $(\mathbf{A} + \mu_j\mathbf{E})\mathbf{X} = \mathbf{E}\mathbf{K}^{(j-1)}$ for $\mathbf{X}$
  end if
  if $\mu_j \in \mathbb{R}$ **then**
    $\mathbf{K}^{(j)} = \mathbf{K}^{(j-1)} - 2\mu_j\mathbf{X}$
    $\mathbf{L} = [\mathbf{L}; \mathbf{X}]$
  else
    $\eta = \sqrt{2}$; $\delta_j = Re(\mu_j)/Im(\mu_j)$;
    $\mathbf{K}^{(j+1)} = \mathbf{K}^{(j-1)} - 4Re(\mu_j)(Re(\mathbf{X}) + \delta_j Im(\mathbf{X}))$
    $\mathbf{L} = [\mathbf{L}; \eta(Re(\mathbf{X}) + \delta_j Im(\mathbf{X})); \eta\sqrt{\delta_j^2 + 1}Im(\mathbf{X})]$
    $j = j + 1$
  end if
  $j = j + 1$
end while
$\mathbf{D} = -2\, blkdiag(Re(\mu_1)\mathbf{R}, \ldots, Re(\mu_j)\mathbf{R}))$

3.6 Complete procedure with the EKS method

For given model matrices $\mathbf{E}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{L}$ and given frequencies ω_1, ω_2, the complete procedure for frequency-limited BT of large-scale models is as follows:

- Evaluate the quantities $\mathbf{F}\mathbf{E}^{-1}\mathbf{B}$ and $\mathbf{L}\mathbf{F}$ of (3.21), with $\mathbf{F}$ being the matrix logarithm of (3.17), through the EKS method of Algorithm 4.

- Compute the low-rank factors $\mathbf{Z}_P$ and $\mathbf{Z}_Q$ of the frequency-limited Gramians through the EKS method of Algorithm 3, by inserting the matrices $\mathbf{B}_\omega$, $\mathbf{L}_\omega^T$ of (3.21) instead of $\mathbf{B}_E$, $\mathbf{L}^T$.
- Execute steps 2 and 3 of the BT Algorithm 2.

It must be noted that there is no guarantee that the reduced-order models by frequency-limited BT preserve passivity or stability. However, the focus of MOR in recent years has been shifted from provably passive models to passivity enforcement *after* efficient reduction. A wealth of passivity enforcement techniques such as [53] have been developed, which can be used to assure passivity (and also stability) of the reduced-order models obtained by frequency-limited BT.

3.7 Complete procedure with the ADI method

For given model matrices $\mathbf{E}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{L}$ and given frequencies ω_1, ω_2, the complete procedure for frequency-limited BT of large-scale models is as follows:

- Evaluate the $\mathbf{B}_\omega$ and $\mathbf{C}_\omega$ parts of the RHS of frequency-limited Lyapunov equations from (3.21), through the extended Krylov subspace method of Algorithm 4.
- Solve the frequency-limited Lyapunov equations (3.15) in low-rank $\mathbf{LDL}^T$ form, through the modified ADI method of Algorithm 5 by inserting the matrices $\mathbf{S}$ and $\mathbf{R}$ of (3.29).
- Compute the low-rank factors of the frequency-limited Gramians $\mathbf{P}_\omega$ and $\mathbf{Q}_\omega$ as $\mathbf{Z}_P = \mathbf{LM}$ and $\mathbf{Z}_Q = \mathbf{E}^T\mathbf{LM}$, where $\mathbf{L}$ is the output of Algorithm 3 and $\mathbf{M}$ is the matrix of (3.30), and then execute steps 2 and 3 of the BT Algorithm 2.

3.8 Experimental Results

TABLE 3.1: Reduction results of frequency-limited BT vs standard BT for various circuit benchmarks.

Ckt	#nodes	#ports	Standard BT			Frequency-limited BT			
			ROM order	Max error	Times (s)	ROM order for same error	Reduction percentage	Max error for same order	Times(s)
MNA_1	578	9	102	6.4142e-04	0.21	85	16.66%	2.7648e-09	0.57
MNA_2	9223	18	480	3.4431e-04	404.89	411	14.37%	7.5592e-06	437.99
MNA_3	4863	22	415	1.3421e-04	136.49	368	11.32%	1.8127e-07	213.64
MNA_4	980	4	122	5.3913e-04	1.24	100	18.03%	7.5183e-11	1.44
MNA_5	10913	9	135	8.5454e-06	947.36	102	24.44%	1.5469e-11	1047.26
TL	3253	22	140	9.5354e-05	10.50	103	26.42%	2.4554e-11	9.34
LNA	29885	79	432	6.4324e-05	2041.84	401	7.17%	1.5314e-10	2171.70
MX3	867	110	133	1.1610e-05	1.77	123	7.51%	8.4520e-10	1.85
MX7	133	66	70	2.5481e-07	0.13	66	5.71%	1.2960e-11	0.15
IS	16862	646	1268	7.9140e-04	1295.44	1136	10.41%	4.6970e-08	1842.14
PG1	100000	370	2554	1.9454e-06	6878.21	2117	17.11%	5.5470e-09	7146.87
PG2	110000	410	2832	1.3240e-06	7612.87	2431	14.15%	4.4791e-09	7853.56

For the experimental evaluation of the proposed methodology we have used the available MNA benchmarks in SLICOT [54] and SparseRC [55], as well as two custom large-scale power grids. Their characteristics are shown in the first three columns of Table 3.1, where the SLICOT benchmarks are MNA_1 to MNA_5, the two power grids are PG1 and PG2, and the SparseRC benchmarks are a transmission line (TL), a low noise amplifier (LNA), two mixers (MX3, MX7) and an interconnect structure (IS).

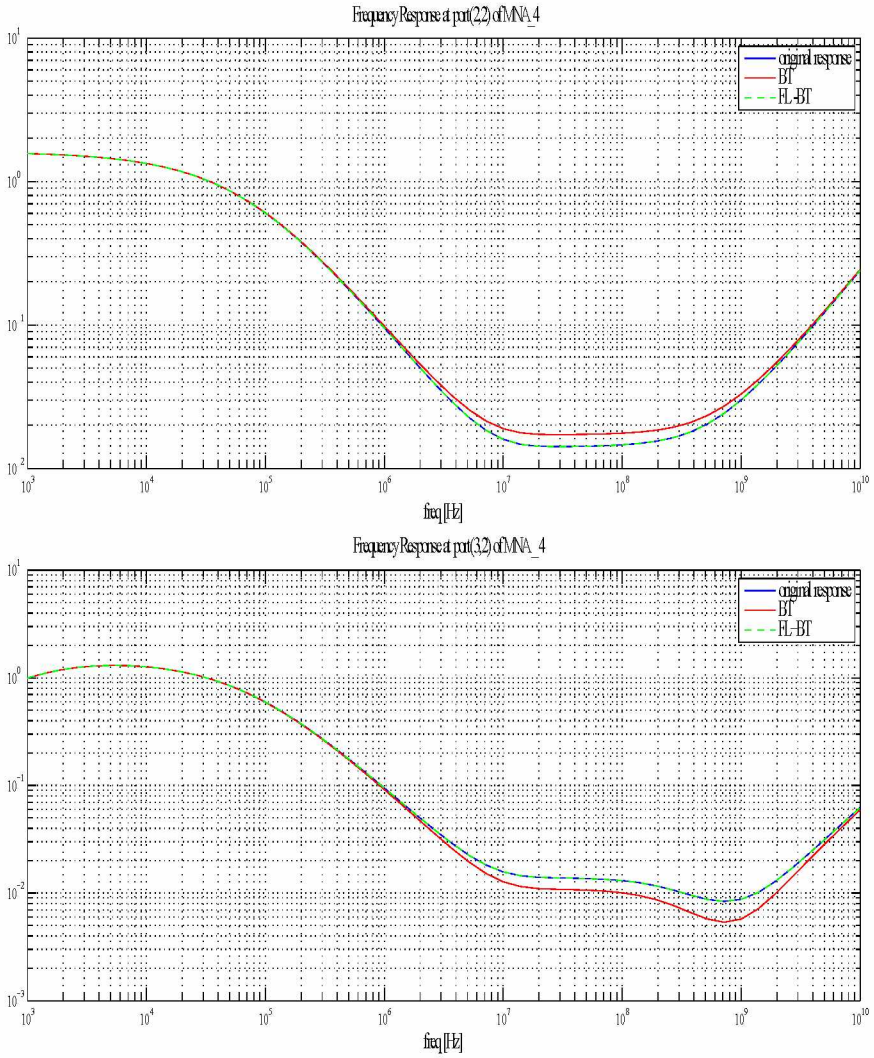


FIGURE 3.1: Comparison of transfer functions of ROMs from standard BT and frequency-limited BT in MNA_4 benchmark at ports (2,2) and (3,2).

The frequency-limited BT was implemented with the procedure of Section 3.6 for a frequency range of $[\omega_1, \omega_2] = [10^3, 10^{10}]$ (purely for testing purposes - one can choose any other range depending on the application), and was compared with the standard (infinite frequency) BT. The reduced-order models (ROMs) of standard and frequency-limited BT were compared both with respect to their order for given tolerance ε, and w.r.t. their accuracy for given ROM order. In the first case the error tolerance was chosen as $\varepsilon = 10^{-4}$, while in the second case the order r was determined by the execution of standard BT and was reused for the truncation of the HSVs of frequency-limited Gramians. All experiments were executed with MATLAB R2015a on a Linux workstation, having a 3.6GHz Intel Core i7 processor with 16GB memory.

The results are reported in the remaining columns of Table 3.1. In the table, *Max error* refers to the maximum error between the transfer functions of the original model and the ROM in the selected frequency range, *Time* refers to the computational time (in seconds) needed to generate the ROMs, while *Reduction percentage* refer to the reduction percentage of the ROMs between the standard BT and the frequency-limited BT. From the table it can be clearly verified that frequency-limited BT can produce ROMs that exhibit either smaller size for given error, or smaller error for given order in comparison to standard BT when restricted

in a finite frequency range. The time needed to generate the frequency-limited ROMs is slightly larger than standard BT because the computation of the matrix logarithm incurs an additional overhead, but the EKS method of Algorithm 4 can effectively mask this overhead to a substantial extent and also makes the procedure applicable to very large circuit models.

Especially for the experiments comparing ROM accuracy for given order, we have plotted in Fig. 3.1 and Fig 3.2 the transfer functions of the original model and the ROMs generated by frequency-limited BT and standard BT for two benchmarks and two and three ports per benchmark. In both figures, the response of the frequency-limited ROM is indistinguishable from the original model in the selected frequency range, while the response of the ROM from standard BT can be seen to exhibit a clear deviation.

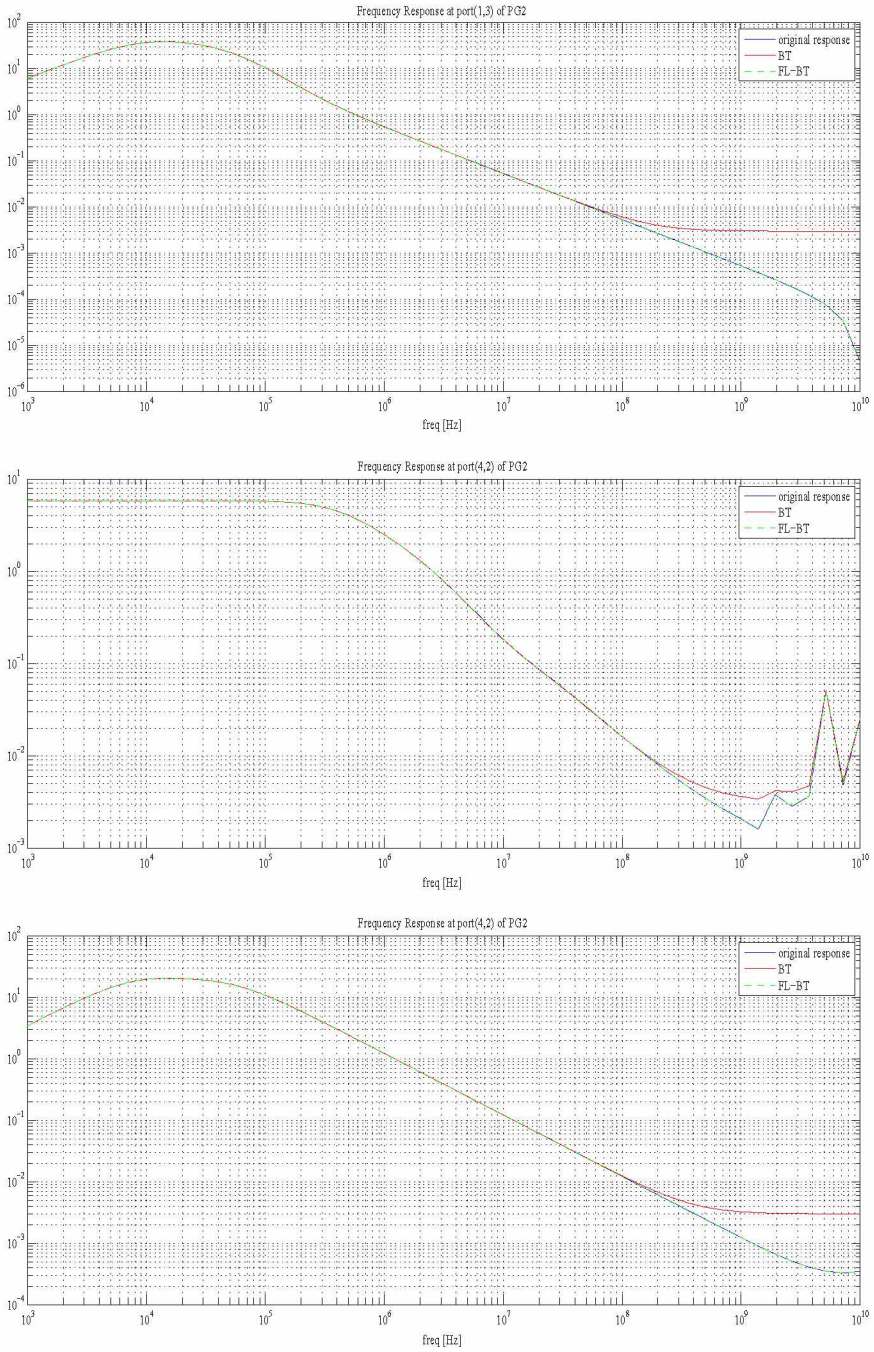


FIGURE 3.2: Comparison of transfer functions of ROMs from standard BT and frequency-limited BT in PG2 benchmark at ports (1,3), (3,2) and (4,2).

Chapter 4

Efficient IC Hotspot Thermal Analysis via Low-Rank Model Order Reduction

4.1 Introduction

Temperature hotspots are an inevitable consequence of the rising power consumption due to the technology downscaling and its nonuniform distribution across the chip. Local temperature gradients have significant impact on chip performance and functionality, leading to slower transistor speed, more leakage power, higher interconnect resistance and reduced reliability. Stacking multiple layers in a 3D chip requires extensive thermal analysis as the power density and temperature of these architectures can be quite high. Moreover, the problem becomes more pronounced due to the use of new device technologies, like FinFETs and Silicon on Insulator (SOI), that are more sensitive to the self-heating effect [56]. Management of the above remains a key issue for future microprocessors and ICs [57].

Therefore, thermal analysis and especially hotspot analysis is one of the most critical challenges arising from the technological evolution. The continuous effort for smaller sizes, in sub 45-nm era, and greater performance as well as the new 3D structures have begun to outspace the ability of heat sinks to dissipate the on-chip power. Due to this fact, ICs thermal analysis problems have drawn considerable attention over the past two decades. To deal with the thermal analysis challenge, most previous methodologies have focused on solving linear systems of equations, that result from modeling approaches such as the Finite Element Method (FEM), the Finite Difference Method (FDM) and the Green's Function method [58–60]. The huge linear systems resulting from thermal modeling approaches require unreasonably long computational times. While the formulation problem, by applying a thermal equivalent circuit, is prevalent and can be easily constructed, the corresponding 3D equations network has an undesirably time consuming numerical simulation over many time-steps.

Previous methodologies, propose efficient solution algorithms for the entire systems of equations that results from any thermal modeling approach but those techniques do not scale well with the equations dimension and can only be applied to a narrow range of problems. However, in many cases, thermal analysis does not need to be performed across the whole chip, but only at some pre-specified hotspots, in order to validate the thermal reliability of the chip or can be useful in addressing the performance of individual devices. In those cases, the very large thermal model can be substituted by a much smaller model with similar behavior at the hotspot points, through a Model Order Reduction (MOR) process.

Moment-matching methods have been reasonably successful in circuit simulation problems to produce reduced order models in a computationally efficient way [2], but they exhibit some drawbacks, the most prominent of which is that they do not offer any a-priori estimation for the approximation error. This can result in reduced models that are not very accurate or of sufficient small order. On the other hand, balancing-type methods (in particular Balanced Truncation [61]) rely on rigorous system theoretic concepts and are capable of providing a global error bound, without having to compute the reduced model first. Thus, models obtained by balancing are generally superior to those obtained by moment-matching [62]. However, they have greater computational complexity than moment-matching methods due to the need for solution of Lyapunov matrix equations which take up $O(n^3)$ time.

In this Section we propose an approach for hotspot thermal analysis that is based on RC equivalent circuit in conjunction with Balanced Truncation for MOR [63, 64], which overcomes the high computational demands by using an Extended Krylov Subspace (EKS) method along with low-rank factorized storage, for solving the Lyapunov matrix equations. The EKS method is a state-of-the-art projection type method which exhibits fast convergence and straightforward implementation.

In particular, the contributions to the problem of thermal analysis are:

- **Compact modeling of hotspots.** The complete equation network is transformed into a reduced model that captures heat flow only in the hotspots, which is sufficient for validating the thermal compliance of the circuit.
- **Accurate results within a predefined error bound.** System theoretic techniques have exact error bound formula allowing to trade off the accuracy and the order of the reduced model.
- **Efficient derivation of the reduced model.** The use of low-rank EKS method method for solving the Lyapunov matrix equations require small computational times and reduced memory storage.
- **Reusable models for different dissipation scenaria.** The reduced model can be computed once but it can be reused with different input sequences that define the heat flow in the IC, and can be solved many times preserving the same error bound.

Experimental results demonstrate around 97% model size reduction as compared to the full 3D network model, with approximation in the order of 10^{-4}.

The rest of the Chapter, is organized as follows. Section 4.2 describes previous work on thermal simulation problem. Section 4.3 introduces the thermal model that was used in the present work. Section 4.4 provides a detailed description of MOR, and more specifically on the theoretical background of low-rank Balanced Truncation and EKS methods. Section 4.5 describes the proposed approach, combining the methods presented in the two previous sections. Moreover, in Sections 4.6, 4.7, and 4.8 we provide an extension of the proposed methodology for limited-time intervals along with efficient computational of the RHS that arise in the particular problem. Finally, Section 4.9 presents the results and a discussion about the advantages of the method.

4.2 Related Work

In this section, we briefly describe some previous works in the area of thermal analysis. As mentioned before most transient thermal analysis methodologies have so far relied on solving of the entire system, using different modeling techniques, based mainly in FEM, the FDM and the Green's Functions.

Research work in [65] adopts the FDM method, with a multigrid approach in order to speed up the simulation process and the FDM method with temporal and spatial adaptation to further accelerate thermal analysis is proposed in [66]. Similarly, in [67], the full-chip thermal transient equations are solved in a similar manner using an Alternating Direction Implicit (ADI) method for enhanced computational efficiency. Also, in [69] the FDM approach and the RC equivalent is used along with modeling of the fluids for microcooling 3D structures. Moreover, parallel approaches with general [70] or dedicated [71, 72] precoditioners was proposed to map the thermal analysis problem in GPUs. In [59] the FEM method is adopted for 2D and 3D geometries along with a multigrid preconditioning method and automatic mesh generation for chip geometries. Finally Green's functions, are used in [60] with discrete cosine transform and its inversion in order to accelerate the numerical computation of the homogeneous and inhomogeneous solution. However, these methods are efficient for limited range of problems, and with the escalation of manufacturing technology can lead to huge systems of equations.

Besides the previous conventional approaches different methods like a Neural Net (NN) approach is used in [73], but since it is based in predictions it does not always provide accurate solution to the crucial problem of thermal analysis. Moreover, a Look Up Table (LUT) method based on the power thermal relation, which develops a double-mesh scheme to capture thermal characteristics and store the results in library files is presented in [74]. However nowadays chips can lead to huge library files due to the highly complex combined heat maps.

The approach in [75], bears a resemblance to the proposed method since a MOR method with moment-matching method is used along with FDM modeling. However, this technique considers only input currents as ports and, as mentioned before, moment-matching techniques do not always provide compact and accurate models. Another moment matching MOR method is described in [76], but it is applicable on the architectural level.

Clearly, the concept of a low-rank system theoretic technique has not yet been introduced in the context of transient thermal analysis. A balancing-type approach, becomes more attractive even for large scale systems with the recent results on low rank approximations that boost the solution of Lyapunov equations, which are the bottleneck of the method.

4.3 On-Chip Thermal Modeling

The primary mechanism of heat transfer in solids is by conduction. The starting point for thermal analysis is Fourier's law of heat conduction [77]:

$$\mathbf{q}(\mathbf{r},t) = -k_t \nabla T(\mathbf{r},t) \tag{4.1}$$

which states that the vector of heat flux density $\mathbf{q}$ (heat flow per unit area and unit time) is proportional to the negative gradient of temperature T at every spacial point $\mathbf{r} = [x,y,z]^T$ and time t, where k_t is the thermal conductivity of the material.
The conservation of energy also states that the divergence of the heat flux $\mathbf{q}$ equals the difference between the power generated by external heat sources and the rate of change of temperature, i.e.

$$\nabla \cdot \mathbf{q}(\mathbf{r},t) = g(\mathbf{r},t) - \rho c_p \frac{\partial T(\mathbf{r},t)}{\partial t} \tag{4.2}$$

where $g(\mathbf{r},t)$ is the power density of the heat sources, c_p is the specific heat capacity of the material, and ρ is the density of the material. By combining (4.1) and (4.2) we have:

$$-k_t \nabla^2 T(\mathbf{r},t) = g(\mathbf{r},t) - \rho c_p \frac{\partial T(\mathbf{r},t)}{\partial t} \tag{4.3}$$

which may be rewritten as the following parabolic Partial Differential Equation (PDE):

$$\begin{aligned} \rho c_p \frac{\partial T(\mathbf{r},t)}{\partial t} &= k_t \nabla^2 T(\mathbf{r},t) + g(\mathbf{r},t) \\ = k_t \left(\frac{\partial^2 T(\mathbf{r},t)}{\partial x^2} + \frac{\partial^2 T(\mathbf{r},t)}{\partial y^2} + \frac{\partial^2 T(\mathbf{r},t)}{\partial z^2}\right) &+ g(\mathbf{r},t) \end{aligned} \tag{4.4}$$

(normally accompanied by appropriate boundary conditions [78]).

A common procedure for the numerical solution of (4.4) is by discretization along the 3 spatial coordinates with steps $\Delta x, \Delta y$ and Δz, and substitution of the spatial second-order derivatives by finite difference approximations, leading to the following expression for temperature $T_{i,j,k}$ at each discrete point (i,j,k) in relation to its neighboring points:

$$\begin{aligned} \rho c_p \frac{dT_{i,j,k}}{dt} &= k_t \frac{T_{i+1,j,k} - 2T_{i,j,k} + T_{i-1,j,k}}{\Delta x^2} \\ &+ k_t \frac{T_{i,j+1,k} - 2T_{i,j,k} + T_{i,j-1,k}}{\Delta y^2} \\ &+ k_t \frac{T_{i,j,k+1} - 2T_{i,j,k} + T_{i,j,k-1}}{\Delta z^2} + g_{i,j,k} \end{aligned} \tag{4.5}$$

or by multiplying by $\Delta x \Delta y \Delta z$:

$$\begin{aligned} \rho c_p (\Delta x \Delta y \Delta z) \frac{dT_{i,j,k}}{dt} \\ - k_t \frac{\Delta y \Delta z}{\Delta x} (T_{i+1,j,k} - 2T_{i,j,k} + T_{i-1,j,k}) \\ - k_t \frac{\Delta x \Delta z}{\Delta y} (T_{i,j+1,k} - 2T_{i,j,k} + T_{i,j-1,k}) \\ - k_t \frac{\Delta x \Delta y}{\Delta z} (T_{i,j,k+1} - 2T_{i,j,k} + T_{i,j,k-1}) \\ = g_{i,j,k} (\Delta x \Delta y \Delta z) \end{aligned} \tag{4.6}$$

There is a well known analogy between thermal and electrical conduction, where temperature corresponds to voltage and heat flow corresponds to current (see Table 4.1).

TABLE 4.1: Analogy between electrical and thermal circuits.

Electrical Circuit	Thermal Circuit
Voltage	Temperature
Current	Heat Flow
Electrical Conductance	Thermal Conductance
Electrical Resistance	Thermal Resistance
Electrical Capacitance	Thermal Capacitance
Current Source	Heat Source

In light of this analogy, eq. (4.6) has a direct correspondence to an electrical circuit where there is a node at every discrete point or cell in the thermal grid (see Fig. 4.1). Every circuit node is connected to spatially neighboring nodes via conductances in the directions x, y, z with values:

$$G_x \equiv \frac{k_t \Delta y \Delta z}{\Delta x}, G_y \equiv \frac{k_t \Delta x \Delta z}{\Delta y}, G_z \equiv \frac{k_t \Delta x \Delta y}{\Delta z} \tag{4.7}$$

and there is a capacitance to ground at every node or thermal cell with value:

$$C \equiv \rho c_p(\Delta x \Delta y \Delta z) \tag{4.8}$$

The heat sources constitute input excitations and are modeled in the equivalent circuit as current sources with values:

$$I_{i,j,k} \equiv g_{i,j,k}(\Delta x \Delta y \Delta z) \tag{4.9}$$

The above current sources are connected at the specific points (i, j, k) or circuit nodes where there is heat flow (i.e. power dissipation from the underlying chip logic blocks).

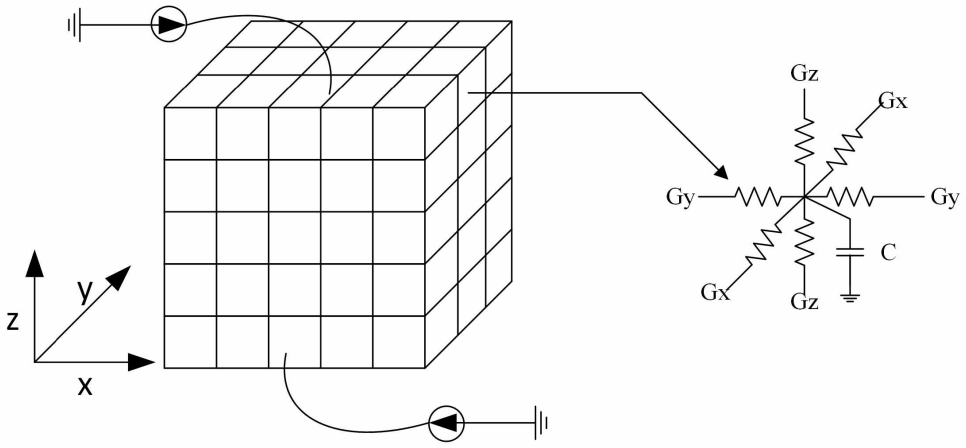


FIGURE 4.1: Spatial discretization of chip for thermal analysis, and formulation of electrical equivalent problem.

It must be noted that in practical IC thermal modeling the thermal conductivity k_t is different for silicon, insulator, and metal materials. However, k_t is typically assumed to be layer-wise uniform, where the substrate has the conductivity of bulk silicon and the interconnect layers, consisting of a mix of insulator and metal materials, take on an average value of k_t depending on the metal density of each layer. If the thickness of the slimmest bottom layer is an integer multiple of the discretization step Δz, then k_t may be taken as constant in (4.6) and (4.7) for points (i, j, k) within the same layer (with proper modifications of the conductances (4.7) for points lying at the interface of two layers - see [79]).

The resulting electrical equivalent circuit is described in the time domain, using the Modified Nodal Analysis (MNA) framework, by a system of Ordinary Differential Equations (ODE):

$$\mathbf{G}\mathbf{x}(t) + \mathbf{C}\frac{d\mathbf{x}(t)}{dt} = \mathbf{E}\mathbf{u}(t) \tag{4.10}$$

where $\mathbf{G} \in \mathbb{R}^{n\times n}$ is a symmetric and positive definite matrix of the conductances (4.7), $\mathbf{C} \in \mathbb{R}^{n\times n}$ is a diagonal matrix of cell capacitances (4.8), $\mathbf{x} \in \mathbb{R}^n$ is the vector of unknown temperatures $T_{i,j,k}$ at all discretization points (constituting internal states of the system), $\mathbf{u} \in \mathbb{R}^p$ is the vector of input excitations from the current sources $I_{i,j,k}$ of (4.9), and $\mathbf{E} \in \mathbb{R}^{n\times p}$ is the input-to-state connectivity matrix.

In many practical scenarios, there is no need for full temperature evaluation at every point (i, j, k) of the 3D discretization, but only at some specific vulnerable or problematic points ("hotspots") that critically affect the operation and reliability of the chip (see Fig. 4.2). In those cases, the hotspot temperatures constitute the output portion of the state vector $\mathbf{x}(t)$ of

temperatures at all possible points, and can be formulated as:

$$\mathbf{y}(t) = \mathbf{L}\mathbf{x}(t) \tag{4.11}$$

where $\mathbf{y} \in \mathbb{R}^q$ is the vector of output hotspot temperatures, and $\mathbf{L} \in \mathbb{R}^{q \times n}$ is the state-to-output connectivity matrix.

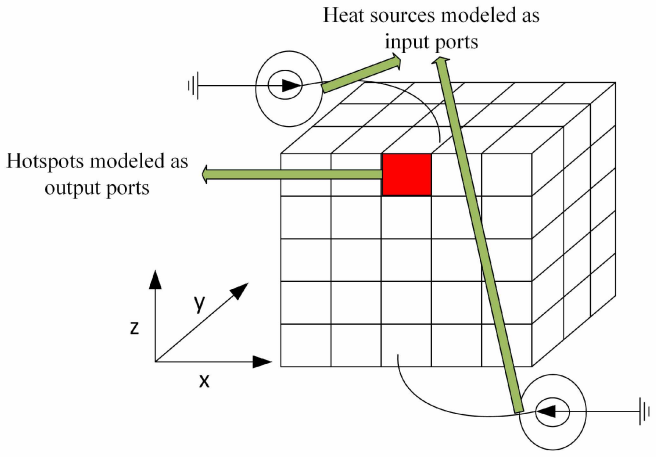


FIGURE 4.2: Schematic depiction of input heat sources and output hotspots in the 3D discretization of a chip.

4.4 Low-Rank Model Order Reduction for Thermal Models

4.4.1 Balanced Truncation for Thermal Models

Eq. (4.10) and (4.11) can be written in standard state space form as:

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{x}(t) \end{aligned} \tag{4.12}$$

with $\mathbf{A} \equiv -\mathbf{C}^{-1}\mathbf{G}$, and $\mathbf{B} \equiv \mathbf{C}^{-1}\mathbf{E}$.
Model Order Reduction (MOR) aims at generating a reduced-order model:

$$\begin{aligned} \frac{d\tilde{\mathbf{x}}(t)}{dt} &= \tilde{\mathbf{A}}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{B}}\mathbf{u}(t), \\ \tilde{\mathbf{y}}(t) &= \tilde{\mathbf{L}}\tilde{\mathbf{x}}(t) \end{aligned} \tag{4.13}$$

with $\tilde{\mathbf{A}} \in \mathbb{R}^{r \times r}$, $\tilde{\mathbf{B}} \in \mathbb{R}^{r \times p}$, $\tilde{\mathbf{L}} \in \mathbb{R}^{q \times r}$, which both exhibits $r << n$ and constitutes a good approximation in the time domain of (4.12), in that the output error is bounded as $||\tilde{\mathbf{y}}(t) - \mathbf{y}(t)||_2 < \varepsilon||\mathbf{u}(t)||_2$ for the given vector of input thermal excitations $\mathbf{u}(t)$ and given small ε. The bound in the output error can be equivalently written in the frequency domain as $||\tilde{\mathbf{y}}(s) - \mathbf{y}(s)||_2 < \varepsilon||\mathbf{u}(s)||_2$ via Plancherel's theorem [4]. If

$$\begin{aligned} \mathbf{H}(s) &= \mathbf{L}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \\ \tilde{\mathbf{H}}(s) &= \tilde{\mathbf{L}}(s\mathbf{I} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{B}} \end{aligned} \tag{4.14}$$

are the transfer functions of the original and the reduced-order model, then the output error in the frequency domain is:

$$\begin{aligned} ||\tilde{\mathbf{y}}(s) - \mathbf{y}(s)||_2 &= ||\tilde{\mathbf{H}}(s)\mathbf{u}(s) - \mathbf{H}(s)\mathbf{u}(s)||_2 \\ &\leq ||\tilde{\mathbf{H}}(s) - \mathbf{H}(s)||_\infty ||\mathbf{u}(s)||_2 \end{aligned} \tag{4.15}$$

where $||.||_\infty$ is the induced $\mathcal{L}_2$ matrix norm, or $\mathcal{H}_\infty$ norm of a rational transfer function. Therefore, the output error can be bounded by bounding the distance between the transfer functions as $||\tilde{\mathbf{H}}(s) - \mathbf{H}(s)||_\infty < \varepsilon$.

Balanced Truncation (BT) and related methods for MOR make use of the controllability and observability Gramian matrices:

$$\begin{aligned} \mathbf{P} &= \int_0^\infty exp(\mathbf{A}t)\mathbf{B}\mathbf{B}^T exp(\mathbf{A}t)^T dt \\ \mathbf{Q} &= \int_0^\infty exp(\mathbf{A}t)^T \mathbf{L}^T \mathbf{L} exp(\mathbf{A}t) dt \end{aligned} \tag{4.16}$$

which are equivalently derived by the solution of the Lyapunov matrix equations [15]:

$$\begin{aligned} \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^T &= -\mathbf{B}\mathbf{B}^T \\ \mathbf{A}^T\mathbf{Q} + \mathbf{Q}\mathbf{A} &= -\mathbf{L}^T\mathbf{L} \end{aligned} \tag{4.17}$$

The controllability Gramian $\mathbf{P}$ characterizes the input-to-state behavior, i.e. the degree to which the states are controllable (reachable) by the inputs, while the observability Gramian $\mathbf{Q}$ characterizes the state-to-output behavior, i.e. the degree to which the states are observable at the outputs. A reduced-order model can, in principle, be obtained by eliminating (truncating) the states that are difficult to reach or observe. However, in the original state-space coordinates there might be states that are difficult to reach but easy to observe, and vice versa. The process of "balancing" is to transform the state vector into a new coordinate system where for every state the degree of difficulty is the same for both reaching and observing it. There exists such a transformation $\mathbf{T}\mathbf{x}(t)$, which leads to a new model:

$$\begin{aligned} \frac{d(\mathbf{T}\mathbf{x}(t))}{dt} &= \mathbf{T}\mathbf{A}\mathbf{T}^{-1}(\mathbf{T}\mathbf{x}(t)) + \mathbf{T}\mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{T}^{-1}(\mathbf{T}\mathbf{x}(t)) \end{aligned} \tag{4.18}$$

(thus preserving the transfer function $\mathbf{H}(s)$) and makes [16]:

$$\mathbf{P} = \mathbf{Q} = diag(\sigma_1, \sigma_2, ..., \sigma_n) \tag{4.19}$$

where $\sigma_i, i = 1, \ldots, n$ are known as the Hankel singular values (HSVs) of the model and are equal to the square roots of the eigenvalues of the product $\mathbf{P}\mathbf{Q}$ (in any coordinate system of state-space), i.e. $\sigma_i = \sqrt{\lambda_i(\mathbf{P}\mathbf{Q})}, i = 1, \ldots, n$. In the balanced model (4.18) the states that are easier to reach and observe correspond to the largest HSVs, and if r of them are kept (truncating the $n - r$ states corresponding to the smallest HSVs) it can be shown that the distance between the original and the reduced-order transfer functions is bounded as [15]:

$$||\mathbf{H}(s) - \tilde{\mathbf{H}}(s)||_\infty \leq 2(\sigma_{r+1} + \sigma_{r+2} + ... + \sigma_n) \tag{4.20}$$

The latter is an "a-priori" criterion for selecting the order of the reduced model for a desired output error tolerance ε, and is a significant advantage of BT-type methods for MOR. The main steps of BT are summarized in the following Algorithm 6 [17]:

Algorithm 6 MOR by Balanced Truncation

1: Solve the Lyapunov equations (4.17) to obtain the Gramian matrices $\mathbf{P}$ and $\mathbf{Q}$ in low-rank format as described in Section 4.4.3
2: Compute the eigenvalue decomposition of $\mathbf{PQ}$, or equivalently the singular value decomposition (SVD) of the product of the Cholesky factors $\mathbf{P} = \mathbf{Z}_P\mathbf{Z}_P^T$ and $\mathbf{Q} = \mathbf{Z}_Q\mathbf{Z}_Q^T$, i.e. $\mathbf{Z}_P^T\mathbf{Z}_Q = \mathbf{U\Sigma V}$
3: Compute the truncated part of the balancing transformations $\mathbf{T}_{(r\times n)} = \mathbf{\Sigma}_{(r\times r)}^{-1/2}\mathbf{V}_{(r\times n)}\mathbf{Z}_Q^T$ and $\mathbf{T}_{(n\times r)}^{-1} = \mathbf{Z}_P\mathbf{U}_{(n\times r)}\mathbf{\Sigma}_{(r\times r)}^{-1/2}$, and the corresponding reduced-order model matrices as
$\tilde{\mathbf{A}} = \mathbf{T}_{(r\times n)}\mathbf{A}\mathbf{T}_{(n\times r)}^{-1}$, $\tilde{\mathbf{B}} = \mathbf{T}_{(r\times n)}\mathbf{B}$, $\tilde{\mathbf{L}} = \mathbf{L}\mathbf{T}_{(n\times r)}^{-1}$

4.4.2 Low-Rank Solution of Lyapunov Equations

The main drawback of BT is the significant computational and memory cost for deriving the reduced-ordered model, which is a serious obstacle in the applicability of BT for the reduction of large-scale models (with n more than a few thousand states or so). That is because the solution of Lyapunov equations, the Cholesky factorization and the SVD are all computationally expensive tasks of complexity $O(n^3)$, and also involve dense matrices since the Gramians $\mathbf{P}, \mathbf{Q}$ are dense even if the system matrices $\mathbf{A}, \mathbf{B}, \mathbf{L}$ are sparse.

However, it is almost always the practical case that the number of inputs and outputs is much smaller than the number of states, i.e. $p, q << n$. This means that the products $\mathbf{BB}^T$ and $\mathbf{L}^T\mathbf{L}$ will have low numerical rank compared to n, and this will also hold for the corresponding Gramians [40], allowing their own approximation by low-rank products instead of full Cholesky factorizations, i.e. $\mathbf{P} \approx \mathbf{Z}_P\mathbf{Z}_P^T$ and $\mathbf{Q} \approx \mathbf{Z}_Q\mathbf{Z}_Q^T$ with $\mathbf{Z}_P, \mathbf{Z}_Q \in \mathcal{R}^{n\times k}$ $(k << n)$. This greatly reduces the memory requirements, as well as the complexity of the factorization and SVD which are now of size k instead of full n, leaving the solution of Lyapunov equations as the main computational task of low-rank BT.

As we mentioned before, the two recent classes of algorithms that have been developed for directly solving the Lyapunov equations in low-rank factorized form are the Alternating Direction Implicit (ADI) [26] and the projection-type or Krylov-subspace methods [41]. The ADI method exhibits fast convergence but requires the input of a number of shift parameters, whose choice greatly affects convergence but relies on unclear heuristics and is very problem-dependent. Projection-type methods do not depend on the selection of specific parameters and their algorithmic implementation is more straightforward and well-studied, having been successfully used for several years for the solution of conventional linear systems of equations. However, they generally have not been competitive with ADI methods, until the recent development of the extended Krylov subspace method (EKS) [25] which employs two complementary subspaces to radically speed up convergence [42]. In this dissertation we propose the use of the EKS method for the low-rank solution of Lyapunov equations arising in the application of BT for the reduction of large-scale thermal models.

4.4.3 Extended Krylov Subspace Method

The essence of low-rank projection type methods is to project the large-scale Lyapunov equations (4.17) onto a lower-dimensional subspace, and then solve the resulting small-scale equations to obtain the low-rank approximate solutions of (4.17).

More specifically, if $\mathbf{K} \in \mathbb{R}^{n\times k}$ $(k << n)$ is a projection matrix whose columns span the k-dimensional Krylov subspace $\mathcal{K}_k(\mathbf{A}, \mathbf{B}) = span\{\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \ldots, \mathbf{A}^{k-1}\mathbf{B}\}$, then the

projected Lyapunov equation (for the controllability Gramian $\mathbf{P}$) onto $\mathcal{K}_k(\mathbf{A},\mathbf{B})$ is:

$$(\mathbf{K}^T\mathbf{A}\mathbf{K})\mathbf{X}+\mathbf{X}(\mathbf{K}^T\mathbf{A}\mathbf{K})^T=-\mathbf{K}^T\mathbf{B}\mathbf{B}^T\mathbf{K} \tag{4.21}$$

The solution of $\mathbf{X}\in\mathbb{R}^{k\times k}$ of (4.21) can be back-projected to the n-dimensional space to give an approximate solution for the original large-scale equation (4.17) as $\mathbf{P}=\mathbf{K}\mathbf{X}\mathbf{K}^T$, and a low-rank factor $\mathbf{Z}\in\mathcal{R}^{n\times k}$ of $\mathbf{P}$ can be obtained as $\mathbf{Z}=\mathbf{K}\mathbf{S}$ where $\mathbf{X}=\mathbf{S}\mathbf{S}^T$ is the Cholesky factorization of $\mathbf{X}$.

The projection process is independent of the subspace selection, but its effectiveness is critically dependent on the chosen subspace and can sometimes take many iterations of subspace updating before converging to the final solution. The convergence problem can be alleviated by enriching the standard Krylov subspace $\mathcal{K}_k(\mathbf{A},\mathbf{B})$ with information from the subspace $\mathcal{K}_k(\mathbf{A}^{-1},\mathbf{B})$ corresponding to the inverse matrix $\mathbf{A}^{-1}$, leading to the extended Krylov subspace:

$$\mathcal{K}_k^E(\mathbf{A},\mathbf{B})=\mathcal{K}_k(\mathbf{A},\mathbf{B})+\mathcal{K}_k(\mathbf{A}^{-1},\mathbf{B})=$$

$$span\{\mathbf{B},\mathbf{A}^{-1}\mathbf{B},\mathbf{A}\mathbf{B},\mathbf{A}^{-2}\mathbf{B},\mathbf{A}^2\mathbf{B},\ldots,\mathbf{A}^{-(k-1)}\mathbf{B},\mathbf{A}^{k-1}\mathbf{B}\}$$

The extended Krylov subspace method starts by the pair $\{\mathbf{B},\mathbf{A}^{-1}\mathbf{B}\}$ and generates a sequence of extended subspaces $\mathcal{K}_k^E(\mathbf{A},\mathbf{B})$ of increasing dimensions, solving the projected Lyapunov equation (4.21) in each iteration, until a sufficiently accurate approximation of the solution of (4.17) is obtained. The complete EKS procedure for the thermal grids is given in Algorithm 7.

Algorithm 7 Extended Krylov Subspace (EKS) Method for low-rank solution of Lyapunov equations arise in thermal models

Input: $\mathbf{A}$, $\mathbf{B}$ (or $\mathbf{A}^T$, $\mathbf{L}^T$)
Output: $\mathbf{Z}$ such that $\mathbf{P}\approx\mathbf{Z}\mathbf{Z}^T$,

$p=size\_col(\mathbf{B})$; $j=1$;
$\mathbf{K}^{(j)}=Orth([\mathbf{B},\mathbf{A}^{-1}\mathbf{B}])$
while $j<maxiter$ **do**
  $\mathbf{M}=\mathbf{K}^{(j)T}\mathbf{A}\mathbf{K}^{(j)}$; $\mathbf{R}=\mathbf{K}^{(j)T}\mathbf{B}$
  Solve $\mathbf{M}\mathbf{X}+\mathbf{X}\mathbf{M}^T=-\mathbf{R}\mathbf{R}^T$ for $\mathbf{X}\in\mathcal{R}^{2pj\times 2pj}$
  if converged **then**
    $\mathbf{S}=Chol(\mathbf{X})$
    $\mathbf{Z}=\mathbf{K}^{(j)}\mathbf{S}$
    break;
  end if
  $s=2p(j-1)$; $e=s+p$; $f=2pj$
  $\mathbf{K}_1=[\mathbf{A}\mathbf{K}^{(j)}(:,s+1:e),\mathbf{A}^{-1}\mathbf{K}^{(j)}(:,e+1:f)]$
  $\mathbf{K}_2=Orth(\mathbf{K}_1)$ w.r.t $\mathbf{K}^{(j)}$
  $\mathbf{K}_3=Orth(\mathbf{K}_2)$
  $\mathbf{K}^{(j+1)}=[\mathbf{K}^{(j)},\mathbf{K}_3]$
  $j=j+1$
end while

4.4.4 EKS Method Implementation Details

In this section we elaborate on some points regarding the efficient implementation of the EKS method Algorithm 7.

- **Matrix products with inverse of sparse matrix.** Algorithm 7 involves the inverse $\mathbf{A}^{-1}$ of the *sparse* system matrix $\mathbf{A}$. Unfortunately, the inverse of a sparse matrix is a dense matrix, and also inversion is a very expensive computational task which should be avoided whenever the inverse is not needed explicitly. Since, however, the inverse $\mathbf{A}^{-1}$ is only applied in products with the $n \times p$ matrix $\mathbf{B}$ (initially) and the $n \times pj$ matrix $\mathbf{K}^{(j)}(:, e+1 : f)$ in each iteration j (where p, $pj << n$, and the iteration count is typically very small), we can compute these products by solving the p and pj sparse systems $\mathbf{AY} = \mathbf{B}$ and $\mathbf{AY} = \mathbf{K}^{(j)}(:, e+1 : f)$. The normal structure of the thermal grids allows the use a highly parallel iterative Preconditioned Conjugate Gradient (PCG) method, which overcomes the computational demands for the very large systems arising from the thermal modeling algorithm like [71, 72].

- **Orthogonalization.** An Arnoldi iteration which uses a modified Gram-Schmidt procedure [46], which exploit sparse-matrix vector products, is employed in Algorithm 7 to handle the orthogonalization ($Orth()$) steps.

- **Solution of small-scale Lyapunov equation.** A direct Schur decomposition method [81] similar to the previous chapter can be employed for the solution of the small-scale ($2pj \times 2pj$) Lyapunov equation in each iteration of Algortihm 7.

- **Convergence criterion.** An appropriate stopping criterion as described in the previous Chapter, for Algorithm 7, is the residual of eq. (4.17) with the approximate solution $\mathbf{P} = \mathbf{KXK}^T$ to reach a certain threshold in magnitude, i.e.

$$||\mathbf{AKXK}^T + \mathbf{KXK}^T\mathbf{A}^T + \mathbf{BB}^T|| \leq tol \tag{4.22}$$

  In fact for this type of problems, it can be shown [25] that the above residual norm is equal to $||\mathbf{R}^T\mathbf{MX}||$ which can be computed more efficiently, and thus the stopping criterion becomes:

$$||\mathbf{R}^T\mathbf{MX}|| \leq tol \tag{4.23}$$

  A tolerance of $tol = 10^{-10}$ is typically employed in practice to acquire a good approximation of the solution.

- **Lower rank solution.** The solution $\mathbf{Z}$ obtained after the termination of Algorithm 7 has rank $2pj$, where j is the final iteration count. This can be reduced even further by employing in step 7 the eigendecomposition $\mathbf{X} = \mathbf{W\Lambda W}^T$, instead of the Cholesky factorization $\mathbf{X} = \mathbf{SS}^T$ for the solution $\mathbf{X}$ of the final projected Lyapunov equation. By keeping only the k eigenvalues above a certain threshold (a fair choice of threshold is 10^{-12}), along with the corresponding eigenvectors, a factor $\mathbf{Z}$ of $\mathbf{P}$ with lower rank $k < 2pj$ can be obtained as $\mathbf{Z} = \mathbf{K}\mathbf{W}_{(2pj \times k)}\mathbf{\Lambda}^{\frac{1}{2}}_{(k \times k)}$.

4.5 Proposed Methodology for Hotspot Thermal Simulation

The proposed methodology for hotspot thermal simulation is summarized in the following steps:

- **3D discretization of the chip.** The spatial steps Δx, Δy in the x- and y-direction are user defined, but the step Δz along the z-direction is typically chosen to coincide with the interface between successive layers (metal and insulator). The discretization procedure naturally covers multiple layers in the z-direction, and can be easily extended to model heterogeneous structures that can be found in modern chips (e.g. heat sinks).

- **Construction of equivalent electrical circuit.** The RC elements of the electrical equivalent are calculated by (4.7) and (4.8).

- **Formulation of equivalent circuit description.** Using Modified Nodal Analysis, the equivalent circuit is described by the ODE system (4.10).

- **Estimation of power consumption profile of chip logic blocks.** This determines the location and the time behavior of heat sources, which in turn specify the structure of the input-to-state connectivity matrix $\mathbf{E}$, and the value of current sources (4.9) that constitute the vector $\mathbf{u}(t)$ in (4.10).

- **Selection of hotspots.** The hotspots are usually the same points where heat sources are applied, but can be any other user-defined points along the layer stack of the chip (a specially inter-layer vias and points on the upper metal layer where power distribution pins are connected). Since the hotspots constitute the outputs $\mathbf{y}(t)$ of the model, their location specifies the structure of the state-to-output connectivity matrix $\mathbf{L}$ in (4.11) (with 1s at the appropriate matrix positions).

- **Formulation of state-space model.** This results from (4.10) and (4.11) as:

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= -\mathbf{C}^{-1}\mathbf{G}\mathbf{x}(t) + \mathbf{C}^{-1}\mathbf{E}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{x}(t) \end{aligned} \tag{4.24}$$

  where the inversion of $\mathbf{C}$ is trivial since it is a diagonal matrix.

- **Construction of reduced-order model.** This is performed by the Balanced Truncation Algorithm 6, with the EKS method Algorithm 7 employed to compute low-rank solutions of the Lyapunov equations (4.17) and analogous approximations of the system Gramians $\mathbf{P}$ and $\mathbf{Q}$. Note that there is no need for passivity preservation in thermal analysis problems, since the reduced-order model is not interconnected with other thermal models but is used individually in multiple transient thermal simulations, and thus it is only needed to be accurate enough to capture all thermal effects at the hotspots.

- **Simulation of the reduced-order model.** This can be performed by the Backward-Euler (BE) differential approximation, where a direct or iterative linear solver can be employed for the solution of the resulting linear systems at each discrete point in time.

4.6 Extension in Limited-Time Intervals

In most practical applications there is typically a final time at which all thermal effects can be considered to have reached steady-state. This means that the reduced-order model can become unnecessarily large to achieve approximation over an infinite time period.

Recalling the controllability and observability Gramians of (4.16) and restricting the integration limits to a time interval $[0, t_1]$, with $0 < t_1 < \infty$, the time-limited Gramians are obtained by

$$\begin{aligned} \mathbf{P}_t &= \int_0^{t_1} exp(\mathbf{A}t)\mathbf{B}\mathbf{B}^T exp(\mathbf{A}t)^T dt \\ \mathbf{Q}_t &= \int_0^{t_1} exp(\mathbf{A}t)^T \mathbf{L}^T \mathbf{L} exp(\mathbf{A}t) dt \end{aligned} \tag{4.25}$$

Equivalently, $\mathbf{P}_t$ and $\mathbf{Q}_t$ can be derived by the solution of the following modified Lyapunov equation [30, 31]:

$$\begin{aligned} \mathbf{A}\mathbf{P}_t + \mathbf{P}_t\mathbf{A}^T &= -\mathbf{B}\mathbf{B}^T + \mathbf{F}\mathbf{B}(\mathbf{F}\mathbf{B})^T \\ \mathbf{A}^T\mathbf{Q}_t + \mathbf{Q}_t\mathbf{A} &= -\mathbf{L}^T\mathbf{L} + (\mathbf{L}\mathbf{F})^T\mathbf{L}\mathbf{F} \end{aligned} \tag{4.26}$$

where

$$\mathbf{F} = e^{\mathbf{A}t_1} \tag{4.27}$$

The time-limited Gramians characterize the controllability and observability of the model in the selected time window, and the process of balancing and truncation will eliminate states that are difficult to reach and observe inside this time range. This means that more states can be eliminated for a given tolerance in (4.20) leading to lower order r in the reduced model, or alternatively to lower error in the time range for a given order r. However, in order to compute $\mathbf{P}_t$ and $\mathbf{Q}_t$ by solving (4.26) we have to deal with the different RHS of time-limited Lyapunov equations which require the computation of a matrix exponential. In order to deal with the computation of the matrix exponential, we adopt the EKS method as described in the next section.

4.7 Computation of the RHS of the Time-Limited Gramians

To construct the RHS of the time-limited Lyapunov equations of (4.26) we need to compute the matrix exponential $\mathbf{F}$ of (4.27). Computing the matrix exponential is generally an intensive procedure. Fortunately, the matrix exponential is not explicitly needed in but only the effects in pre-multiplying p, $(e^{\mathbf{Y}}\mathbf{B})$ or post-multiplying q vectors $(\mathbf{L}e^{\mathbf{Y}})$ are required. Similarly, these can be effectively computed by projection algorithms for evaluating matrix functions, which iteratively project the large-scale input matrices onto lower dimensional subspaces and compute the analogous small-scale problem of matrix functions times a vector. The dimension of the projection subspace is increased in every iteration until convergence is achieved. The most effective modern algorithm for evaluating large-scale matrix functions is a similar procedure based on the EKS method [51], where the two complementary subspaces are employed in the same way as we described them in the previous section. The implementation is shown in Algorithm 8.

Note that in steps 4 and 7 the operations $\mathbf{Y}\mathbf{K}^{(j)}$ and $\mathbf{Y}^{-1}\mathbf{K}^{(j)}$ we use the same parallel iterative Preconditioned Conjugate Gradient (PCG) method as previous.

Algorithm 8 Extended Krylov subspace method for matrix exponential multiplying p vectors $\mathbf{B}$

Input: $\mathbf{Y}, \mathbf{B}$, Convergence tolerance ϵ
Output: $\mathbf{v} = exp(\mathbf{Y})\mathbf{B}$

1: $j = 1$;
2: $\mathbf{K}^{(j)} = Orth([\mathbf{B}, \mathbf{Y}^{-1}\mathbf{B}])$;
3: **repeat**
4: $\quad \mathbf{X} = \mathbf{K}^{(j)T}\mathbf{Y}\mathbf{K}^{(j)}; \quad \mathbf{R} = \mathbf{K}^{(j)T}\mathbf{B}$
5: $\quad$ Compute $\mathbf{y} = exp(\mathbf{X})\mathbf{R}$
6: $\quad k_1 = 2p(j-1); k_2 = k_1 + p; k_3 = 2pj$
7: $\quad \mathbf{K}_1 = [\mathbf{Y}\mathbf{K}^{(j)}(:, k_1+1 : k_2); \mathbf{Y}^{-1}\mathbf{K}^{(j)}(:, k_2+1 : k_3)]$
8: $\quad \mathbf{K}_2 = Orth(\mathbf{K}_1) \quad$ w.r.t $\quad \mathbf{K}^{(j)}$
9: $\quad \mathbf{K}_3 = Orth(\mathbf{K}_2)$
10: $\quad \mathbf{K}^{(j+1)} = [\mathbf{K}^{(j)}, \mathbf{K}_3]$
11: $\quad j = j + 1$
12: **until** $\frac{||\mathbf{y}^{(j)} - \mathbf{y}^{(j-1)}||}{||\mathbf{y}^{(j)}||} < \epsilon$
13: $\mathbf{v} = \mathbf{K}^{(j)}\mathbf{y}$

4.8 Proposed Methodology for Time-Limited Hotspot Thermal Simulation

In addition to the previously described steps, only a small modification is needed in the *construction of the reduced-order model*, where the following steps need to be employed:

- The quantity of (4.27) are evaluated through the extended Krylov subspace method of Algorithm 8.
- The low-rank factors $\mathbf{Z}_p$ and $\mathbf{Z}_q$ of the time-limited Gramians are computed through the EKS method of Algorithm 7.
- The states that are difficult to reach or observe in the selected time interval are eliminated (truncated) by executing steps 2 and 3 of BT algorithm 6.

4.9 Experimental Results

In order to evaluate the efficiency of the proposed methodology for thermal analysis, we have created a set of artificial benchmark circuits that represent simplified microprocessor designs with random control logic and datapath. The characteristics of the constructed benchmarks are shown in Table 4.2.

The 3D discretization of each benchmark was performed with 10000 points along each material layer, and the RC equivalent electrical circuit was constructed. All hotspots were taken at the same points as the heat sources (with no loss of generality), i.e. every input port was also an output port in the state space model (4.12). Finally, the dissipation excitation of the heat sources where random piece-wise-linear functions.

The Reduced-Order Models (ROMs) were obtained by the BT Algorithm 6, with the EKSM Algorithm 7 for the computation of Gramians in low-rank form. All experiments were run using Matlab R2015a on a Linux workstation having an Intel Core i7 processor with 8 cores at 3.6GHz and 16GB memory. The tolerance error for BT was selected as $\varepsilon = 10^{-4}$, which is very strict but still leads to compact ROMs. The reduction results are shown in Table 4.3.

TABLE 4.2: Statistics of benchmark circuits. Material layers include both metal and insulator layers, and heat sources represent sources of power dissipation from chip logic blocks.

Benchmark	Metal Layers	Material Layers	Heat Sources
ckt1	3	5	200
ckt2	3	6	220
ckt3	4	7	260
ckt4	4	8	295
ckt5	5	9	330
ckt6	5	10	370
ckt7	6	11	410

TABLE 4.3: Model Order Reduction results.

Benchmark	Original Size	Size of ROM	Reduction Percentage
ckt1	50000	1543	96.91%
ckt2	60000	1753	97.07%
ckt3	70000	1976	97.17%
ckt4	80000	2177	97.27%
ckt5	90000	2321	97.42%
ckt6	100000	2540	97.46%
ckt7	110000	2732	97.51%

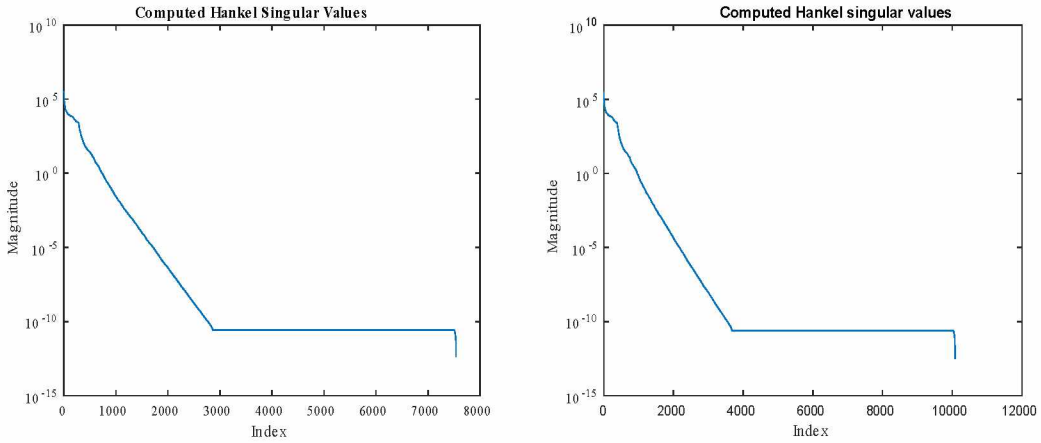
The above results demonstrate that system theoretic techniques like BT can achieve very high reduction percentages, of about 97%, and thus can lead to very compact ROMs for the efficient capture of thermal effects at the hotspots. The resulting ROMs exhibit low memory requirements for storing the system matrices and significantly faster transient simulation.

Fig. 4.3 also displays graphically the magnitude of the Hankel Singular Values (HSVs) in decreasing order for two benchmark circuits, where it can be clearly seen that more than 90% have negligible contribution to the system dynamics.

Furthermore, to evaluate the accuracy and efficiency of the resulting ROMs, we compared their transient simulation against the original, with both state-of-the-art direct methods and iterative methods with zero-fill incomplete factorization preconditioners [44, 45]. Both the original models and the ROMs were simulated over 200 time-steps, from 0 to 0.2 seconds, and the runtime results are reported in Tables 4.4 and 4.5 for direct and iterative methods respectively. In the tables, *Simul. Time* refers to the average time (in seconds) per time-step required for the transient solution. It can be observed, that the compact models provide a significant acceleration which ranges from 26.33X to 46.33X for direct methods, and from 3.75X to 5.63X for iterative methods. The speedup is significantly more pronounced for direct methods since the size of the resulting ROMs is quite small.

Also, Fig. 4.4 depicts the simulated waveforms of temperature over time at certain hotspots of benchmarks ckt4 and ckt5, where the responses of the ROMs show a nearly perfect match with the responses of the original models.

Finally, in order to achieve enhanced accuracy the time-limited BT was implemented with the procedures that was described in the previous sections for a time range $[0, 2]$ (purely for testing purposes - one can choose any other range depending on the application), and was compared with the standard (infinite time) BT that we calculated before. The ROMs of standard and time-limited BT were compared both with respect to their order for given



(A) HSVs of ckt3. (B) HSVs of ckt5.

FIGURE 4.3: Magnitude of Hankel Singular Values for two benchmark circuits.

TABLE 4.4: Runtime results for transient thermal simulation of the original and the reduced-order model with direct methods.

Benchmark	Original Model Simul. Time	ROM Simul. Time	Speedup
ckt1	1.58	0.06	26.33X
ckt2	1.78	0.06	29.66X
ckt3	2.34	0.07	33.42X
ckt4	3.04	0.09	33.77X
ckt5	3.87	0.10	28.77X
ckt6	4.89	0.11	44.45X
ckt7	5.56	0.12	46.33X

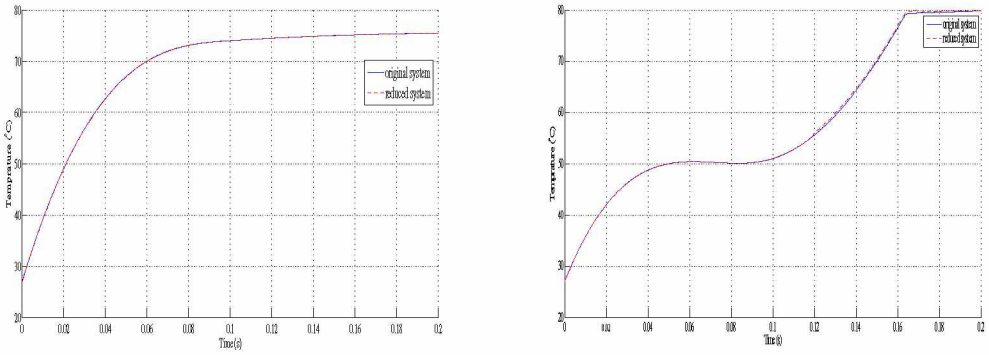
tolerance ε, and w.r.t. their accuracy for given ROM order. In the first case the error tolerance was chosen as $\varepsilon = 10^{-4}$, while in the second case the order r was determined by the execution of standard BT and was reused for the truncation of the HSVs of time-limited Gramians. The results are reported in the of Table 4.6 where *Max error* refers to the absolute maximum error between the temperature waveforms of the original model and the ROM in the selected time range, while *Reduction Percentage* refers to the reduction percentage of the original model and the time-limited BT. From the table it can be clearly verified that time-limited BT can produce ROMs that exhibit either smaller size for given error, or smaller error for given order in comparison to standard BT when restricted in a finite frequency range.

The above results demonstrate that the proposed enchaced time-limited methodology can achieve slightly higher reduction percentages, of more than 97%, but provides a clear improvemnt in the error metric.

Finally, Fig. 4.5 depicts the simulated waveforms with the time-limited methodology of temperature over time at a hotspot of ckt3 benchmark, where the response of the ROM show a perfect match with the responses of the original model in the selected time window (with error less than 10^{-8}).

TABLE 4.5: Runtime results for transient thermal simulation of the original and the reduced-order model with iterative methods.

Benchmark	Original Model Simul. Time	ROM Simul. Time	Speedup
ckt1	1.05	0.28	3.75X
ckt2	1.11	0.31	3.58X
ckt3	1.62	0.37	4.37X
ckt4	2.05	0.40	5.12X
ckt5	2.22	0.41	5.41X
ckt6	2.46	0.43	5.72X
ckt7	2.76	0.49	5.63X



(A) Transient simulation of a hotspot in benchmark ckt4.

(B) Transient simulation of a hotspot in benchmark ckt5.

FIGURE 4.4: Results of transient analysis for original and reduced-order model at a hotspot point in two benchmark circuits.

TABLE 4.6: Reduction results of time-limited BT vs standard BT for various circuit benchmarks.

Bench.	Stadard BT		Time-limited BT		
	ROM order	Max error	ROM order for same error	Percent %	Max error for same order
ckt1	1543	5.58e-06	1265	97.47%	2.48e-09
ckt2	1753	1.64e-05	1442	97.59%	1.64e-08
ckt3	1976	7.71e-05	1542	97.79%	3.76e-10
ckt4	2177	7.14e-05	1775	97.78%	6.44e-08
ckt5	2321	4.66e-06	1898	97.89%	9.12e-09
ckt6	2540	7.89e-06	2114	97.86%	7.33e-09

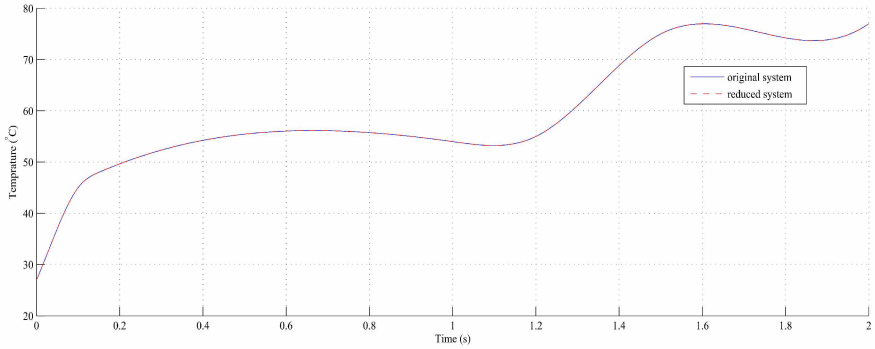


FIGURE 4.5: Transient simulation of a hotspot in benchmark ckt3 with time-limited BT.

Chapter 5

Conclusions and Future Directions

5.1 Conclusions

This dissertation focus was given in circuit and thermal models derived from integrated circuits and the calculation of the reduced model on specific frequency or time windows, as well as on the efficient implementation of Balanced Truncation and Lyapunov matrix equations solution methods. More specifically we have presented:

- An efficient method for reduction of large-scale circuit models in finite frequency windows, which only requires the specification of the end frequencies and avoids the need for ambiguous frequency weighting functions. The method has been shown to provide clear improvements in model accuracy or size with respect to standard Balanced Truncation, while retaining its benefits of specified error bounds. The implementation of the proposed method has been made with efficient computational choices, as well as adaptations and modifications of large-scale methods for matrix equations.
- A compact modeling for hotspot transient thermal simulation. Unlike traditional methods that focus on the solution of the full thermal problem in an IC, this work focused on the observation that the simulation of the thermal phenomena in hotspots is sufficient to define the proper functionality of the circuit and additional nodes contain thermal information with little value to reliability issues. Experimental results shown that this method can provide very compact models for large 3D structures with acceptable error in transient analysis. Summarizing, MOR techniques can provide attractive solution to the problem of thermal analysis.

5.2 Future Directions

In the future, we plan to extend the research presented in this dissertation towards the following directions:

- The computational implementation of MOR methods can be improved by exploiting the massive parallelism of modern heterogeneous systems with simultaneous multicore processors and graphic processing units (GPUs). Existing implementations in the literature are scarce, and perhaps the greatest benefits arise in matrix equation solvers, which can handle really large models with the advantage of fast computation of the reduced model.
- Generating linear models that can be easily synthesized. Concerning the synthesis of the reduced models, the industry requirement is the passive composition without transformers, but this usually results in a much larger number of RLC elements compared to the reduced-order model. This is directly related to the fact that the matrix loses its sparse structure, so tackling the problem with graph-theoretic techniques may also

solve the problem of non-minimal composition. Another important problem is that the synthesis procedure produces circuits with negative RLC values, which can cause various problems when simulating the resulting models.

- Generate guaranteed passive models. As for the passivity of the reduced-order model, all existing methods try moving a small subset of poles that initially make the model non-passive. Usually, however, this adversely affects the accuracy of the approach, and the best solution is to move a wider set of poles (not just those that affect the passivity) so that the reduced model becomes passive without loss of accuracy.

Appendix A

Relevant Publications

Conference publications:

- George Floros, Nestor Evmorfopoulos and George Stamoulis. "THANOS: Eliminating Redundant States in Transient Thermal Analysis". International Workshop Thermal Investigations Of ICs And Systems (THERMINIC), September 25-27, 2019, at Lecco, Italy.
- George Floros, Nestor Evmorfopoulos and George Stamoulis. "Efficient Circuit Reduction in Limited Frequency Windows". International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), July 15-18, 2019, at Lausanne, Switzerland.
- George Floros, Nestor Evmorfopoulos and George Stamoulis. "Efficient Reduction of Large Circuit Models Over Limited Frequency Windows". Design Automation Conference (DAC), June 2-6, 2019, Work-in-Progress (WIP) Poster Session, Las Vegas, NV, USA.
- George Floros, Nestor Evmorfopoulos and George Stamoulis. "Efficient Hotspot Thermal Simulation via Low Rank Model Order Reduction". International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), July 2-5, 2018, at Prague, Czech Republic.

Journal publications:

- George Floros, Nestor Evmorfopoulos and George Stamoulis. "Frequency-Limited Reduction of Large Circuit Models via Low-Rank Sparse ADI Method". IEEE Transactions on VLSI. (**submitted**)
- George Floros, Nestor Evmorfopoulos and George Stamoulis. "Efficient Hotspot Thermal Simulation via Low-Rank Model Order Reduction". Integration, the VLSI Journal, volume 66, May 2019, pp 1–8.

Bibliography

[1] J. Lillis, C. Cheng, S. Lin, and N. Chang, "Interconnect analysis and synthesis". *John Wiley*, 1999.

[2] A. Odabasioglu, M. Celik and L. T. Pileggi, "PRIMA: passive reduced-order interconnect macromodeling algorithm," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 8, pp. 645–654, 1998.

[3] J. Phillips, L. Daniel and L. Miguel Silveira, "Guaranteed passive balancing transformations for model order reduction," in *Design Automation Conference (DAC)*, 2002.

[4] K. Gröchenig, "Foundations of Time-Frequency Analysis," in *Applied and Numerical Harmonic Analysis*, 2001.

[5] Pillage, Lawrence T., and Ronald A. Rohrer. "Asymptotic waveform evaluation for timing analysis," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol 9, no 4, pp 352-366, 1990.

[6] P. Feldmann, R. W. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol 14, no 5, pp 639-649, 1994.

[7] R. W. Freund, "SPRIM: structure-preserving reduced-order interconnect macromodeling," in *International Conference on Computer Aided Design (ICCAD)*, pp. 80–87, 2004.

[8] Hao Yu, Lei He and S. X. D. Tar, "Block structure preserving model order reduction," in *Proceedings of the 2005 IEEE International Behavioral Modeling and Simulation Workshop,* pp. 1-6, 2005.

[9] Z. Zhang, X. Hu, C. Cheng and N. Wong, "A block-diagonal structured model reduction scheme for power grid networks," in *Design, Automation & Test in Europe*, pp. 1-6, 2011.

[10] Pu Liu, Sheldon X.-D. Tan, Boyuan Yan, Bruce McGaughy, "An efficient terminal and model order reduction algorithm," in *Integration*, vol 41 pp. 210–218, 2008.

[11] Pu Liu et al., "An efficient method for terminal reduction of interconnect circuits considering delay variations," in *IEEE/ACM International Conference on Computer-Aided Design,*, pp. 821–826, 2005.

[12] P. Miettinen, M. Honkala, J. Roos and M. Valtonen, "PartMOR: Partitioning-Based Realizable Model-Order Reduction Method for RLC Circuits," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 3, pp. 374–387, 2011.

[13] G. De Luca, G. Antonini and P. Benner, "A parallel, adaptive multi-point model order reduction algorithm," in *IEEE 22nd Conference on Electrical Performance of Electronic Packaging and Systems*, San Jose, CA, 2013, pp. 115-118.

[14] W. Zhao, G. K. H. Pang and N. Wong, "Automatic adaptive multi-point moment matching for descriptor system model order reduction," in *International Symposium on VLSI Design, Automation, and Test*, pp. 1-4, 2013.

[15] A.C. Antoulas "Approximation of large-scale dynamical systems," in *SIAM*, 2005.

[16] A.C. Antoulas "An overview of approximation methods for large-scale dynamical systems," in *Annual Reviews in Control,* , vol. 29, no. 2, pp. 181–190, 2005.

[17] A. Laub, M. Heath, C. Paige and R. Ward, "Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms," in *IEEE Transactions on Automatic Control*, vol. 32, no. 2, pp. 115–122, February 1987.

[18] Z. Zhang, Q. Wang, N. Wong and L. Daniel, "A moment-matching scheme for the passivity-preserving model order reduction of indefinite descriptor systems with possible polynomial parts," in *Asia and South Pacific Design Automation Conference*, pp. 49-54, 2011.

[19] T. Reis and T. Stykel, "PABTEC: Passivity-Preserving Balanced Truncation for Electrical Circuits," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 29, no. 9, pp. 1354-1367, Sept. 2010.

[20] B. Yan, S. X. -. Tan, P. Liu and B. McGaughy, "Passive Interconnect Macromodeling Via Balanced Truncation of Linear Systems in Descriptor Form," in *Asia and South Pacific Design Automation Conference*, pp. 355–360, 2007.

[21] M. Imran, A. Ghafoor and M. Imran, "Frequency Limited Model Reduction Techniques With Error Bounds," in *IEEE Transactions on Circuits and Systems II: Express Briefs,* vol. 65, no. 1, pp. 86–90, Jan. 2018.

[22] U. Zulfiqar, M. Imran, A. Ghafoor and M. Liaquat, "A New Frequency-Limited Interval Gramians-Based Model Order Reduction Technique," in *IEEE Transactions on Circuits and Systems II: Express Briefs,* vol. 64, no. 6, pp. 680–684, 2017.

[23] D. Li, S. X. -. Tan and B. McGaughy, "ETBR: Extended Truncated Balanced Realization Method for On-Chip Power Grid Network Analysis," in *Design, Automation and Test in Europe,* pp. 432–437, 2008.

[24] v. Simoncini, "Computational Methods for Linear Matrix Equations," in *SIAM Review,* vol. 58, no. 3, pp. 337–441, 2016.

[25] V. Simoncini, "A New Iterative Method for Solving Large-Scale Lyapunov Matrix Equations, " in *SIAM Journal on Scientific Computing*, vol. 29, no. 3, pp. 1268–1288, 2007.

[26] J. R. Li and J. White, "Low Rank Solution of Lyapunov Equations," in *SIAM Journal on Matrix Analysis and Applications*, vol. 24, no. 1, pp. 260–280, 2002.

[27] J. Phillips and L. M. Silveira, "Poor man's TBR: a simple model reduction scheme," in *Design, Automation and Test in Europe Conference and Exhibition (DATE)*, 2004.

[28] V. Vasudevan and M. Ramakrishna, "An efficient algorithm for frequency-weighted balanced truncation of VLSI interconnects in descriptor form," in *Design Automation Conference (DAC)*, 2015.

[29] S. Gugercin and A. C. Antoulas, "A Survey of Model Reduction by Balanced Truncation and Some New Results," in *International Journal of Control,*, vol. 77, no. 8, pp. 748–766, 2004.

[30] W. Gawronski and J. Juang, "Model reduction in limited time and frequency intervals," in *International Journal of Systems Science*, vol. 21, no. 2, pp. 349–376, 1990.

[31] P. Benner, P. Kurschner and J. Saak, "Frequency-Limited Balanced Truncation with Low-Rank Approximations," in *SIAM Journal on Scientific Computing*, vol. 38, no. 1, pp. 471–499, 2016.

[32] S. Tan and L. He, *Advanced model order reduction techniques in VLSI design*, Cambridge University Press, 2007.

[33] B. Yan, S. X. -D. Tan, L. Zhou, J. Chen and R. Shen, "Decentralized and Passive Model Order Reduction of Linear Networks With Massive Ports," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 5, pp. 865-877, 2012.

[34] P. Feldmann, "Model order reduction techniques for linear systems with large numbers of terminals," in *Proceedings Design, Automation and Test in Europe Conference and Exhibition (DATE)*, 2004.

[35] P.Li and W. Shi, "Model order reduction of linear networks with massive ports via frequency-dependent port packing," in *ACM/IEEE Design Automation Conference (DAC)*, 2006.

[36] B. Yan, S. X. -D. Tan and B. McGaughy, "Second-Order Balanced Truncation for Passive-Order Reduction of RLCK Circuits," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 9, pp. 942-946, Sept. 2008.

[37] D. Li, S. X.-D. Tan, and B. McGaughy, "ETBR: extended truncated balanced realization method for on-chip power grid network analysis," in *Proceedings of the conference on Design, automation and test in Europe (DATE)*, 2008.

[38] P. Heydari and M. Pedram, "Model-order reduction using variational balanced truncation with spectral shaping," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 4, pp. 879-891, April 2006.

[39] A. Ghafoor and V. Sreeram, "Model Reduction Via Limited Frequency Interval Gramians," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 9, pp. 2806–2812, Oct. 2008.

[40] A.C. Antoulas, D.C. Sorensen, Y. Zhou, "On the decay rate of Hankel singular values and related issues," in *Systems and Control Letters*, vol. 46, no. 5, pp. 323–342, 2002.

[41] K. Jbilou and A.J. Riquet, "Projection methods for large Lyapunov matrix equations, " in *Linear Algebra and its Applications*, vol. 415, no. 2-3, pp. 344–358, 2006.

[42] L. Knizhnerman and V. Simoncini, "Convergence analysis of the extended Krylov subspace method for the Lyapunov equation, " in *Numerische Mathematik*, vol. 118, no. 3, pp. 567–586, 2011.

[43] Stykel, T., "Gramian-Based Model Reduction for Descriptor Systems" *T. Math. Control Signals Systems*, vol 16, pp 297–319, 2004.

[44] UMFPACK, http://www.cise.ufl.edu/research/sparse/umfpack/

[45] Barrett, R. and Berry, M. and Chan, T. and Demmel, J. and Donato, J. and Dongarra, J. and Eijkhout, V. and Pozo, R. and Romine, C. and van der Vorst, H, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, 1994.

[46] G. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1996.

[47] N. Higham, *Functions of Matrices: Theory and Computation*, SIAM, 2008.

[48] N. Lang, H. Mena and J. Saak, "On the benefits of the LDL^T factorization for large-scale differential matrix equation solvers," in *Linear Algebra and its Applications,* vol. 480, pp. 44–71, 2015.

[49] G. Floros, N. Evmorfonoulos and G. Stamoulis, "Efficient Circuit Reduction in Limited Frequency Windows," in *nternational Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD),* pp. 129–132, 2019.

[50] P. Benner et al, "Self-Generating and Efficient Shift Parameters in ADI Methods for Large Lyapunov and Sylvester Equations," in *Electronic Transaction on Numerical Analysis,* vol. 43, pp. 142–162, 2014.

[51] L. Knizhnerman and V. Simoncini, "A new investigation of the extended Krylov subspace method for matrix function evaluations ," in *Numerical Linear Algebra Appl.,* vol. 17, pp. 615-638, 2010.

[52] F. D. Freitas, J. Rommes and N. Martins, "Gramian-Based Reduction Method Applied to Large Sparse Power System Descriptor Models," in *IEEE Transactions on Power Systems,* vol. 23, no. 3, pp. 1258-1270, 2008.

[53] S. Grivet-Talocia and A. Ubolli, "A Comparative Study of Passivity Enforcement Schemes for Linear Lumped Macromodels," in *IEEE Transactions on Advanced Packaging,* vol. 31, no. 4, pp. 673-683, 2008.

[54] http://slicot.org/20-site/126-benchmark-examples-for-model-reduction

[55] R. Ionutiu, J. Rommes and W. H. A. Schilders, "SparseRC: Sparsity Preserving Model Reduction for RC Circuits With Many Terminals, " in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 30, no. 12, pp. 1828-1841, 2011.

[56] C. Xu, S. K. Kolluri, K. Endo and K. Banerjee, "Analytical Thermal Model for Self-Heating in Advanced FinFET Devices With Implications for Design and Reliability," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 7, pp. 1045–1058, 2013.

[57] "International technology roadmap for semiconductors (ITRS) 2015 Edition -ERD," 2015.

[58] Peng Li, L. T. Pileggi, M. Asheghi and R. Chandra, "IC thermal simulation and modeling via efficient multigrid-based approaches," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 9, pp. 1763–1776, 2006.

[59] S. Ladenheim , Y. C. Chen, M. Mihajlovic, V. Pavlidis, "IC thermal analyzer for versatile 3-D structures using multigrid preconditioned Krylov methods," in *IEEE/ACM International Conference on Computer-Aided Design*, pp. 1–8, 2016.

[60] Y. Zhan and S. S. Sapatnekar, "High-Efficiency Green Function-Based Thermal Simulation Algorithms," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 19, pp. 1661–1675, 2007.

[61] L. Codecasa, D. D'Amore and P. Maffezzoni, "Compact modeling of electrical devices for electrothermal analysis," in *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,*, vol. 50, no. 4, pp. 465-476, April 2003.

[62] A.C. Antoulas, D.C. Sorensen, and S. Gugercin, "A Survey of Model Reduction Methods for Large-Scale Systems," in *Contemporary Mathematics.*, vol. 280, pp. 193–201, 2001.

[63] G. Floros, N. Evmorfopoulos and G. Stamoulis, "Efficient Hotspot Thermal Simulation Via Low-Rank Model Order Reduction," in *2018 15th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, Prague, 2018, pp. 205-208.

[64] G. Floros, N. Evmorfopoulos and G. Stamoulis, "Efficient IC hotspot thermal analysis via low-rank Model Order Reduction," in *Integration,*, vol 66, pp. 1–8, 2019.

[65] Peng Li, L. T. Pileggi, M. Asheghi and R. Chandra, "Efficient full-chip thermal modeling and analysis," in *IEEE/ACM International Conference on Computer Aided Design.*, pp. 319–326, 2004.

[66] Y. Yang, Z. Gu, C. Zhu, R. P. Dick and L. Shang, "ISAC: Integrated Space-and-Time-Adaptive Chip-Package Thermal Analysis," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 1, pp. 86–99, 2007.

[67] T. Y. Wang and C. C. P. Chen, "3-D Thermal-ADI: a linear-time chip level transient thermal simulator," in *IIEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 12, pp. 1434–1445, 2002.

[68] Y. Yang, C. Zhu, Z. Gu and L. Shang and R. P. Dick, "Adaptive multi-domain thermal modeling and analysis for integrated circuit synthesis and design," in *IEEE/ACM International Conference on Computer Aided Design.*, pp. 575–582, 2006.

[69] A. Sridhar, A. Vincenzi, D. Atienza and T. Brunschwiler, "3D-ICE: A Compact Thermal Model for Early-Stage Design of Liquid-Cooled ICs," in *IEEE Transactions on Computers*, vol. 63, no. 10, pp. 2576–2589, 2014.

[70] X. X. Liu, K. Zhai, Z. Liu, K. He, S. X. D. Tan and W. Yu, "Parallel Thermal Analysis of 3-D Integrated Circuits With Liquid Cooling on CPU-GPU Platforms," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 3, pp. 575–579, 2015.

[71] G. Floros, K. Daloukas, N. Evmorfopoulos and G. Stamoulis, "A parallel iterative approach for efficient full chip thermal analysis," in *7th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, Thessaloniki, pp. 1–4, 2018.

[72] G. Floros, K. Daloukas, N. Evmorfopoulos, G. Stamoulis, "A Preconditioned Iterative Approach for Efficient Full Chip Thermal Analysis on Massively Parallel Platforms," in *Technologies* 2019, 7, 1.

[73] A. Vincenzi, A. Sridhar, M. Ruggiero and D. Atienza, "Fast thermal simulation of 2D/3D integrated circuits exploiting neural networks and GPUs," in *IEEE/ACM International Symposium on Low Power Electronics and Design.*, pp. 151–156, 2011.

[74] Y. M. Lee, C. W. Pan, P. Y. Huang and C. P. Yang, "LUTSim: A Look-Up Table-Based Thermal Simulator for 3-D ICs," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems.*, vol. 34, no. 8, pp. 1250–1263, 2015.

[75] T. Y. Wang and C. C. P. Chen, "SPICE-compatible thermal simulation with lumped circuit modeling for thermal reliability analysis based on modeling order reduction," in *International Symposium on Signals, Circuits and Systems*, pp. 357–362, 2004.

[76] P. Liu, Z. Qi, H. Li, L. Jin, W. Wu, S. X. D. Tan and J. Yang, "Fast thermal simulation for architecture level dynamic thermal management," in *IEEE/ACM International Conference on Computer-Aided Design.*, pp. 639–644, 2005.

[77] N. Ouzisik, "Heat transfer - A basic approach," in *Mcgraw-Hill College book company*, 1985.

[78] T. Bergman, B. Lavine, P. Incropera and P. DeWitt, "Fundamentals of Heat and Mass Transfer. ," in *Wiley New York*, 2017.

[79] Y.-K. Cheng, P. Raha, C.-C. Teng, E. Rosenbaum and S.-M. Kang, "ILLIADS-T: an electrothermal timing simulator for temperature-sensitive reliability diagnosis of CMOS VLSI chips," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 8, pp. 668-681, 1998.

[80] K. Jbilou and A.J. Riquet, "Projection methods for large Lyapunov matrix equations, " in *Linear Algebra and its Applications*, vol. 415, no. 2-3, pp. 344–358, 2006.

[81] D. Sorensen and Y. Zhou, "Direct methods for matrix Sylvester and Lyapunov equations, " in *J. Appl. Math.,* 2003, pp. 277-303.