Lehrstuhl für Hochfrequenztechnik
der Technischen Universität München

Model Order Reduction (MOR)

in Transmission Line Matrix (TLM) Method

Dzianis Lukashevich

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor–Ingenieurs

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Ph.D. Paolo Lugli
Prüfer der Dissertation:	1. Univ.-Prof. Dr. techn. Peter Russer 2. Prof. Dr. Andreas C. Cangellaris, University of Illinois at Urbana-Champaign, USA

Die Dissertation wurde am 21.12.2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 04.05.2005 angenommen.

Abstract

The Finite-Difference Time-Domain (FDTD) method, the Finite-Integration Technique (FIT) and the Transmission Line Matrix (TLM) method provide for discrete approximations of electromagnetic boundary value problems cast in state-space forms. The dimension of the generated state-space models is usually very large. In general terms, Model Order Reduction (MOR) enhances computational efficiency. Application of the reduced-order modeling to the FDTD, FIT and TLM methods yields considerable reduction in the computational effort necessary for the solution of the discrete models. Furthermore, MOR can be used to generate compact, broadband discrete models of the original electromagnetic systems. Such compact broadband macro-modeling enables the abstraction of the discretized electromagnetic system in terms of a frequency-dependent transfer function matrix representation, which, in turn, provides for efficient implementation of the model in both general-purpose network-analysis oriented simulators, and full-wave, time- and frequency-domain electromagnetic field solvers aimed at system-level electromagnetic modeling.

Reduced-order modeling in TLM is an approximation of the discrete model of electromagnetic field obtained through the application of the TLM method, in terms of a model of considerable lower dimension. The development of such a reduced-order model (ROM) can be achieved using two classes of methods, namely, singular value decomposition (SVD) methods and moment matching methods. A representative member of the former class of methods is the so-called balanced model reduction, which is aimed at removing from the original system those eigenstates that are difficult to observe and control. SVD-based MOR-techniques have the attractive attribute that bounds for the approximation error in the reduced model can be established. One disadvantage is that they tend to be computationally more expensive than moment matching methods.

The main focus of this work is on moment matching methods. While techniques for MOR based on moment matching have been studied extensively in the case of FDTD and FIT methods, they have not yet been considered in detail in the context of TLM approximations of electromagnetic systems. The application of moment matching MOR-techniques to TLM is presented, with emphasis placed on KRYLOV subspace methods that utilize the LANCZOS and ARNOLDI processes. The attributes of such methods, both in terms of computational efficiency and solution accuracy, are examined through their application to the analysis of several electromagnetic structures. For this purpose the implicit time evolution TLM-scheme is utilized. Also, a comparison is provided between FDTD, FIT and TLM with regards to the computational efficiency of MOR-methods for expediting their numerical integration.

Dedicated to my parents
Прысвечана маім бацькам

Acknowledgment

First of all, I would like to thank PROF. PETER RUSSER for the opportunity to prepare my dissertation in the Institute for High-Frequency Engineering at Munich University of Technology and for his guidance on the scientific work presented here. He has offered me certain freedom in choosing my research field, and he always had an open ear for any kinds of problems. I am also very grateful for the financial support and for the possibility to travel to several conferences for presenting parts of this work.

I would further like to thank PROF. ANDREAS CANGELLARIS for his great support during my stay for three months in Urbana-Champaign at the Department of Electrical and Computer Engineering at the University of Illinois, USA in 2003. He has been a constant source of advice and motivation since the day I met him. It was a great pleasure to work together with him and to prepare several collaborated papers.

My special thanks are to my colleague and friend KLAUS HEPPENHEIMER for numerous discussions on the theoretical aspects of the TLM method and fundamental questions referred to it. He always has been opened for discussions and he always shared his knowledge on the theory, physics and mathematics of electromagnetism. He always patiently corrected my mistakes in German whilst we worked together for three years.

Next, I thank my friend MARKUS BURGER for his critical remarks to a number of my ideas and for any kind of help I needed. My thanks are to my friend BORYS BROIDO for always to be there for me. I am grateful to GABRIELA PREINDL for giving an initial motivation to do my doctorate in Munich, for her support and for teaching me German. I would also like to thank ROSALIA BROIDO for taking care of me during my stay in Munich.

I thank all colleagues from the Institute for High-Frequency Engineering for their stimulating discussions and friendly collaboration. Among whom I would like to thank especially DR. GERHARD OLBRICH for solving many organizing and administrative problems. Thanks to all of you for your friendship!

I am grateful to all the people who helped me to prepare this manuscript. My special thanks are to CARMEN OKHMATOVSKI for her careful proofreading some parts of my thesis.

Вялікі дзякуй маім бацькам за іхняе разуменне і ўсялякую падтрымку праз усе гады і, у першую чаргу, за маё існаванне. Бязмежная ўдзячнаць і хвала Вам!

Dzianis Lukashevich
Munich, Germany
in December 2004

Contents

Chapter 1

Introduction

Over the past few years Model Order Reduction (MOR) is receiving significant attention by the computational electromagnetics community as a systematic and effective method for the development of compact representations of passive electromagnetic systems of high complexity. Since the modeling of an electromagnetic system involves the spatial discretization of MAXWELL's equations over the computational domain of interest, the number of state variables tends to be very large for systems exhibiting high geometric complexity. For example, discrete electromagnetic models of very large dimension result from the discretization of structures that contain very fine (sub-wavelength) geometric features, yet their spatial extent spans several wavelengths in one or more dimensions in space. For such systems, MOR allows for significant increase in the efficiency of the numerical solution of the discrete model through the construction of reduced-order models (ROMs) with a number of state variables much smaller than that of the original system. Furthermore, for those cases for which the structure under modeling happens to be a component of a larger system, the construction of a reduced-order macro-model of the component helps expedite the simulation of the overall system. More specifically, the generated reduced-order macro-model facilitates the abstraction of a multi-port component of the original system in terms of a frequency-dependent, transfer function matrix representation. In this manner, a decomposition of the system in terms of low-order transfer function matrices is achieved. Furthermore, the elements of these transfer functions are cast in terms of rational functions of the angular frequency, thus enabling their seamless and efficient integration in both general-purpose, network-analysis oriented circuit simulators and full-wave electromagnetic solvers aimed at system-level modeling.

The Transmission Line Matrix (TLM) method has proven to be a powerful tool for solving MAXWELL's equations and has been successfully applied to the analysis of various complicated planar and three-dimensional electromagnetic structures. Like related differential equation-based numerical methods, the TLM approximation of a linear electromagnetic boundary value problem results a system of linear equations of large dimension. For practical electromagnetic devices the dimension of the resulting models is in the order of millions. Since higher spatial resolution requires a shorter time step to guarantee stability, the solution cost may easily become overwhelming and eventually prohibitive. Therefore, the exploitation of MOR as a means of expediting the TLM solution process makes sense.

1.1 State of the Art

The Transmission Line Matrix (TLM) method, developed and first published in 1971 by JOHNS and BEURLE [60] has emerged as a powerful method for computer modeling of electromagnetic fields [23],[56],[99]. The TLM method exhibits excellent numerical stability and is also suitable for the modeling of lossy, dispersive, active and nonlinear media [81], [101],[110]. However, the state-space representation of TLM is usually very large, requiring significant computational resources and long computation times. It is the objective of the proposed TLM Model Order Reduction (TLM-MOR) to address these computational complexity difficulties of standard TLM through the reduction of the original model into one of significantly smaller dimension. The dimension of the reduced model, also called model order, is associated with a subset of the eigenmodes of the system that influence its response over the desirable frequency bandwidth of interest. Furthermore, in the pole-residue representation of the generated ROM transfer function, the model order is equal to the number of complex natural frequencies or poles of the transfer function.

One of the early applications of MOR in numerical electromagnetics was its use for the expedient calculation of the broadband response of an electromagnetic device (see, for example, [12],[129]). However, the most attractive attribute of model order reduction is its close relationship to the development of compact, broadband *macro-models* of electromagnetic devices, described in terms of rational function approximations of the associated multi-port transfer functions [14],[43]. Availability of such macro-models facilitates highly accurate, linear or non-linear simulation at the subsystem and system level by means of general-purpose, network analysis-oriented transient simulators such as SPICE [36],[85]. Once cast in the form of such rational function macro-model, the matrix transfer function of the electromagnetic device being modeled can be interfaced conveniently and seamlessly with other frequency- and/or time-domain electromagnetic and electromagnetic/circuit solvers. A particularly attractive special case of such an implementation is in conjunction with the development of effective, broadband and computationally efficient sub-gridding strategies in the finite difference modeling of complex electromagnetic structures [35], [69], [68].

The development of such a reduced-order model can be achieved using two classes of methods, namely,

- singular value decomposition (SVD) methods, and
- moment matching methods.

A representative member of the former class of methods is the so-called balanced model reduction, aimed at removing from the original system those eigenstates that are difficult to observe and control [82]. SVD-based model order reduction methods have the attractive attribute that bounds for the approximation error in the reduced model can be established. One disadvantage is that they tend to be computationally more expensive than moment matching methods. An assessment of the state of the art in the development of SVD-based model order reduction methods can be found in [4]. The SVD-technique with SCHUR decomposition has been already adopted to the TLM method [80]. However, the huge size of the TLM state-space system restricts the application

of the proposed approach to 2-D cases with a small number of discretization cells.

The key attribute of moment matching techniques is the efficiency of their numerical implementation. The Asymptotic Waveform Evaluation (AWE) method, based on the explicit moment matching technique, has been used for analysis of large linear networks and interconnect structures (see, for example [9], [111]). Recently, the AWE technique has been adopted to the state-space systems resulting from approximations of differential equation-based description of electromagnetic boundary value problems [11],[129]. From the available moment matching MOR-techniques KRYLOV subspace methods appear to be the most robust and effective. More specifically, the LANCZOS- and ARNOLDI-algorithm based MOR are most commonly used to expedite the solution of large linear circuits, and FDTD and FIT approximations of electromagnetic problems [16], [84],[85],[96]. One of the best-known MOR techniques is PADÉ-VIA-LANCZOS (PVL), which provides a PADÉ approximation for the TAYLOR expansion of the transfer function around a point in the LAPLACE domain [41], [84]. The PVL matrix approximation is becoming a common approach for the Fast Frequency Sweep (FFS) in Finite-Element Method (FEM) and Method of Moments (MoM) based commercial simulators [17], [127]. The direct application of the LANCZOS algorithm to a MOR process yields a PADÉ approximation for the expansion point placed at infinity. The approximations of the transfer function by KRYLOV subspace methods with expansion at either a finite complex-frequency point or at infinite frequency have advantages and drawbacks, which can be briefly summarized as follows:

- KRYLOV *subspace methods with finite expansion point* (classical PVL and/or PADÉ-VIA-ARNOLDI (PVA)) - low order reduced models, high computational time effort and memory requirements in order to obtain the inverse system matrix;
- KRYLOV *subspace methods with expansion at infinity* (direct application of Krylov subspace methods) - low computational effort, large dimension of reduced-order models, rapid increase of computational costs for post-processing due to the large dimension of the reduced system matrix.

In order to exploit their advantages these approaches can be easily integrated into a two-step reduction process. This combination yields a low-order reduced model with the reasonable computational effort [126].

Once a reduced-order model is available, generated by means of applying KRYLOV subspace methods to TLM, the time-domain response of the electromagnetic structure under investigation can be recovered expediently. In principle, the transient response of the electromagnetic structure may be also predicted efficiently from a limited number of time samples of the TLM-generated time series through the application of high-resolution parametric model estimation techniques, based on System Identification (SI) techniques [87]. These methods allow the expedition of TLM simulations and the determination of the network equivalent model directly from the simulated results [24], [25]. Although SI techniques have been applied extensively for time extrapolation of numerically obtained electromagnetic responses, a systematic and efficient approach for their robust implementation in conjunction with transient electromagnetic field solvers is still missing. SI methods can be regarded as digital signal post-processing based methods for obtaining the approximate impulse response of

the device under modeling. On the other hand, moment matching MOR techniques are aimed at providing for such impulse response extraction capability not through post-processing of the numerical solution process but during the solution process.

In particular, MOR-techniques are attractive for expedient analysis of electrically large, highly-resonant structures, like filters and resonators. Time-domain based simulators require long time iterations to achieve steady state, and frequency-domain methods demand fine frequency resolutions in order to gain sharp resonances. In both cases, the required solution time is very long. In the RF area, for example, reduced-order models are used to capture the behavior of a complex interconnect system in the form of a smaller RF macro-model, which is then embedded into the main circuit and co-simulated with it [15], [122]. Similar approaches are used in other application areas, such as the design of multi-port Microwave Monolithic Integrated Circuits (MMICs), optical transceiver ICs and electromagnetic interference (EMI). Recently, MOR-techniques have been applied successfully to distributed, passive structures such as connectors, transmission lines and antennas. The ARNOLDI methods are used to generate accurate reduced-order macro-models for coupled energy domain nonlinear Micro-Electro-Mechanical Systems (MEMS) [18], [19]. Obtained results demonstrate that the reduced nonlinear model offers much better accuracy in capturing the original device behavior than the simple linearization method. A computationally inexpensive approach to nonlinear model reduction, based on representing a nonlinear system with a piecewise-linear system and then reducing each of the pieces with a KRYLOV projection, is described in [97].

Parameter optimization based on MOR-techniques has been suggested in [64]. This significantly reduces the number of simulation runs, because poles and residues of a reduced model are presented as linearized function of its geometry parameters, and actual optimization is performed on the reduced-order model; hence, the model itself is updated after each optimization step.

In summary, an increasing level of research has been reported in the MOR field over the past decade addressing three primary issues:

- reduced-order model accuracy,
- numerically stable and computationally efficient generation of models of arbitrary order,
- generation of macro-models that are "well-behaved" when embedded into a system-level simulation tool for co-simulation with other models for the remaining part of the system under analysis.

1.2 Definition of the Problem

The moment matching MOR-techniques have been studied extensively in the case of FDTD and FIT (e.g. see [16], [84], [96]); however they have not yet been considered in detail in the context of TLM approximations of electromagnetic systems.

This work focuses on the application of moment matching MOR-techniques to TLM with emphasis placed on KRYLOV subspace methods based on the

LANCZOS and ARNOLDI processes. The KRYLOV subspace methods are attractive for MOR in the TLM framework because the construction of the orthonormal and/or bi-orthogonal bases used for the Krylov subspaces can be carried out through the direct application of the implicit TLM solver, i.e. the TLM operator need not be known explicitly, as only its impact on the sequence of KRYLOV vectors must be taken into account, whereas MOR in context of FDTD and FIT usually deals with system matrices in an explicit form. Thus, the computation of a reduced model can be integrated in the usual iterative TLM process in a very efficient manner.

The TLM method differs from conventional finite difference schemes, e.g. FDTD, in the sense that its kernel is a discrete time evolution scheme. So, reciprocity and energy conservation of a loss-free model provide the skew-symmetry (or/and the symmetry via some modifications [96], [125]) of FDTD system matrix; however, they do not yield symmetry of the discrete time evolution TLM-operator. Consequently, the symmetric LANCZOS algorithm cannot be applied and the use of more expensive general ARNOLDI or LANCZOS procedures becomes necessary.

The main goals to be set and achieved by MOR applied to TLM are:

- acceleration of TLM simulations and enhancement of computational efficiency,
- generation of a compact macro-model of the electromagnetic structure under investigation.

A generated macro-model can by incorporated into other simulators by means of its matrix transfer function representation in one of several possible forms (e.g., General Impedance Matrix (GIM) $\boldsymbol{Z}$, scattering matrix $\boldsymbol{S}$ or multi-port impulse response $\tilde{\boldsymbol{\mathcal{H}}}$). Thus, we seek a model of reduced order that approximates the impulse response $\tilde{\boldsymbol{\mathcal{H}}}$ or (and/or GIM $\boldsymbol{Z}$) or, in an equivalent sense, the initial TLM-system. This may be done in an efficient manner using moment matching techniques based on KRYLOV subspace methods. The basic idea is to project the initial TLM-system of large dimension N onto a space of significantly lower dimension n, constructed in such a manner that only the first n moments in the LAURENT series expansions of the impulse response (and/or GIM) of the original and the reduced systems are matched [14], [100]. The generated reduced-order model contains, predominantly, those eigenstates that correspond to the physical eigenfrequencies of an electromagnetic system and, in particular, those that impact the states involved in the excitation and observation of the system.

1.3 Overview

KRYLOV subspace based MOR-techniques in the context of the TLM framework are discussed in detail and evaluated in this work. The present doctoral thesis is organized as follows.

In Chapter 2, the principles of the TLM method with symmetrical condensed nodes for the solution of three-dimensional electromagnetic problems are presented. The HILBERT space formulation of the TLM is given, using the DIRAC notation; thus the TLM formulation is expressed through operator equations.

The scattering, connection and general TLM operators in the TLM time evolution scheme for both lossless and lossy structures in unbounded and bounded domains are described. Some modifications of the TLM scattering algorithm, introduced originally by JOHNS, and field mappings needed for MOR-techniques are also discussed in this Chapter.

In Chapter 3, the novel and effective hybrid Transmission Line Matrix-Mode Matching (TLM-MM) approach is proposed for full-wave analysis of passive transmission structures in MMICs. Using the hybrid TLM-MM approach, time and memory requirements for structure response generation can be reduced in comparison to pure TLM simulations, while improving solution accuracy. A comparison between both methods by means of numerical simulations is provided. In addition, this method can be easily extended to facilitate macro-modeling of MMICs, in the sense that through its application the generation of a reduced-order macro-model can be simplified and accelerated.

Chapter 4 provides an overview of the published theoretical work on KRYLOV subspace methods for iterative solution of large linear eigensystems, such as the ones arising in circuit simulation and, in general, state-space representations of large-scale time-invariant linear systems. These iterative techniques, based on the orthogonal and/or oblique projection onto KRYLOV subspaces, allow to reduce an arbitrary general matrix to a condensed form. We also review variants of the ARNOLDI and LANCZOS algorithms for generating basis vectors for KRYLOV subspaces. The main focus is on the robust algorithms for the solution of general non-HERMITIAN eigenvalue problems.

In Chapter 5, specific methodologies for MOR, suitable for the discretized electromagnetic TLM system, are discussed. The opportunities for enhancing the computational efficiency of TLM through the application of MOR-techniques are considered in detail. The emphasis of the reduced-order modeling is on the nonsymmetric version of the LANCZOS algorithm and its modifications. Beneficial properties of the TLM-matrix allow to qualify the nonsymmetric LANCZOS procedure so that a novel algorithm, called scattering-symmetric LANCZOS algorithm, results, in which the computational cost per iteration is cut in half. The presentation begins with the general mathematical framework for MOR of discrete electromagnetic systems. This is followed by the development of TLM-MOR techniques and their associated algorithms. The TLM reduced-order models in the z-domain and LAPLACE domain are presented here. Thereafter, we give a comparison of the FDTD-, FIT- and TLM-based MOR-techniques. The two-step reduction approach is described in details as well.

Chapter 6 contains numerical studies from the application of MOR to the TLM analysis of various lossless and lossy electromagnetic structures. The computational efficiency of TLM through the application of MOR-techniques based on the ARNOLDI and LANCZOS processes is discussed there. The convergence properties of proposed algorithms and the accuracy of generated reduced-order models are considered in detail. Finally, results are presented from the application of TLM-MOR for the extraction of matrix transfer function representations of the reduced-order model, such as generalized impedance or scattering-parameter matrices.

Chapter 7 concludes this doctoral thesis by summarizing and discussing its main results.

Chapter 2

The Transmission Line Matrix Method

The basics of the Transmission Line Matrix (TLM) method with symmetrical condensed node for the solution of three-dimensional electromagnetic problems are presented in this chapter. After a brief overview of different derivations of the TLM and its variants the original form of JOHNS [60] is described. The HILBERT space formulation of the TLM is introduced using the DIRAC notation; the TLM method is expressed by operator equations. The scattering, connection and general TLM operators in the TLM time evolution scheme for both lossless and lossy structures in boundless and bounded propagation media are described here. Some modifications of the TLM scattering algorithm and field mappings needed for Model Order Reduction (MOR) techniques are introduced. Influence of various boundary conditions on the convergence of MOR-algorithms is summarized here as well.

2.1 Foundations of the TLM Method

The TLM method is a powerful time-domain method for computer modeling of electromagnetic fields. The TLM is founded on the propagation of electromagnetic waves according to the HUYGENS principle and, in general, based on the analogy between the discretized electromagnetic field and a mesh of transmission lines and nodes. The space is discretized by the subdivision into rectangular cells and the tangential components of the electromagnetic field are sampled at the center of each boundary surface of a cell (or at the cell center according to JOHNS), at so-called ports according to Fig. 2.1. The transmission lines connect the center of a cell (the node) with its ports at the boundary surfaces. The continuous space is approximated by a mesh of TLM nodes interconnected by transmission lines (or by a mesh of transmission lines interconnecting TLM nodes). The time evolution of the electromagnetic field is modeled by wave pulses propagating between adjacent cells and scattered within the cells, i.e. the wave pulses are scattered at the nodes and these scattered pulses are propagated through the transmission lines to the adjacent nodes, where these pulses are scattered again. Thus, the TLM method is a time-space-discretizing method and closely related to the Finite Difference Time-Domain (FDTD) method. Roughly speaking, discretized electromagnetic fields are mapped onto wave amplitudes propagating

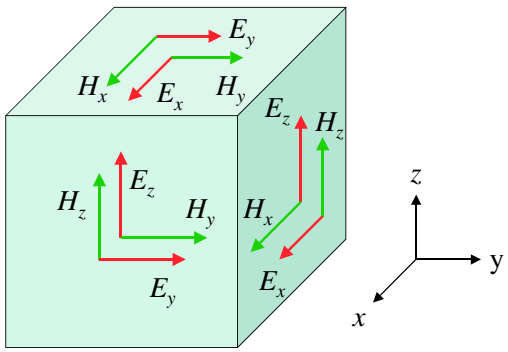


Figure 2.1: The TLM cell.

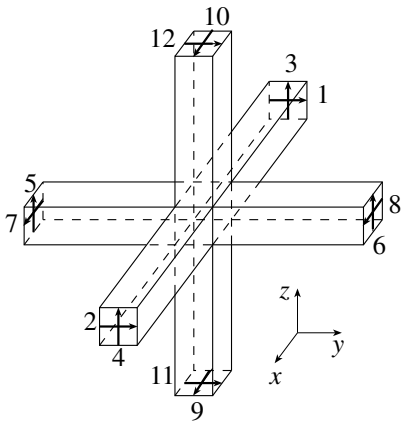


Figure 2.2: The three-dimensional SCN.

in the TLM-mesh in accord with some rules.

In general, the TLM cell is represented by a respective abstract network model containing a scattering center with six stubs which is connected via twelve transmission lines with ports of adjacent cells. When all transmission lines have the same lengths and the node is located in the cell center this model is called as the *Symmetrical Condensed Node* (SCN). The lossless three-dimensional SCN shown in Fig. 2.2 was introduced by JOHNS in 1987 [60] first.

On the basis of the JOHNS formulation of TLM some generalizations and modifications of the SCN are proposed. So the *Hybrid Symmetrical Condensed Node* (HSCN) proposed in [104] has three instead of six stubs, the other three stubs can be modeled through the variation of transmission line parameters. The *General Symmetrical Condensed Node* (GSCN) introduced in [115], [119] does not have stubs at all and allows to save additional memory. To this class of the symmetrical condensed nodes belong the *Symmetrical Super-Condensed Node* (SSCN) and the *Adaptable Symmetrical Condensed Node* (ASCN) given in [115], [116] and [117] with variable stub arrangements and transmission line parameters. A different kind of the TLM method are the *Alternating Transmission Line Matrix* method [6] reducing the number of cells by a factor of two and the *Alternating Rotated Transmission Line Matrix* method [98] decreasing

the numerical effort for homogenous media by a factor of two through a transformation of the scattering matrix. However, these alternating methods are not used widely due to their complicated implementation. First proposals of the TLM with unstructured two-dimensional meshes upon triangular elements can be found in [107]. The generalization of the TLM for a non-orthogonal mesh and for dispersive anisotropic materials is presented in HEIN'S works [50], [51], [52]. Dispersive [81], active and nonlinear media [101] can be well modeled by the SCN. The TLM in the frequency-domain is proposed also in [59]. Almost radiating boundary conditions, e.g. reflection free walls, matched [83] and perfectly matched layers [38], [90] are already used in the TLM, which yield approximate solutions of a radiating problem and require some distance between the structure to be modeled and boundaries of the computational domain. A novel TLM Multipole Expansion (TLM-ME) method for a potentially exact modeling of the radiating boundary condition is proposed in [72], where an electromagnetic structure is embedded in a spherical region and modeled with the TLM method. Then outside the spherical region the field is expanded into analytically known spherical waves. The problems of the local subgridding in the TLM scheme are investigated in [37] and [55]. Hybridizations of the TLM method with other numerical methods such as the TLM-Integral Equation (TLM-IE) method [92] for the solving of radiation problems and the TLM-Mode Matching (TLM-MM) approach [75] for a full-wave analysis of transmission structures in multilayered MMICs occurred to be very efficient. The improved skin effect [58] and thin wire [70] models allowed to spread the usage of the TLM additionally.

While the original derivation of the TLM with the SCN by JOHNS is mainly based on the analogy between the electromagnetic wave propagation in free space and the propagation of voltages and currents in the TLM mesh, several formal derivation are suggested in the literature. So the Method of Moments (MoM) is applied to the TLM without stubs by KRUMPHOLZ and RUSSER [67], [65] in order to derive fundamental TLM equations. In addition, the authors proposed the *Cell Boundary Mapping* between the electromagnetic field and wave amplitudes. This bijective mapping requires four tangential field components at a cell wall only in contrast to the non-bijective JOHNS *Center Field Mapping* with six field components in the cell center. CHEN, NEY and HOEFER presented a new finite-difference formulation in [20], which is equivalent to the TLM with SCNs without stubs. In order to prove the SCN-TLM without stubs LOVETRI and SIMONS started from MAXWELL'S curl equations written as a system of conservation laws [73]. Starting from the integral formulation of MAXWELL'S equations AIDAM and RUSSER obtained the lossless TLM basic equations in [1], [2], and TARDIOLI and HOEFER presented a derivation of the scattering matrix for the SCN without stubs [112]. JIN and VAHLDIECK derived the TLM method with stubs from MAXWELL'S equations using centered differencing and averaging. HEIN'S derivation in [50] trough a discrete propagator integral leads to a slightly different interpretation of the TLM algorithm. The convergence of the classical TLM with SCNs is proved by REBEL, AIDAM and RUSSER in [95], [94] for the first time. The dispersion behavior of the TLM with various nodes is thoroughly studied in TRENKIC'S works [115] and [118].

The electromagnetic field and wave amplitude states in the TLM are represented by an enumerable set of real quantities and the field evolution is governed by linear mapping rules. For these reasons, the HILBERT space representation of the field state and the field evolution is possible [67], [99]. So, the field map-

pings, propagation and the scattering of the wave amplitudes may be expressed in the operator notation in the HILBERT space.

As a basis for the present dissertation the JOHNS original form of the TLM method with SCNs is preferred. The MOR-techniques suggested in this doctoral thesis can be easily expanded on the TLM with HSCNs, GSCNs, SSCNs and ASCNs. The reduced-order modeling is applicable to the alternating TLM variants as well. Further described MOR theory can be naturally transferred to the treatment of two-dimensional electromagnetic problems by the two-dimensional TLM method.

2.2 The TLM Time Evolution Scheme

Now we introduce the HILBERT space $\mathcal{H}_\mu$ (see APPENDIX A) spanned by the sequence of the grid points with normalized space coordinates l, m and n. The indices l, m and n are linked to the space coordinates with unitary spatial steps $\Delta l_x, \Delta l_y$ and Δl_z through $x = l\Delta l_x$, $y = m\Delta l_y$ and $z = n\Delta l_z$. The ket-vectors $|l, m, n\rangle$– given in Dirac notation – represent an orthonormal base of the space $\mathcal{H}_\mu$. Their HERMITIAN conjugate bra-vectors are denoted by $\langle l, m, n|$. The basis vectors $|l, m, n\rangle$, assigned to the space points (l, m, n), fulfil the orthogonality relation [65, 67]

$$\langle l_i, m_i, n_i \,|\, l_j, m_j, n_j \rangle = \delta_{l_i,l_j} \delta_{m_i,m_j} \delta_{n_i,n_j} \,. \tag{2.1}$$

The state vectors ${}_k\boldsymbol{a}_{l,m,n} \in \mathcal{R}^{12}$ and ${}_k\boldsymbol{b}_{l,m,n} \in \mathcal{R}^{12}$, which summarize the incident and the scattered wave pulses at the single SCN (l, m, n) at the discrete time k corresponding to the unit time interval Δt are given by

$$\begin{aligned} {}_k\boldsymbol{a}_{l,m,n} &= {}_k\left[a_1, a_2, a_3, \ldots a_{10}, a_{11}, a_{12}\right]^T_{l,m,n} \,, \\ {}_k\boldsymbol{b}_{l,m,n} &= {}_k\left[b_1, b_2, b_3, \ldots b_{10}, b_{11}, b_{12}\right]^T_{l,m,n} \,. \end{aligned} \tag{2.2}$$

The incident wave amplitudes a_1 to a_{12} and the scattered wave amplitudes b_1 to b_{12} in the arms (or at the ports) of the single SCN (l, m, n) (at the boundary surfaces of the TLM cell (l, m, n)) are elements of the real vector in $\mathcal{R}^{12}$.

The instant scattering of the wave pulses at the single SCN is represented by

$${}_{k+1}\boldsymbol{b}_{l,m,n} = {}_k\boldsymbol{S}_{l,m,n} \;\; {}_k\boldsymbol{a}_{l,m,n} \,, \tag{2.3}$$

where the operator ${}_k\boldsymbol{S}_{l,m,n} \in \mathcal{R}^{12}$ is related to the scattering rules of wave amplitudes at the single SCN (l, m, n).

Next we construct the state-space of the wave amplitudes $\mathcal{H}^{12}_\nu$ from the direct product of $\mathcal{H}_\mu$ and $\mathcal{C}^{12}$ as

$$H^{12}_\nu = C^{12} \otimes \mathcal{H}_\mu. \tag{2.4}$$

The vectors $|{}_k\boldsymbol{a}\rangle \in \mathcal{H}^{12}_\nu$ and $|{}_k\boldsymbol{b}\rangle \in \mathcal{H}^{12}_\nu$ denoting all incident and scattered wave amplitudes of the complete TLM mesh at the time k can thus be written in the forms

$$\begin{aligned} |{}_k\boldsymbol{a}\rangle &= \sum_{l,m,n=-\infty}^{+\infty} {}_k\boldsymbol{a}_{l,m,n} \,|l, m, n\rangle \,, \\ |{}_k\boldsymbol{b}\rangle &= \sum_{l,m,n=-\infty}^{+\infty} {}_k\boldsymbol{b}_{l,m,n} \,|l, m, n\rangle \,. \end{aligned} \tag{2.5}$$

Application of the projection operator $|l,m,n\rangle\langle l,m,n|$ onto the vectors $|_k\boldsymbol{a}\rangle$ and $|_k\boldsymbol{b}\rangle$ yields the vectors incident and scattered wave amplitudes at the SCN (l,m,n), i.e.

$$|l,m,n\rangle\langle l,m,n|_k\boldsymbol{a}\rangle = {}_k\boldsymbol{a}_{l,m,n}\,|l,m,n\rangle\ , \tag{2.6}$$

$$|l,m,n\rangle\langle l,m,n|_k\boldsymbol{b}\rangle = {}_k\boldsymbol{b}_{l,m,n}\,|l,m,n\rangle\ . \tag{2.7}$$

Accordingly to eq. (2.3) the simultaneous scattering at all TLM mesh nodes is described by the operator equation

$$|_{k+1}\boldsymbol{b}\rangle = {}_k\boldsymbol{S}\,|_k\boldsymbol{a}\rangle\ , \tag{2.8}$$

where the scattering matrix ${}_k\boldsymbol{S}$ is an operator in $\mathcal{H}_\nu^{12}$ and given by

$${}_k\boldsymbol{S} = \sum_{l,m,n=-\infty}^{+\infty} |l,m,n\rangle\,{}_k\boldsymbol{S}_{l,m,n}\,\langle l,m,n|\ . \tag{2.9}$$

The operator ${}_k\boldsymbol{S} \in \mathcal{H}_\nu^{12}$ represents the scattering operations at the time k for the complete TLM mesh.

The compositions of the projection operator $|l,m,n\rangle\langle l,m,n|$ and operator ${}_k\boldsymbol{S}$ give the scattering operator at the single node (l,m,n), e.g.

$$|l,m,n\rangle\langle l,m,n| \circ {}_k\boldsymbol{S} = {}_k\boldsymbol{S} \circ |l,m,n\rangle\langle l,m,n| = \tag{2.10}$$
$$= |l,m,n\rangle\,{}_k\boldsymbol{S}_{l,m,n}\,\langle l,m,n|\ .$$

Introducing the connection operator ${}_k\boldsymbol{\Gamma} \in \mathcal{H}_\nu^{12}$ to specify the interaction between adjacent cells (or nodes) we obtain *the TLM time evolution scheme* for a lossless case

$$|_{k+1}\boldsymbol{a}\rangle = {}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\,|_k\boldsymbol{a}\rangle\ . \tag{2.11}$$

Now we introduce the HILBERT space $\mathcal{H}_t$ with an orthonormal basis $|k\rangle$, i.e.

$$\langle k_i\,|\,k_j\rangle = \delta_{k_i,k_j}\ , \tag{2.12}$$

in order to describe the time state in the TLM time evolution scheme.

Then the time-spatial state-space of incident and/or scattered wave amplitudes can be defined by

$$\mathcal{H}_\xi^{12} = \mathcal{H}_t \otimes \mathcal{H}_\nu^{12} = \mathcal{H}_t \otimes \mathcal{C}^{12} \otimes \mathcal{H}_\mu\ . \tag{2.13}$$

We can present the orthonormal basis of $\mathcal{H}_t \otimes \mathcal{H}_\mu$ space through

$$|\,k,l,m,n\rangle = |\,k\rangle \otimes |\,l,m,n\rangle\ . \tag{2.14}$$

The complete time-spatial state vectors of the incident $|\boldsymbol{a}\rangle \in \mathcal{H}_\xi^{12}$ and the scattered waves $|\boldsymbol{b}\rangle \in \mathcal{H}_\xi^{12}$ in the TLM system we construct by

$$|\boldsymbol{a}\rangle = \sum_{k=-\infty}^{+\infty} |\,k\rangle \otimes |_k\boldsymbol{a}\rangle\ , \tag{2.15}$$

$$|\boldsymbol{b}\rangle = \sum_{k=-\infty}^{+\infty} |\,k\rangle \otimes |_k\boldsymbol{b}\rangle\ . \tag{2.16}$$

In order to describe the time evolution of electromagnetic fields in the TLM scheme we define an unitary time shift operator in $\mathcal{H}_t \otimes \mathcal{H}_\mu$ as

$$\boldsymbol{T} \,|\, k,l,m,n\rangle = |\, k+1,l,m,n\rangle \ , \tag{2.17}$$

which does not describe the time evolution of the vector $|\, l,m,n\rangle$, it just shifts the vector $|l,m,n\rangle$ for a time step Δt, i.e.

$$\boldsymbol{T}(\boldsymbol{T} \,|\, k,l,m,n\rangle) = |\, k+2,l,m,n\rangle \,. \tag{2.18}$$

Its inverse shift operator

$$\boldsymbol{T}^\dagger \,|\, k,l,m,n\rangle = |\, k-1,l,m,n\rangle \tag{2.19}$$

satisfies

$$\boldsymbol{T}\boldsymbol{T}^\dagger = \boldsymbol{T}^\dagger\boldsymbol{T} = \sum_{k,l,m,n=-\infty}^{+\infty} |\, k,l,m,n\rangle \langle k,l,m,n| = \mathbf{1}. \tag{2.20}$$

Therefore, according to eq. (2.11) the TLM algorithm can be summarized in $\mathcal{H}_\xi^{12}$ as following

$$|\boldsymbol{b}\rangle = \boldsymbol{T}\boldsymbol{S}\,|\boldsymbol{a}\rangle \,, \tag{2.21}$$

$$|\boldsymbol{a}\rangle = \boldsymbol{\Gamma}\,|\boldsymbol{b}\rangle \,. \tag{2.22}$$

The global scattering $\boldsymbol{S}$ and connection $\boldsymbol{\Gamma}$ operators in (2.21) and (2.22) are defined as

$$\boldsymbol{S} = \sum_{k,l,m,n=-\infty}^{+\infty} |k,l,m,n\rangle \, {}_k\boldsymbol{S}_{k,l,m,n} \langle k,l,m,n| \ , \tag{2.23}$$

and

$$\boldsymbol{\Gamma} = \sum_{k=-\infty}^{+\infty} |k\rangle \, {}_k\boldsymbol{\Gamma} \langle k| \ , \tag{2.24}$$

which summarize scattering and connection operations for the complete TLM mesh at all discrete times k.

The system of equations (2.21) and (2.22) gives the complete time-spatial state description of the TLM scheme. The first equation (2.21) describes the scattering of all incident wave amplitudes at all times k, the second equation (2.22) describes the connection between all cells at all times k. Since the time shift operator commutates with both scattering and connection operators the complete time-spatial state of the TLM system can be written

$$|\boldsymbol{a}\rangle = \boldsymbol{T}\boldsymbol{\Gamma}\boldsymbol{S}\,|\boldsymbol{a}\rangle \,, \tag{2.25}$$

and or as

$$|\boldsymbol{b}\rangle = \boldsymbol{T}\boldsymbol{S}\boldsymbol{\Gamma}\,|\boldsymbol{b}\rangle \,. \tag{2.26}$$

From eq. (2.25) and eq. (2.26) it follows that the TLM system can be fully charaterized by one of the state vectors of incident or scattered amplitudes only, i.e. one of the wave vectors has to be stored to represent state vectors for the electromagnetic field.

2.3 The Scattering Operator

The scattering operator $\boldsymbol{S}_{l,m,n}$ of dimension 12×12 for the single cubic SCN for a lossless homogeneous medium has been determined by JOHNS [60] as follows

$$\boldsymbol{S}_{l,m,n} = \begin{bmatrix} a & c & 0 & 0 & 0 & 0 & d & -d & 0 & 0 & b & b \\ c & a & 0 & 0 & 0 & 0 & -d & d & 0 & 0 & b & b \\ 0 & 0 & a & c & b & b & 0 & 0 & d & -d & 0 & 0 \\ 0 & 0 & c & a & b & b & 0 & 0 & -d & d & 0 & 0 \\ 0 & 0 & b & b & a & c & 0 & 0 & 0 & 0 & d & -d \\ 0 & 0 & b & b & c & a & 0 & 0 & 0 & 0 & -d & d \\ d & -d & 0 & 0 & 0 & 0 & a & c & b & b & 0 & 0 \\ -d & d & 0 & 0 & 0 & 0 & c & a & b & b & 0 & 0 \\ 0 & 0 & d & -d & 0 & 0 & b & b & a & c & 0 & 0 \\ 0 & 0 & -d & d & 0 & 0 & b & b & c & a & 0 & 0 \\ b & b & 0 & 0 & d & -d & 0 & 0 & 0 & 0 & a & c \\ b & b & 0 & 0 & -d & d & 0 & 0 & 0 & 0 & c & a \end{bmatrix} . \tag{2.27}$$

The structure of the scattering matrix satisfies the next considerations:

- Each incident wave pulse is associated with one E-field and one H-field components, and therefore linked with some other ports according to the MAXWELL's equations (see Fig. 2.1 and 2.2). For example, following from

$$\frac{dH_z}{dy} - \frac{dH_y}{dz} = \varepsilon \frac{dE_x}{dt}\,, \qquad \frac{dE_y}{dx} - \frac{dE_x}{dy} = -\mu \frac{dH_z}{dt}, \tag{2.28}$$

  the pulse a_7 associated with E_x and H_z is reflected to the ports 7 (i.e. self-reflected), 9 and 10 (since they are associated with E_x and H_y), 1 and 2 (related to E_y and H_z) and 8 (associated with E_x and H_z). Then each row and/or column of $\boldsymbol{S}_{l,m,n}$ only consists of six non-zero unknown elements.

- Due to the symmetry of the node the total number of unknown parameters in eq. (2.27) may be reduced down to four, and

$$\boldsymbol{S}_{l,m,n} = \boldsymbol{S}_{l,m,n}^T \,. \tag{2.29}$$

- The scattering process has to preserve the charge at the node.

- While the SCN exhibits no losses, the total reflected power must be equal to the total incident power, i.e.

$$\boldsymbol{S}_{l,m,n}^T \boldsymbol{S}_{l,m,n} = \mathbf{1} \,. \tag{2.30}$$

The last two considerations lead to one of four possible solutions with

$$a = 0, \quad b = \frac{1}{2}, \quad c = 0, \quad d = \frac{1}{2}. \tag{2.31}$$

The scattering operator eq. (2.27) may be written in a compact form as

$$\boldsymbol{S}_{l,m,n} = \begin{bmatrix} 0 & \boldsymbol{S}_0 & \boldsymbol{S}_0{}^T \\ \boldsymbol{S}_0{}^T & 0 & \boldsymbol{S}_0 \\ \boldsymbol{S}_0 & \boldsymbol{S}_0{}^T & 0 \end{bmatrix}, \text{ with } \boldsymbol{S}_0 = \begin{bmatrix} 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix} . \tag{2.32}$$

The scattering operator $\boldsymbol{S}_{l,m,n}$ is symmetric, in general – self-adjoint or HERMITIAN, and unitary, i.e.

$$\boldsymbol{S}_{l,m,n} = \boldsymbol{S}^{\dagger}_{l,m,n} = \boldsymbol{S}^{-1}_{l,m,n} \,. \tag{2.33}$$

In this case all twelve SCN transmission lines (or six SCN arms) of length $\Delta l/2$ have the same characteristic impedance of the free space

$$Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} = \sqrt{\frac{L_0}{C_0}} \,, \tag{2.34}$$

and the propagation velocity in the TLM grid (the velocity on the transmission lines between adjacent nodes) is

$$c_0 = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = \frac{1}{\sqrt{L_0 C_0}} \,. \tag{2.35}$$

Hence, the propagation time on each arm is $\Delta t/2$, and according to eq. (2.34) and (2.35) the characteristic impedance and the propagation velocity in each arm can be modeled by network with one capacitance C_0 and one inductance L_0 [95, 78]

$$C_0 = Y_0 \frac{\Delta t}{2} = \varepsilon_0 \frac{\Delta l}{h} \quad \text{and} \quad L_0 = Z_0 \frac{\Delta t}{2} = \mu_0 \frac{\Delta l}{h} \,, \tag{2.36}$$

where h is a stability factor introduced as

$$h \geq \frac{2\Delta l}{c_0 \Delta t} \,. \tag{2.37}$$

We note that the propagation velocity of a dispersion free wave along a cubic cell diagonal is $c_0/2$ (for a two-dimensional case $c_0/\sqrt{2}$), whereas the propagation time between two scattering centers is fixed by $\Delta t = \Delta l/c_0$ [23]. Due to the SCN symmetry the axial wave propagation can be modeled by a network with double capacitance C_0 and double inductance L_0, and voltages and currents can be scanned and/or sampled with $\Delta t/2$. So the synchronization of wave pulses is kept to be guaranteed.

In order to model an inhomogeneous medium with the characteristic impedance

$$Z_m = \sqrt{\frac{\mu}{\varepsilon}} = \sqrt{\frac{L}{C}} \tag{2.38}$$

and with the propagation velocity

$$c_m = \frac{1}{\sqrt{\mu \varepsilon}} = \frac{1}{\sqrt{LC}} \tag{2.39}$$

and to use the non-uniform cubic cells with dimensions

$$(\Delta l_x, \Delta l_y, \Delta l_z) \equiv (i_x \Delta l, i_y \Delta l, i_z \Delta l) \tag{2.40}$$

JOHNS added six circuited stubs to the 12×12 node. The first three opened stubs are related to the E_x, E_y, and E_z field components and cause an additional capacitance. The second three shorted stubs add an inductance to the node and coupled to H_x, H_y, and H_z. For more details see [23] and [60].

Starting at the cubic SCN, the prolongation (or the reduction) of a node arm length causes the changing of the propagation time through the node. The equal time variation can be achieved through the corresponding scaling of capacitances and inductances as well. According to eq. (2.38) and (2.39) the deformation of the node and/or the scaling of the reactivities are equivalent to the local variation of the permittivity ε and permeability μ. The modeling of material properties by means of heterogeneous lengths of transmission lines is presented in TRENKIC'S works and not discussed in the present thesis deeply. We only note that this approach yields a more complicated structure of the scattering matrix because of additional scattering centers in the TLM cell.

The required total capacitances at a SCN of dimension $(i_x\Delta l, i_y\Delta l, i_z\Delta l)$ introduced by JOHNS in [60] are

$$C_x = \varepsilon_x \frac{i_y i_z}{i_x}\Delta l \ , \ C_y = \varepsilon_y \frac{i_x i_z}{i_y}\Delta l \ , \ C_z = \varepsilon_z \frac{i_x i_y}{i_z}\Delta l \, , \tag{2.41}$$

and inductances are

$$L_x = \mu_x \frac{i_y i_z}{i_x}\Delta l \ , \ L_y = \mu_y \frac{i_x i_z}{i_y}\Delta l \ , \ L_z = \mu_z \frac{i_x i_y}{i_z}\Delta l \, . \tag{2.42}$$

For example, the modeled capacitance C_x is associated with lines 7, 8, 9 and 10, and as following, is equivalent to $4C_0$ for a cubic TLM cell in free space. The modeled inductance L_x is linked to the lines 5, 6, 11 and 12, and is equivalent to $4L_0$. Therefore, in order to model non-uniform cells the capacitance and inductance have to be added to the symmetric node, e.g. for the x–components we have

$$C_x^{stub} = \varepsilon_x \frac{i_y i_z}{i_x} - 4C_0 \quad \text{and} \quad L_x^{stub} = \mu_x \frac{i_y i_z}{i_x} - 4L_0 \, . \tag{2.43}$$

From eq. (2.36), (2.41) - (2.43) and from relations normalized with respect to Y_0 and/or Z_0

$$C^{stub} = Y_0 Y^{stub} \frac{\Delta t}{2} \quad \text{and} \quad L^{stub} = Z_0 Z^{stub} \frac{\Delta t}{2} \, , \tag{2.44}$$

where the admittances Y^{stub} of the stubs 13-15 are calculated by

$$Y_x = 2(\varepsilon_{rx} h \frac{i_y i_z}{i_x} - 2) \ , \ Y_y = 2(\varepsilon_{ry} h \frac{i_x i_z}{i_y} - 2) \ , \ Y_z = 2(\varepsilon_{rz} \frac{i_x i_y}{i_z} - 2) \, , \tag{2.45}$$

and the impedances Z^{stub} of the stubs 16-18 are given by

$$Z_x = 2(\mu_{rx} h \frac{i_y i_z}{i_x} - 2) \ , \ Z_y = 2(\mu_{ry} h \frac{i_x i_z}{i_y} - 2) \ , \ Z_z = 2(\mu_{rz} \frac{i_x i_y}{i_z} - 2) \, , \tag{2.46}$$

where ε_r and μ_r denote relative permittivity and permeability, respectively.

Obviously, the stability factor in eq. (2.37) is chosen to preserve the *positivity* for C^{stub} and L^{stub}.

These additional three capacitive and three inductive stubs are connected to the SCN and are not linked with adjacent nodes directly. The connection process to the neighboring SCNs occurs through 12 link lines without energy loss. If ohmic and magnetic losses have to be included, six additional stubs

are connected to the center of the SCN, so the 18×18 scattering matrix gets six additional lines. The scattering in these stubs has not to be considered explicitly, since the energy dissipated in the center node is lost. These six lines are necessary if the loss power has to be calculated, else 18×18 matrix in eq. (2.49) is used. The normalized values of the ohmic loss stubs are presented by

$$G_x = \frac{i_y i_z}{i_x}\sigma_{ex}\Delta l\, Z_0\ , \quad G_y = \frac{i_x i_z}{i_y}\sigma_{ey}\Delta l\, Z_0\ , \quad G_z = \frac{i_x i_y}{i_z}\sigma_{ez}\Delta l\, Z_0\,, \tag{2.47}$$

and the magnetic loss stubs are calculated as

$$R_x = \frac{i_y i_z}{i_x}\sigma_{mx}\Delta l\, Y_0\ , \quad R_y = \frac{i_x i_z}{i_y}\sigma_{my}\Delta l\, Y_0\ , \quad R_z = \frac{i_x i_y}{i_z}\sigma_{mz}\Delta l\, Y_0\,, \tag{2.48}$$

where σ_e and σ_m are electric and magnetic conductivities, respectively.

Considering stubs the general scattering 18×18 matrix of any deformed SCN for modeling of inhomogeneous media with losses has been introduced by JOHNS as following

$$\boldsymbol{S}^{stubs}_{l,m,n} = \begin{bmatrix}
a & c & 0 & 0 & 0 & 0 & d & -d & 0 & 0 & b & b & 0 & g & 0 & 0 & 0 & -d \\
c & a & 0 & 0 & 0 & 0 & -d & d & 0 & 0 & b & b & 0 & g & 0 & 0 & 0 & d \\
0 & 0 & a & c & b & b & 0 & 0 & d & -d & 0 & 0 & 0 & 0 & g & 0 & d & 0 \\
0 & 0 & c & a & b & b & 0 & 0 & -d & d & 0 & 0 & 0 & 0 & g & 0 & -d & 0 \\
0 & 0 & b & b & a & c & 0 & 0 & 0 & 0 & d & -d & 0 & 0 & g & -d & 0 & 0 \\
0 & 0 & b & b & c & a & 0 & 0 & 0 & 0 & -d & d & 0 & 0 & g & d & 0 & 0 \\
d & -d & 0 & 0 & 0 & 0 & a & c & b & b & 0 & 0 & g & 0 & 0 & 0 & 0 & d \\
-d & d & 0 & 0 & 0 & 0 & c & a & b & b & 0 & 0 & g & 0 & 0 & 0 & 0 & -d \\
0 & 0 & d & -d & 0 & 0 & b & b & a & c & 0 & 0 & g & 0 & 0 & 0 & -d & 0 \\
0 & 0 & -d & d & 0 & 0 & b & b & c & a & 0 & 0 & g & 0 & 0 & 0 & d & 0 \\
b & b & 0 & 0 & d & -d & 0 & 0 & 0 & 0 & a & c & 0 & g & 0 & d & 0 & 0 \\
b & b & 0 & 0 & -d & d & 0 & 0 & 0 & 0 & c & a & 0 & g & 0 & -d & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & b & b & b & b & 0 & 0 & h & 0 & 0 & 0 & 0 & 0 \\
b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b & b & 0 & h & 0 & 0 & 0 & 0 \\
0 & 0 & b & b & b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -f & f & 0 & 0 & 0 & 0 & f & -f & 0 & 0 & 0 & j & 0 & 0 \\
0 & 0 & f & -f & 0 & 0 & 0 & 0 & -f & f & 0 & 0 & 0 & 0 & 0 & 0 & j & 0 \\
-f & f & 0 & 0 & 0 & 0 & f & -f & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & j
\end{bmatrix} \tag{2.49}$$

with coefficients

$$\begin{aligned}
a &= -\frac{G+Y}{2(G+Y+4)} + \frac{R+Z}{2(R+Z+4)}\ , & b &= \frac{2}{G+Y+4}\ , \\
c &= -\frac{G+Y}{2(G+Y+4)} - \frac{R+Z}{2(R+Z+4)}\ , & d &= \frac{2}{Z+R+4}\ , \\
f &= \frac{2Z}{R+Z+4}\ , & g &= \frac{2Y}{G+Y+4}\ , \\
h &= -\frac{G-Y+4}{G+Y+4}\ , & j &= -\frac{-R+Z-4}{R+Z+4}\ .
\end{aligned} \tag{2.50}$$

For the general scattering matrix (2.49) with losses we have that

$$\boldsymbol{S}_{l,m,n}^{stub} \neq (\boldsymbol{S}_{l,m,n}^{stub})^{\dagger} \neq (\boldsymbol{S}_{l,m,n}^{stub})^{-1} . \tag{2.51}$$

The symmetry of $\boldsymbol{S}_{l,m,n}^{stub}$ is not preserved due to the stub parameter normalization proposed by JOHNS. The unitarity of the scattering matrix is not only lost due to the losses in an analyzed structure. Indeed, even though the scattering matrix in eq. (2.49) comprises no losses, it is also non-unitary. In other words, the unitarity and symmetry are lost due to the normalization in accord with eq. (2.49). However, in this case the complete energy during the scattering process is preserved, and $\boldsymbol{S}_{l,m,n}^{stub}$ has nine eigenvalues being ones and nine eigenvalues being equal to -1 (For more details see Chapter 5).

However, the symmetry of the general scattering operator in eq. (2.49) can be reconstructed by a new normalization without its energy conservation property being changed by following

$$\hat{\boldsymbol{S}}_{l,m,n}^{stub} = \boldsymbol{N}^{-1} \, \boldsymbol{S}_{l,m,n}^{stub} \, \boldsymbol{N} , \tag{2.52}$$

here the diagonal matrix $\boldsymbol{N}$ is given as

$$\boldsymbol{N} = \mathrm{diag}(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,(\sqrt{Y})^{-1},\sqrt{Z}) . \tag{2.53}$$

Now the modified general scattering operator for the lossy SCN in eq. (2.52) satisfies next relations

$$\hat{\boldsymbol{S}}_{l,m,n}^{stub} = (\hat{\boldsymbol{S}}_{l,m,n}^{stub})^{\dagger} \neq (\hat{\boldsymbol{S}}_{l,m,n}^{stub})^{-1} . \tag{2.54}$$

The general scattering operator ${}_k\boldsymbol{S}$ in

$$\mathcal{H}_{\nu}^{18} = \mathcal{C}^{18} \otimes \mathcal{H}_{\mu} , \tag{2.55}$$

describing the scattering at all nodes of the TLM mesh at the time k, has the block-diagonal matrix structure, i.e.

$${}_k\boldsymbol{S} = \mathrm{diag}\,(\hat{\boldsymbol{S}}_{l,m,n}^{stub}) . \tag{2.56}$$

From eq. (2.54) and (2.56) we obtain essential properties of the operator $\boldsymbol{S}$, and namely

$${}_k\boldsymbol{S} = {}_k\boldsymbol{S}^{\dagger} \neq {}_k\boldsymbol{S}^{-1} . \tag{2.57}$$

Further, we will deal the modified scattering operators in accord with eq. (2.52) and eq. (2.56), which are self-adjoint (or HERMITIAN) and non-unitary. The modification of the $\boldsymbol{S}_{l,m,n}$ in eq. (2.52) does not infringe the scattering operation proposed by JOHNS and has only impact on the field mapping between the electromagnetic field and the wave amplitudes.

That is clear that the matrix representing the scattering at all TLM nodes at all times k is operating in

$$\mathcal{H}_{\xi}^{18} = \mathcal{H}_t \otimes \mathcal{H}_{\nu}^{18} = \mathcal{H}_t \otimes \mathcal{C}^{18} \otimes \mathcal{H}_{\mu} . \tag{2.58}$$

2.4 The Connection Operator

Next, we describe the connection operator $\boldsymbol{\Gamma}$ representing the interaction between all adjacent nodes in $\mathcal{H}_\nu^{18}$. For this purpose we define the shift operators $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ and their adjoints $\boldsymbol{X}^\dagger$, $\boldsymbol{Y}^\dagger$ and $\boldsymbol{Z}^\dagger$ so that

$$\begin{aligned} \boldsymbol{X}\ |l,m,n\rangle &= |l+1,m,n\rangle\ , \\ \boldsymbol{X}^\dagger|l,m,n\rangle &= |l-1,m,n\rangle\ , \end{aligned} \tag{2.59}$$

$$\begin{aligned} \boldsymbol{Y}\ |l,m,n\rangle &= |l,m+1,n\rangle\ , \\ \boldsymbol{Y}^\dagger|l,m,n\rangle &= |l,m-1,n\rangle\ , \end{aligned} \tag{2.60}$$

$$\begin{aligned} \boldsymbol{Z}\ |l,m,n\rangle &= |l,m,n+1\rangle\ , \\ \boldsymbol{Z}^\dagger|l,m,n\rangle &= |l,m,n-1\rangle\ . \end{aligned} \tag{2.61}$$

The spatial operators $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ shift the position of the node (l,m,n) by Δl in positive $x-$, $y-$ or $z-$ direction respectively. The operators $\boldsymbol{X}^\dagger$, $\boldsymbol{Y}^\dagger$ and $\boldsymbol{Z}^\dagger$ make the opposite moving by one step. For example, the second incident wave amplitudes of all nodes are linked to the first reflected wave amplitudes of the neighboring nodes, i.e. corresponding to Fig. 2.2 we have

$$\begin{aligned} \sum_{l,m,n=-\infty}^{+\infty} {}_k[a_2]_{l,m,n}\,|l,m,n\rangle &= \sum_{l,m,n=-\infty}^{+\infty} {}_k[b_1]_{l+1,m,n}\,|l,m,n\rangle = \\ = \sum_{l,m,n=-\infty}^{+\infty} {}_k[b_1]_{l,m,n}\,|l+1,m,n\rangle &= \boldsymbol{X} \sum_{l,m,n=-\infty}^{+\infty} {}_k[b_1]_{l,m,n}\,|l,m,n\rangle\ . \end{aligned} \tag{2.62}$$

The shift operators (2.61) can be written as proposed in [67] in the next form

$$\begin{aligned} \boldsymbol{X} &= \sum_{l,m,n=-\infty}^{+\infty} |l+1,m,n\rangle\langle l,m,n|\ , \\ \boldsymbol{X}^\dagger &= \sum_{l,m,n=-\infty}^{+\infty} |l-1,m,n\rangle\langle l,m,n|\ , \\ \boldsymbol{Y} &= \sum_{l,m,n=-\infty}^{+\infty} |l,m+1,n\rangle\langle l,m,n|\ , \\ \boldsymbol{Y}^\dagger &= \sum_{l,m,n=-\infty}^{+\infty} |l,m-1,n\rangle\langle l,m,n|\ , \\ \boldsymbol{Z} &= \sum_{l,m,n=-\infty}^{+\infty} |l,m,n+1\rangle\langle l,m,n|\ , \\ \boldsymbol{Z}^\dagger &= \sum_{l,m,n=-\infty}^{+\infty} |l,m,n-1\rangle\langle l,m,n|\ . \end{aligned} \tag{2.63}$$

From eq. (2.63) it follows that the shift operators are self-adjoint, unitary and commutative with itself, i.e.

$$\boldsymbol{X}\boldsymbol{X}^\dagger = \boldsymbol{X}^\dagger\boldsymbol{X} = \mathbf{1}\,, \tag{2.64}$$

$$\boldsymbol{Y}\boldsymbol{Y}^{\dagger} = \boldsymbol{Y}^{\dagger}\boldsymbol{Y} = \mathbf{1}\,, \tag{2.65}$$

$$\boldsymbol{Z}\boldsymbol{Z}^{\dagger} = \boldsymbol{Z}^{\dagger}\boldsymbol{Z} = \mathbf{1}\,. \tag{2.66}$$

Therefore, the connection between all nodes in the infinite space can be summarized in the following matrix form

$$ {}_k\boldsymbol{\Gamma} = \begin{bmatrix} 0 & \boldsymbol{X} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \boldsymbol{X}^{\dagger} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{X} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{X}^{\dagger} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boldsymbol{Y} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{Y}^{\dagger} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boldsymbol{Y} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boldsymbol{Y}^{\dagger} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boldsymbol{Z} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boldsymbol{Z}^{\dagger} & 0 & 0 & \boldsymbol{Z} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boldsymbol{Z}^{\dagger} & 0 \end{bmatrix} . \tag{2.67}$$

According to eq. (2.67) the connection operator ${}_k\boldsymbol{\Gamma}$ is self-adjoint and unitary

$${}_k\boldsymbol{\Gamma} = {}_k\boldsymbol{\Gamma}^{\dagger} = {}_k\boldsymbol{\Gamma}^{-1}\,, \tag{2.68}$$

and commutates with itself

$${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{\Gamma}^{\dagger} = {}_k\boldsymbol{\Gamma}^{\dagger}\,{}_k\boldsymbol{\Gamma} = \mathbf{1}\,. \tag{2.69}$$

Actually, it is self-evident that the matrix ${}_k\boldsymbol{\Gamma}$ for the infinite space is filled with ones, since no losses come into play during the linking operations.

In practical implementations we are dealing with electromagnetic problems with boundary conditions, i.e. with finite-dimensional vector spaces. So, in TLM a structure under consideration is embedded in a box of finite extension, and, therefore, the boundary condition must be included. In order to describe the connection between SCNs in a region with certain boundary conditions we introduce the $N \times N$ connection operator ${}_k\hat{\boldsymbol{\Gamma}}$ in HILBERT space $\mathcal{H}_{\nu}^{18}$ of dimension N as

$${}_k\hat{\boldsymbol{\Gamma}} = {}_k\boldsymbol{\Gamma} + {}_k\boldsymbol{\Upsilon}\,. \tag{2.70}$$

The matrix ${}_k\boldsymbol{\Gamma}$ is only responsible for the connection between nodes (without taking into account boundary conditions), the matrix ${}_k\boldsymbol{\Upsilon}$ represents boundary conditions, e.g. if reflection free walls have to be modeled, the matrix ${}_k\boldsymbol{\Upsilon}$ consists of local reflection coefficients at its diagonal, i.e.

$${}_k\boldsymbol{\Upsilon} = \mathrm{diag}\,[\rho_i^{loc}]\,, i = 1..N\,. \tag{2.71}$$

The diagonal elements of ${}_k\boldsymbol{\Upsilon}$ do not vanish in accord with cells located on the boundaries, therefore ${}_k\boldsymbol{\Upsilon}$ is a diagonal matrix mostly filled with zeros.

We remark that the structure of the ${}_k\boldsymbol{\Gamma}$ in eq. (2.67) for bounded media is slightly changed, because not all nodes are connected with their adjacent, i.e. several ones elements vanish in the matrix. Even though the unitary property of ${}_k\boldsymbol{\Gamma}$ is lost, the matrix ${}_k\hat{\boldsymbol{\Gamma}}$ in eq. (2.70) yields the energy conservation for a structure with PEC/PMC. ${}_k\boldsymbol{\Upsilon}$ is a diagonal matrix filled with ones and/or minus ones and zeros; ${}_k\boldsymbol{\Upsilon}$ and ${}_k\hat{\boldsymbol{\Gamma}}$ are both unitary operators. In some cases

the structure of the ${}_k\hat{\boldsymbol{\Gamma}}$-matrix has to be additionally modified, e.g. in order to model very thin layers (metal sheets) with reflection and transmission coefficients. One part of the scattered wave pulse will be reflected from a layer back to the nodes, and the second one will be transmitted (or linked) to the adjacent nodes on the other side of the layer. In this case the scattering matrix ${}_k\boldsymbol{\Gamma}$ has several elements (instead ones before) according to the transmission coefficients, and the diagonal matrix ${}_k\boldsymbol{\Upsilon}$ gets additional diagonal elements to commensurate with the reflection coefficients.

If we understand ${}_k\boldsymbol{\Gamma}$ in eq. (2.67) as a matrix describing the connection between nodes only (transmission coefficients), and ${}_k\boldsymbol{\Upsilon}$ is a matrix consisting of reflection coefficients at the node boundaries, then the properties of the linking operators on the assumption of a bounded medium with losses are given via

$$ {}_k\boldsymbol{\Gamma} = {}_k\boldsymbol{\Gamma}^{\dagger} \neq {}_k\boldsymbol{\Gamma}^{-1} , \tag{2.72} $$

$$ {}_k\boldsymbol{\Upsilon} = {}_k\boldsymbol{\Upsilon}^{\dagger} \neq {}_k\boldsymbol{\Upsilon}^{-1} = \text{diagonal} , \tag{2.73} $$

and from eq. (2.70) and (2.72), (2.73) follows that

$$ {}_k\hat{\boldsymbol{\Gamma}} = {}_k\hat{\boldsymbol{\Gamma}}^{\dagger} \neq {}_k\hat{\boldsymbol{\Gamma}}^{-1} . \tag{2.74} $$

In the following chapters we are dealing with ${}_k\hat{\boldsymbol{\Gamma}}$.

2.5 Mappings

The six components of the electromagnetic field at the TLM-SCN are represented by eighteen incident and/or scattered wave amplitudes. Thus, a correspondence, so called mapping, between the electromagnetic field and the wave pulses has to be established. A mapping is understood as a projection of the field state vector in

$$ \mathcal{H}_F = C^6 \otimes \mathcal{H}_\mu \tag{2.75} $$

onto the vector of incident wave amplitudes in

$$ \mathcal{H}_\nu = C^{18} \otimes \mathcal{H}_\mu , \tag{2.76} $$

and an vice versa.

In the present thesis the Centered Field Mapping (CFM) and the Cell Boundary Mapping (CBM) are considered.

2.5.1 Centered Field Mapping

We introduce the state vector of the electromagnetic field in $\mathcal{H}_F$ as

$$ |\boldsymbol{F}\rangle = \sum_{l,m,n=-\infty}^{+\infty} \begin{bmatrix} {}_kE^x_{l,m,n} \\ {}_kE^y_{l,m,n} \\ {}_kE^z_{l,m,n} \\ Z_k H^x_{l,m,n} \\ Z_k H^y_{l,m,n} \\ Z_k H^z_{l,m,n} \end{bmatrix} |l,m,n\rangle . \tag{2.77} $$

Each field component is determined from the voltages of four transmission lines according to the field component polarization, e.g. the E_x component is dependent on the wave pulses a_7, a_8, a_9, a_{10} for a x-direction polarized wave. Then introducing the part charge $Q_x = Q^a + Q^b = 2(C_0 a_7 + C_0 a_8 + C_0 a_9 + C_0 a_{10})$, which has to be conserved at the center of SCN, we have the equal homogeneous electric field

$$E_x = \frac{Q_x}{C_x \Delta l} = \frac{2C_0(a_7 + a_8 + a_9 + a_{10})}{4C_0 \Delta l} = \frac{1}{2\Delta l}(a_7 + a_8 + a_9 + a_{10})\,. \quad (2.78)$$

Introducing the magnetic part flux $\Phi_x = \Phi^a + \Phi^b = 2L_0(i_5 - L_0 i_6 - L_0 i_{11} + L_0 i_{12})$, which has to be conserved, we obtain in an analogous manner

$$I_x = \frac{\Phi_x}{L_x} = \frac{2L_0(i_5 - i_6 - i_{11} + i_{12})}{4L_0} = \frac{1}{2Z_0}(i_5 - i_6 - i_{11} + i_{12})\,, \quad (2.79)$$

and, consequently, the H_x component as

$$H_x = \frac{1}{2\Delta l Z_0}(i_5 - i_6 - i_{11} + i_{12}). \quad (2.80)$$

The disadvantage of this proposed approach is that an inverse mapping from the eighteen wave amplitudes onto the six field components can not be clearly defined. JOHNS proposed to distribute the wave amplitudes uniformly, e.g. if the E_x field has to be induced, one of possible mesh excitations can be

$$a_7 = a_8 = a_9 = a_{10} = \frac{1}{2}\Delta l E_x\,. \quad (2.81)$$

In order to describe the projection of the wave amplitude state vector onto the field state vector we introdece the projection operator $\boldsymbol{M} : \mathcal{H}_\nu \longrightarrow \mathcal{H}_F$ as

$$|\boldsymbol{F}\rangle = \frac{1}{\Delta l}\,\boldsymbol{M}\,|\boldsymbol{a}\rangle\,, \quad (2.82)$$

where

$$\boldsymbol{M} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (2.83)$$

Then the inverse mapping $\boldsymbol{M}^T : \mathcal{H}_F \longrightarrow \mathcal{H}_\nu$ is given as

$$|\boldsymbol{a}\rangle = \Delta l\,\boldsymbol{M}^T\,|\boldsymbol{F}\rangle\,. \quad (2.84)$$

From eq. (2.82) and (2.84) it follows that

$$|\boldsymbol{F}'\rangle = \boldsymbol{M}\boldsymbol{M}^T\,|\boldsymbol{F}\rangle \quad (2.85)$$

and

$$|\boldsymbol{a}'\rangle = \boldsymbol{M}^T\boldsymbol{M}\,|\boldsymbol{a}\rangle\,. \quad (2.86)$$

Since the operator $\boldsymbol{M}\boldsymbol{M}^T$ has ones at its diagonal, i.e.

$$\boldsymbol{M}\boldsymbol{M}^T = \mathrm{diag}\,[1,1,1,1,1,1]\,, \tag{2.87}$$

the application of the operator $\boldsymbol{M}\boldsymbol{M}^T : \mathcal{H}_F \longrightarrow \mathcal{H}_F$ onto the field state vector does not change it, or $|\boldsymbol{F}'\rangle = |\boldsymbol{F}\rangle$.

A different situation occurs for $\boldsymbol{M}^T\boldsymbol{M}$, due to the relation

$$\boldsymbol{M}^T\boldsymbol{M} = \boldsymbol{U}\,\mathrm{diag}\,[0,0,0,0,0,0,1,1,1,1,1,1]\,\boldsymbol{U}^{-1}, \tag{2.88}$$

where $\boldsymbol{U}$ is an unitary matrix.

We have that $|\boldsymbol{a}\rangle \in C^{12} \otimes \mathcal{H}_\mu$, whereas $|\boldsymbol{a}'\rangle \in C^6 \otimes \mathcal{H}_\mu$. The space $C^{12} \otimes \mathcal{H}_\mu$ can be presented by two orthogonal subspaces $C^{12} \otimes \mathcal{H}_\mu = C^6 \otimes \mathcal{H}_\mu \oplus C^6 \otimes \mathcal{H}_\mu$, the operator $\boldsymbol{M}^T\boldsymbol{M}$ transfers the vector $|\boldsymbol{a}\rangle$ from the space $C^{12} \otimes \mathcal{H}_\mu$ to its subspace $C^6 \otimes \mathcal{H}_\mu$. In other words, the operator $\boldsymbol{M}^T\boldsymbol{M} : \mathcal{H}_{C^{12} \otimes \mathcal{H}_\mu} \longrightarrow \mathcal{H}_{C^6 \otimes \mathcal{H}_\mu}$ leads to $|\boldsymbol{a}'\rangle \neq |\boldsymbol{a}\rangle$.

An important property, which has to be noticed, is that the whole energy in the wave amplitudes is conserved during a TLM simulation, whereas the energy in the electromagnetic field obtained by mapping in eq. (2.82) is not [65].

In an analogous way we can derive a mapping including stubs. For example, we consider the E_x and H_x components again. Considering the part charge, which is given by $Q_x = 2(C_0 a_7 + C_0 a_8 + C_0 a_9 + C_0 a_{10} + C^{stub} a_{13})$, we obtain in accord to [60]

$$\begin{aligned} E_x = \frac{Q_x}{C_x \Delta l} &= \frac{2C_0(a_7 + a_8 + a_9 + a_{10}) + 2C^{stub} a_{13}}{(4C_0 + C^{stub})\Delta l} = \\ &= \frac{2}{\Delta l} \cdot \left(\frac{a_7 + a_8 + a_9 + a_{10} + Y_x a_{13}}{4 + Y_x} \right), \end{aligned} \tag{2.89}$$

and using $\Phi_x = 2(L_0 i_5 - L_0 i_6 - L_0 i_{11} + L_0 i_{12} + L^{stub} i_{16})$, we have

$$H_x = \frac{2}{\Delta l Z_0} \left(\frac{a_5 - a_6 - a_{11} + a_{12} + \sqrt{Z_x} a_{16}}{4 + Z_x} \right). \tag{2.90}$$

Next, taking into account

- losses in the media (eq. (2.47) and eq. (2.48)), or to be precise, e.g. for the E_x-component

$$a_7 + a_8 + a_9 + a_{10} + Y_x a_{13} = b_7 + b_8 + b_9 + b_{10} + Y_x b_{13} + G_x b_{G_x}, \tag{2.91}$$

  and for the H_x-component

$$a_5 - a_6 - a_{11} + a_{12} + a_{16} = -(b_5 - b_6 - b_{11} + b_{12} + b_{16} + R_x b_{R_x}), \tag{2.92}$$

- deformations of the TLM mesh in eq. (2.45) and eq. (2.46), and
- normalization of the scattering matrix given by eq. (2.52)

we obtain all six components of the electromagnetic field as

$$E_x = \frac{2}{i_x \Delta l} \left(\frac{a_7 + a_8 + a_9 + a_{10} + \sqrt{Y_x} a_{13}}{4 + Y_x + G_x} \right), \tag{2.93}$$

$$E_y = \frac{2}{i_y \Delta l} \left(\frac{a_1 + a_2 + a_{11} + a_{12} + \sqrt{Y_y} a_{14}}{4 + Y_y + G_y} \right), \tag{2.94}$$

$$E_z = \frac{2}{i_z \Delta l} \left(\frac{a_3 + a_4 + a_5 + a_6 + \sqrt{Y_z} a_{15}}{4 + Y_z + G_z} \right), \tag{2.95}$$

$$H_x = \frac{2}{i_x \Delta l Z_0}\left(\frac{a_5 - a_6 - a_{11} + a_{12} + \sqrt{Z_x}a_{16}}{4 + Z_x + R_x}\right), \tag{2.96}$$

$$H_y = \frac{2}{i_y \Delta l Z_0}\left(\frac{-a_3 + a_4 + a_9 - a_{10} + \sqrt{Z_x}a_{17}}{4 + Z_y + R_y}\right), \tag{2.97}$$

$$H_z = \frac{2}{i_z \Delta l Z_0}\left(\frac{a_1 - a_2 - a_7 + a_8 + \sqrt{Z_x}a_{18}}{4 + Z_z + R_z}\right). \tag{2.98}$$

And an excitation of the TLM mesh can be done uniformly as

$$a_1 = \frac{1}{2}\Delta l(i_y E_y + i_z Z_0 H_z), \quad a_7 = \frac{1}{2}\Delta l(i_x E_x - i_z Z_0 H_z), \tag{2.99}$$

$$a_2 = \frac{1}{2}\Delta l(i_y E_y - i_z Z_0 H_z), \quad a_8 = \frac{1}{2}\Delta l(i_x E_x + i_z Z_0 H_z), \tag{2.100}$$

$$a_3 = \frac{1}{2}\Delta l(i_z E_z - i_y Z_0 H_y), \quad a_9 = \frac{1}{2}\Delta l(i_x E_x + i_y Z_0 H_y), \tag{2.101}$$

$$a_4 = \frac{1}{2}\Delta l(i_z E_z + i_y Z_0 H_y), \quad a_{10} = \frac{1}{2}\Delta l(i_x E_x - i_y Z_0 H_y), \tag{2.102}$$

$$a_5 = \frac{1}{2}\Delta l(i_z E_z + i_x Z_0 H_x), \quad a_{11} = \frac{1}{2}\Delta l(i_y E_y - i_x Z_0 H_x), \tag{2.103}$$

$$a_6 = \frac{1}{2}\Delta l(i_z E_z - i_x Z_0 H_x), \quad a_{12} = \frac{1}{2}\Delta l(i_y E_y + i_x Z_0 H_x), \tag{2.104}$$

$$a_{13} = \frac{1}{2} i_x \Delta l \sqrt{Y_x} E_x, \quad a_{16} = \frac{1}{2} i_x \Delta l Z_0 \sqrt{Z_x} H_x, \tag{2.105}$$

$$a_{14} = \frac{1}{2} i_y \Delta l \sqrt{Y_y} E_y, \quad a_{17} = \frac{1}{2} i_y \Delta l Z_0 \sqrt{Z_y} H_y, \tag{2.106}$$

$$a_{15} = \frac{1}{2} i_y \Delta l \sqrt{Y_z} E_z, \quad a_{18} = \frac{1}{2} i_z \Delta l Z_0 \sqrt{Z_z} H_z. \tag{2.107}$$

The operator $\boldsymbol{M}_{stubs}^{a \to F} : \mathcal{H}_\nu \longrightarrow \mathcal{H}_F$ can be summarized in the form

$$\boldsymbol{M}_{stub}^{a \to F} = \begin{bmatrix} 0 & a_y^E & 0 & 0 & 0 & a_z^H \\ 0 & a_y^E & 0 & 0 & 0 & -a_z^H \\ 0 & 0 & a_z^E & 0 & -a_y^H & 0 \\ 0 & 0 & a_z^E & 0 & a_y^H & 0 \\ 0 & 0 & a_z^E & a_x^H & 0 & 0 \\ 0 & 0 & a_z^E & -a_x^H & 0 & 0 \\ a_x^E & 0 & 0 & 0 & 0 & -a_z^H \\ a_x^E & 0 & 0 & 0 & 0 & a_z^H \\ a_x^E & 0 & 0 & 0 & a_y^H & 0 \\ a_x^E & 0 & 0 & 0 & -a_y^H & 0 \\ 0 & a_y^E & 0 & -a_x^H & 0 & 0 \\ 0 & a_y^E & 0 & a_x^H & 0 & 0 \\ a_x^E\sqrt{Y_x} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_y^E\sqrt{Y_y} & 0 & 0 & 0 & 0 \\ 0 & 0 & a_z^E\sqrt{Y_z} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_x^H\sqrt{Z_x} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_y^H\sqrt{Z_y} & 0 \\ 0 & 0 & 0 & 0 & 0 & a_z^H\sqrt{Z_z} \end{bmatrix}^T, \tag{2.108}$$

where

$$a_x^E = \frac{2}{i_x(4+Y_x+G_x)}, \quad a_y^E = \frac{2}{i_y(4+Y_y+G_y)}, \quad a_z^E = \frac{2}{i_z(4+Y_z+G_z)},$$
$$a_x^H = \frac{2}{i_x(4+Z_x+R_x)}, \quad a_z^H = \frac{2}{i_y(4+Z_y+R_y)}, \quad a_z^H = \frac{2}{i_z(4+Z_z+R_z)}. \tag{2.109}$$

The operator $\boldsymbol{M}_{stubs}^{F\to a} : \mathcal{H}_F \longrightarrow \mathcal{H}_\nu$ can be presented as

$$\boldsymbol{M}_{stub}^{F\to a} = \begin{bmatrix} 0 & e_y & 0 & 0 & 0 & h_z \\ 0 & e_y & 0 & 0 & 0 & -h_z \\ 0 & 0 & e_z & 0 & -h_y & 0 \\ 0 & 0 & e_z & 0 & h_y & 0 \\ 0 & 0 & e_z & h_x & 0 & 0 \\ 0 & 0 & e_z & -h_x & 0 & 0 \\ e_x & 0 & 0 & 0 & 0 & -h_z \\ e_x & 0 & 0 & 0 & 0 & h_z \\ e_x & 0 & 0 & 0 & h_y & 0 \\ e_x & 0 & 0 & 0 & -h_y & 0 \\ 0 & e_y & 0 & -h_x & 0 & 0 \\ 0 & e_y & 0 & h_x & 0 & 0 \\ e_x\sqrt{Y_x} & 0 & 0 & 0 & 0 & 0 \\ 0 & e_y\sqrt{Y_y} & 0 & 0 & 0 & 0 \\ 0 & 0 & e_z\sqrt{Y_z} & 0 & 0 & 0 \\ 0 & 0 & 0 & h_x\sqrt{Z_x} & 0 & 0 \\ 0 & 0 & 0 & 0 & h_y\sqrt{Z_y} & 0 \\ 0 & 0 & 0 & 0 & 0 & h_z\sqrt{Z_z} \end{bmatrix}, \tag{2.110}$$

with coefficients

$$e_x = \frac{i_x}{2}, \quad e_y = \frac{i_y}{2}, \quad e_z = \frac{i_z}{2}, \tag{2.111}$$
$$h_x = \frac{i_x}{2}Z_0, \quad h_y = \frac{i_y}{2}Z_0, \quad h_z = \frac{i_z}{2}Z_0. \tag{2.112}$$

2.5.2 Cell Boundary Mapping

Unlike the centered field mapping (CFM) the electric and magnetic components for cell boundary mapping (CBM) are defined not in the center of the node, but at the ends of the six arms (or the twelve transmission lines). We define the wave amplitudes at the cell boundaries $(l\pm\frac{1}{2},m,n)$, $(l,m\pm\frac{1}{2},n)$ and $(l,m,n\pm\frac{1}{2})$ as

$$|\boldsymbol{a}\rangle = 1/2\left(-|\boldsymbol{E}\rangle + Z\cdot\boldsymbol{n}\times|\boldsymbol{H}\rangle\right), \tag{2.113}$$

$$|\boldsymbol{b}\rangle = -1/2\left(|\boldsymbol{E}\rangle + Z\cdot\boldsymbol{n}\times|\boldsymbol{H}\rangle\right), \tag{2.114}$$

where $\boldsymbol{n}$ is the normal unity vector at a cell boundary, $|\boldsymbol{E}\rangle$ and $|\boldsymbol{H}\rangle$ are electric and magnetic vectors sampled at a boundary.

For example, we take a look at the mapping operation for the arm $(l+\frac{1}{2},m,n)$ at the time k. From eq. (2.113) and (2.114) we obtain following relations for

$\boldsymbol{n} = [1,0,0]^T$

$$
\begin{aligned}
&{}_k\left[a_x\right]_{l+1/2,m,n} = 0,\\
&{}_k\left[a_y\right]_{l+1/2,m,n} = {}_k\left[a_2\right]_{l,m,n} = \frac{\Delta l}{2}\left(i_y\,{}_kE^y_{l+1/2,m,n} - i_z Z\,{}_kH^z_{l+1/2,m,n}\right),\\
&{}_k\left[a_z\right]_{l+1/2,m,n} = {}_k\left[a_4\right]_{l,m,n} = \frac{\Delta l}{2}\left(i_z\,{}_kE^z_{l+1/2,m,n} + i_y Z\,{}_kH^y_{l+1/2,m,n}\right),\\
&{}_k\left[b_x\right]_{l+1/2,m,n} = 0,\\
&{}_k\left[b_y\right]_{l+1/2,m,n} = {}_k\left[b_2\right]_{l,m,n} = \frac{\Delta l}{2}\left(i_y\,{}_kE^y_{l+1/2,m,n} + i_z Z\,{}_kH^z_{l+1/2,m,n}\right),\\
&{}_k\left[b_z\right]_{l+1/2,m,n} = {}_k\left[b_4\right]_{l,m,n} = \frac{\Delta l}{2}\left(i_z\,{}_kE^z_{l+1/2,m,n} - i_y Z\,{}_kH^y_{l+1/2,m,n}\right)
\end{aligned}
\tag{2.115}
$$

and

$$
\begin{aligned}
{}_kE^y_{l+1/2,m,n} &= \frac{1}{i_y\Delta l}\left({}_k\left[a_2\right]_{l,m,n} + {}_k\left[b_2\right]_{l,m,n}\right),\\
{}_kE^z_{l+1/2,m,n} &= \frac{1}{i_z\Delta l}\left({}_k\left[a_4\right]_{l,m,n} + {}_k\left[b_4\right]_{l,m,n}\right),\\
{}_kH^y_{l+1/2,m,n} &= \frac{1}{i_y\Delta l Z}\left({}_k\left[a_4\right]_{l,m,n} - {}_k\left[b_4\right]_{l,m,n}\right),\\
{}_kH^z_{l+1/2,m,n} &= \frac{1}{i_z\Delta l Z}\left({}_k\left[b_2\right]_{l,m,n} - {}_k\left[a_2\right]_{l,m,n}\right),
\end{aligned}
\tag{2.116}
$$

respectively.

In order to obtain CBM in a matrix form we define the electromagnetic field state vectors in $\mathcal{H}_F^{12} = C^{12}\otimes\mathcal{H}_\mu$ as

$$
|{}_k\boldsymbol{F}_E\rangle_{\frac{1}{2}} = \sum_{l,m,n=-\infty}^{+\infty}
\begin{bmatrix}
{}_kE^y_{l-1/2,m,n}\\
{}_kE^y_{l+1/2,m,n}\\
{}_kE^z_{l-1/2,m,n}\\
{}_kE^z_{l+1/2,m,n}\\
{}_kE^z_{l,m-1/2,n}\\
{}_kE^z_{l,m+1/2,n}\\
{}_kE^x_{l,m-1/2,n}\\
{}_kE^x_{l,m+1/2,n}\\
{}_kE^x_{l,m,n-1/2}\\
{}_kE^x_{l,m,n+1/2}\\
{}_kE^y_{l,m,n-1/2}\\
{}_kE^y_{l,m,n+1/2}
\end{bmatrix}
|l,m,n\rangle\,,
\tag{2.117}
$$

$$|{}_k\boldsymbol{F}_M\rangle_{\frac{1}{2}} = Z \sum_{l,m,n=-\infty}^{+\infty} \begin{bmatrix} {}_kH^z_{l-1/2,m,n} \\ {}_kH^z_{l+1/2,m,n} \\ {}_kH^y_{l-1/2,m,n} \\ {}_kH^y_{l+1/2,m,n} \\ {}_kH^x_{l,m-1/2,n} \\ {}_kH^x_{l,m+1/2,n} \\ {}_kH^z_{l,m-1/2,n} \\ {}_kH^z_{l,m+1/2,n} \\ {}_kH^y_{l,m,n-1/2} \\ {}_kH^y_{l,m,n+1/2} \\ {}_kH^x_{l,m,n-1/2} \\ {}_kH^x_{l,m,n+1/2} \end{bmatrix} |l,m,n\rangle\,. \tag{2.118}$$

From eq. (2.113) and (2.114) we obtain the mapping operations in the next operator expressions

$$\begin{aligned} |{}_k\boldsymbol{F}_E\rangle_{\frac{1}{2}} &= \frac{1}{\Delta l}\boldsymbol{M}_E^{-1}\left(|{}_k\boldsymbol{a}\rangle + |{}_k\boldsymbol{b}\rangle\right), \\ |{}_k\boldsymbol{F}_H\rangle_{\frac{1}{2}} &= \frac{1}{\Delta l}\boldsymbol{M}_H^{-1}\left(|{}_k\boldsymbol{a}\rangle - |{}_k\boldsymbol{b}\rangle\right), \end{aligned} \tag{2.119}$$

and

$$\begin{aligned} |{}_k\boldsymbol{a}\rangle &= \frac{\Delta l}{2}\left(\boldsymbol{M}_E\,|{}_k\boldsymbol{F}_E\rangle_{\frac{1}{2}} + \boldsymbol{M}_H\,|{}_k\boldsymbol{F}_H\rangle_{\frac{1}{2}}\right), \\ |{}_k\boldsymbol{b}\rangle &= \frac{\Delta l}{2}\left(\boldsymbol{M}_E\,|{}_k\boldsymbol{F}_E\rangle_{\frac{1}{2}} - \boldsymbol{M}_H\,|{}_k\boldsymbol{F}_H\rangle_{\frac{1}{2}}\right), \end{aligned} \tag{2.120}$$

where the operators $\boldsymbol{M}_E : \mathcal{H}_F^{12} \longrightarrow \mathcal{H}_\nu$ and $\boldsymbol{M}_H : \mathcal{H}_F^{12} \longrightarrow \mathcal{H}_\nu$ are given as

$$\boldsymbol{M}_E = \mathrm{diag}\left[i_y, i_y, i_z, i_z, i_z, i_z, i_x, i_x, i_x, i_x, i_y, i_y\right] \tag{2.121}$$

and

$$\boldsymbol{M}_H = \mathrm{diag}\left[i_z, -i_z, -i_y, i_y, i_x, -i_x, -i_z, i_z, i_y, -i_y, -i_x, i_x\right]. \tag{2.122}$$

Since 24 equations are involved in order to establish the relations between the 24 wave amplitudes and the 24 electromagnetic field components, the CBM operation is definite and bijective in contrast to CFM.

We note, that we do not need to utilize the wave amplitudes at stubs during the mapping. Thus, the normalizing presented in eq. (2.52) does not have any influence on CBM.

2.6 The General TLM Operator

We denote the operator ${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}$ in the TLM time evolution scheme

$$|{}_{k+1}\boldsymbol{a}\rangle = {}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\,|{}_k\boldsymbol{a}\rangle\ ,\ k = 0, 1, 2, \ldots\ , \tag{2.123}$$

as the *general TLM operator*, which, in general, is non-unitary and non-symmetric due to the its non-commutativity, i.e.

$${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S} \neq {}_k\boldsymbol{S}\,{}_k\boldsymbol{\Gamma}. \tag{2.124}$$

The non-commutativity property in eq. (2.124) is valid for the composition of two arbitrary operators. In addition, the next theorem is valid for the TLM-system, which leads to a very interesting interpretation of the TLM time evolution scheme.

Theorem 2.6.1. Let us assume that eq. (2.123) describes the time evolution of the discretized electromagnetic field in the infinite free space. Then the general TLM operator ${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}$ is non-symmetric.

Proof. *a*) In the contest of the theorem we have to show that

$$\left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)^T \neq {}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\,. \tag{2.125}$$

We start with the transpose general TLM operator,

$$\left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)^T = \left({}_k\boldsymbol{S}\right)^T \left({}_k\boldsymbol{\Gamma}\right)^T = {}_k\boldsymbol{S}\,{}_k\boldsymbol{\Gamma}\,. \tag{2.126}$$

The last equation follows from the symmetry of the scattering and connection operators. Because a matrix product is generally non-commutative and due to the eq. (2.124) we have that the general TLM operator is generally non-symmetric.

b) However, in special cases two matrices can be commutative. Next we consider, what happens if we suppose the general TLM operator to be symmetric. Consequently, we have

$${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S} = {}_k\boldsymbol{S}\,{}_k\boldsymbol{\Gamma} = \left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)^{-1}. \tag{2.127}$$

From eq. (2.123) follows that

$$\begin{aligned} |{}_{k+2}\boldsymbol{a}\rangle = {}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\,|{}_{k+1}\boldsymbol{a}\rangle &= \left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)\left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)|{}_k\boldsymbol{a}\rangle = \\ = \left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)\left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)^{-1}|{}_k\boldsymbol{a}\rangle &= |{}_k\boldsymbol{a}\rangle\,. \end{aligned} \tag{2.128}$$

This way we have that after each two time steps the TLM system is returned to its initial state $|{}_0\boldsymbol{a}\rangle$, or in other words, the electromagnetic waves are not propagating at all. In this case

$$\begin{aligned} |{}_{2k+1}\boldsymbol{a}\rangle &= {}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\,|{}_0\boldsymbol{a}\rangle\,, \\ |{}_{2k}\boldsymbol{a}\rangle &= |{}_0\boldsymbol{a}\rangle\,, \end{aligned} \tag{2.129}$$

and our supposition ${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}$ being symmetric is false. □

Table 2.1: Properties of operators in TLM.

Operator	lossless structure	lossy structure
${}_k\boldsymbol{S}$	unitary, symmetric	non-unitary, symmetric[1]
${}_k\boldsymbol{\Gamma}$	unitary, symmetric	non-unitary, symmetric
${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}$	unitary, non-symmetric	non-unitary, non-symmetric

[1] The modified scattering operator is considered

The important properties of the scattering, connection and TLM operators are summarized in TABLE 2.6 The scheme (2.123) can be written in the form

$$|{}_{k+1}\boldsymbol{a}\rangle = \left({}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}\right)^k |{}_0\boldsymbol{a}\rangle\,,\ k = 0, 1, 2, \ldots\,. \tag{2.130}$$

The two-level time evolution scheme (2.123) is said to be stable with respect to the norm $||\cdot||$ (see [94, 95]) if

$$||_{k+1}\boldsymbol{a}|| \leq \|(_k\boldsymbol{\Gamma}\,_k\boldsymbol{S})^k\|\,||_0\boldsymbol{a}\,|| \tag{2.131}$$

for $0 < \Delta l < \Delta l_0$ and $0 < t < t_0$. The passivity of the TLM scheme (2.130) is guaranteed by satisfaction of (2.131) as well.

Introducing the eigenvalue composition of the TLM operator as

$$_k\boldsymbol{\Gamma}\,_k\boldsymbol{S} = {}_k\boldsymbol{U}\,_k\boldsymbol{\Lambda}\,_k\boldsymbol{U}^{-1} \tag{2.132}$$

we can write eq.(2.131) in the next form

$$||_{k+1}\boldsymbol{a}|| \leq \|_k\boldsymbol{U}(_k\boldsymbol{\Lambda})^k\,_k\boldsymbol{U}^{-1}\|\,||_0\boldsymbol{a}\,||\,. \tag{2.133}$$

Eq. (2.131) is fulfilled if eigenvalues of the matrix $_k\boldsymbol{\Gamma}\,_k\boldsymbol{S}$ have a magnitude less or equal to zero, i.e. $|\lambda_i| \leq 1, i = 1..N$.

Convergence and stability are connected via the LAX theorem, which states that a consistent two-level difference scheme for a well-posed linear initial-value problem is convergent, if and only if it is stable. For more details see [114].

Consequently we can summarize convergence, stability and passivity as following.

Proposition 2.6.1. The stability, passivity and convergence of the TLM time evolution scheme given by eq. (2.123) are guaranteed if the eigenvalues of the general TLM operator $_k\boldsymbol{\Gamma}\,_k\boldsymbol{S}$ have magnitudes less or equal to one.

Proof. See Chapter 5. □

2.7 Practical Issues

2.7.1 Alternative Scattering Algorithms

In addition to the JOHNS scattering matrix there is an alternative scattering procedure trough total node values. The scattering algorithm from JOHNS requires 108 additions/ subtractions operation and 75 multiplications for one scattering operation. HERRING proposed in [54] an alternative scattering process requiring 54 additions/subtractions and 12 multiplications that correspond to a reduction of the values to be stored from 36 to 12. Through algebraic manipulations it can be proofed that two scattering formulations are equivalent.

In the present doctoral work the HERRING'S scattering process is used as a basis for TLM program implementations. It does not affect on the MOR theory presented further, but allows to decrease computational effort per scattering operation.

Using the charge conservation law we obtain the total voltages V_x, V_y and V_z at the SCN as

$$V_x = \frac{2(a_7 + a_8 + a_9 + a_{10} + \sqrt{Y_x}a_{13})}{4 + Y_x + G_x}, \tag{2.134}$$

$$V_y = \frac{2(a_1 + a_2 + a_{11} + a_{12} + \sqrt{Y_y}a_{14})}{4 + Y_y + G_y}, \tag{2.135}$$

$$V_z = \frac{2(a_3 + a_4 + a_5 + a_6 + \sqrt{Y_z}a_{15})}{4 + Y_z + G_z}. \tag{2.136}$$

Involving the magnetic flux conservation law we have the total ring currents I_x, I_y and I_z at the SCN in

$$Z_0 I_x = \frac{2(a_5 - a_6 - a_{11} + a_{12} + \sqrt{Z_x} a_{16})}{4 + Z_x + R_x}, \tag{2.137}$$

$$Z_0 I_y = \frac{2(-a_3 + a_4 + a_9 - a_{10} + \sqrt{Z_y} a_{17})}{4 + Z_y + R_y}, \tag{2.138}$$

$$Z_0 I_z = \frac{2(a_1 - a_2 - a_7 + a_8 + \sqrt{Z_z} a_{18})}{4 + Z_z + R_z}. \tag{2.139}$$

Taking eq. (2.134) and (2.137) into account we obtain from the next relations for total voltages and currents

$$V = a + b \quad \text{and} \quad Z_0 I = a - b\,, \tag{2.140}$$

and twelve scattered wave amplitudes by algebraic manipulations

$$b_1 = V_y - Z_0 I_z - a_2, \tag{2.141}$$
$$b_2 = V_y + Z_0 I_z - a_1, \tag{2.142}$$
$$b_3 = V_z + Z_0 I_y - a_4, \tag{2.143}$$
$$b_4 = V_z - Z_0 I_y - a_3, \tag{2.144}$$

$$b_5 = V_z - Z_0 I_x - a_6, \tag{2.145}$$
$$b_6 = V_z + Z_0 I_x - a_5, \tag{2.146}$$
$$b_7 = V_x + Z_0 I_z - a_8, \tag{2.147}$$
$$b_8 = V_x - Z_0 I_z - a_7, \tag{2.148}$$

$$b_9 = V_x - Z_0 I_y - a_{10}, \tag{2.149}$$
$$b_{10} = V_x + Z_0 I_y - a_9, \tag{2.150}$$
$$b_{11} = V_y + Z_0 I_x - a_{12}, \tag{2.151}$$
$$b_{12} = V_y - Z_0 I_x - a_{11}, \tag{2.152}$$

$$b_{13} = \sqrt{Y_x} V_x - a_{13}, \tag{2.153}$$
$$b_{14} = \sqrt{Y_y} V_y - a_{14}, \tag{2.154}$$
$$b_{15} = \sqrt{Y_z} V_z - a_{15}, \tag{2.155}$$

$$b_{16} = -\sqrt{Z_x} Z_0 I_x + a_{16}, \tag{2.156}$$
$$b_{17} = -\sqrt{Z_y} Z_0 I_y + a_{17}, \tag{2.157}$$
$$b_{18} = -\sqrt{Z_z} Z_0 I_z + a_{18}. \tag{2.158}$$

In order to save connection operations for the short stubs 16, 17 and 18 corresponding to the negative reflection at the stub ends, the last three reflected wave amplitudes have to be taken as their opposite, or to be precise

$$b_{16} = -(-\sqrt{Z_x} Z_0 I_x + a_{16}), \tag{2.159}$$
$$b_{17} = -(-\sqrt{Z_y} Z_0 I_y + a_{17}), \tag{2.160}$$
$$b_{18} = -(-\sqrt{Z_z} Z_0 I_z + a_{18}). \tag{2.161}$$

The above presented modification of the scattering algorithm is used for the straight forward TLM scheme dealing with the ($\boldsymbol{\Gamma S}$) operator. If the scattering procedure follows after the connection (the backward TLM scheme with the operator ($\boldsymbol{S\Gamma}$) is necessary for ROM realizations based on the non-symmetric LANCZOS algorithm) the last three reflected stub amplitudes have to be exchanged for

$$b_{16} = -\sqrt{Z_x} Z_0 I_x - a_{16}, \tag{2.162}$$
$$b_{17} = -\sqrt{Z_y} Z_0 I_y - a_{17}, \tag{2.163}$$
$$b_{18} = -\sqrt{Z_z} Z_0 I_z - a_{18}, \tag{2.164}$$

that allows to save negative reflection at the short stubs. In addition, the total ring currents have to be calculated as

$$Z_0 I_x = \frac{2(a_5 - a_6 - a_{11} + a_{12} - \sqrt{Z_x} a_{16})}{4 + Z_x + R_x}, \tag{2.165}$$
$$Z_0 I_y = \frac{2(-a_3 + a_4 + a_9 - a_{10} - \sqrt{Z_y} a_{17})}{4 + Z_y + R_y}, \tag{2.166}$$
$$Z_0 I_z = \frac{2(a_1 - a_2 - a_7 + a_8 - \sqrt{Z_z} a_{18})}{4 + Z_z + R_z}. \tag{2.167}$$

2.7.2 Boundary Conditions

In practice an electromagnetic structure is embedded in a finite computational region with boundary conditions. While PEC/PMC do not pose a challenge, the modeling of radiating boundary conditions remains a big issue. Improper absorbing boundary conditions result in back scattering of electromagnetic waves from boundaries into computational region. These back scattered waves interfere with propagating modes and effect the field solution if not corrupt it. Potentially the impact of back scattering gives rise to wrong interpretations of gained results. For these reasons a lot of different methods are proposed to model radiating conditions of an open structure. So the discretization of the analytical conditions [83] and TAYLOR'S expansion for the plane wave solution [21] are described in literature. Application of GREEN'S functions to model bounded regions with various media is proposed [66], but in general the calculation of GREEN'S functions is a complicated procedure that limits their usage. One of the best choice for truncating media are matched layered absorbers, where electric and magnetic losses are introduced in layers to absorb incident electromagnetic waves. Artificial *Perfectly matched layers* (PML) in the TLM (see [38], [77] and [40]) allow to achieve better absorption than matched layer absorbers, but their instabilities make the usage of unconditionally stable *matched layers* (ML) more attractive. The simplest absorbing boundary conditions like a matched load with the zero reflection coefficient (a reflection free wall) yield very good approximations in many cases as well. However these methods only yield approximate solutions of the problem and require some distance between the structure to be modeled and the boundary of the computational domain. The potentially exact TLM-Multipole Expansion (TLM-ME) approach, where the electromagnetic field outside of the spherical boundary of the computational domain is expanded into spherical waves originating at the center of the sphere, is proposed and studied in [72].

In the framework of present dissertation three kinds of boundary conditions are presented - one- and two-side reflection walls, ML and TLM-ME. It does not make a big difference, which kind is used for MOR-techniques, because boundary conditions change the structure of the connection operator but not its properties. For more details see the section about the connection operator.

One-Sided Reflection Wall

In order to model absorbing boundary conditions a matched reflection free wall must have the surface impedance

$$Z_{load} = \sqrt{\frac{\mu_r}{\varepsilon_r}} Z_0 \,. \tag{2.168}$$

Incident waves falling normal to the wall with square cells and with the impedance Z_{load} are fully absorbed. In general, the local reflection coefficient is dependent on sizes of the wall cell and can be expressed through the effective reflection coefficient (the wanted reflection coefficient) as

$$r_{loc} = \frac{r_0 + r_{eff}}{1 + r_0 r_{eff}} \,, \tag{2.169}$$

where

$$r_0 = \frac{\frac{i_\parallel}{i_\perp}\sqrt{\frac{\mu_r}{\varepsilon_r}} - 1}{\frac{i_\parallel}{i_\perp}\sqrt{\frac{\mu_r}{\varepsilon_r}} + 1} \tag{2.170}$$

with normed cell sizes

$$i_\parallel \in \{i_x, i_y, i_z\} \,, \quad i_\parallel \parallel a_i \,, \tag{2.171}$$

$$i_\perp \in \{i_x, i_y, i_z\} \,, \quad i_\perp \perp a_i \,. \tag{2.172}$$

The variables $i_\parallel$ and $i_\perp$ mean the cell sizes in directions parallel and perpendicular to the incident wave amplitude respectively. For example, if the a_4 amplitude is considered, in accordance with Fig. 2.2 we have $i_\parallel = i_z$ and $i_\perp = i_y$. We note, that even though the effective reflection coefficient is set to zero, the local reflection is not, that, actually, conserves the TLM nature. For PEC/PMC we have $r_{eff} = \pm 1$ that yields $r_{loc} = \pm 1$.

Boundary conditions with fixed reflection coefficients r_{loc} do not take into consideration the dependence on the falling wave angle and, thus, present a first order approximation. However, they can be efficiently used for treatment of numerous electromagnetic problems.

Two-Sided Reflection-Transmission Wall

A two-side reflection-transmission wall is used if a very thin layer with constant surface impedance has to be modeled. Actually, a layer with thickness d can be simulated only by one node. However, in this case the smallest space step is often defined by the thickness of the layer, i.e. $\Delta l = d$ that leads to a significant increase of computational effort. The use of a two-sided reflection wall allows to analyze electromagnetic structures with discretization of lower order and, therefore, to save computational time and memory. This wall is represented via

additional boundary conditions, which model a change of the wave impedance. One part of an incident wave is reflected back and the second one is transmitted through the wall to the adjacent cell. In the case, a thin layer is discretized, the wave impedance jump is taken into account through material properties in process of the connection, and the loss in the thin layer are considered in the scattering matrix (or operation).

The required reflection r_{loc} and transmission t_{loc} coefficients are calculated as

$$r_{loc} = -\frac{1}{1 + 2Y_0\rho_\square \frac{i_\parallel}{i_\perp}}\,, \tag{2.173}$$

$$t_{loc} = \frac{Y_0\rho_\square}{Y_0\rho_\square + \frac{i_\parallel}{i_\perp}}\,, \tag{2.174}$$

where $\rho_\square$ is the surface impedance of the thin layer to be modeled. If the skin-effect plays a big role, then more complicated models have to be used, e.g. in [58].

The two-sided reflection-transmission wall can be sensibly and efficiently used for analysis of planar circuits, especially for microstrip line based types. However, the situation is more complicated, if coplanar waveguide based circuits have to be simulated.

Matched Layer Absorbers

In matched layers (MLs) the loss of the absorbing material raises from layer to layer, whereas the wave impedance is kept constant, or to be precise

$$Z_{ML_1} = \sqrt{\frac{\mu}{\varepsilon}} \quad = \quad Z_{ML_i} = \sqrt{\frac{\mu + \frac{\sigma_{mi}}{j\omega}}{\varepsilon + \frac{\sigma_{ei}}{j\omega}}}\,. \tag{2.175}$$

The eq. (2.175) holds if

$$\frac{\mu}{\varepsilon} \quad = \quad \frac{\sigma_{mi}}{\sigma_{ei}}\,. \tag{2.176}$$

The electric conductivity in the matched layers is increased according to

$$\sigma_i = \sigma_{max}\left(\frac{i}{N_{ML}}\right)^p\,, \tag{2.177}$$

where i denotes the index of the ML, σ_{max} is the maximum conductivity in the outer layer N_{ML} which is calculated as

$$\sigma_{max} = -\frac{\alpha(p+1)\ln(R_0)}{2N_{ML}\Delta l Z_0}\,. \tag{2.178}$$

Here α is a constant changing dispersion behavior in neighboring absorbing layers with different media parameters, R_0 is theoretical reflection, p denotes the profile coefficients. Comprehensive investigations on the optimal parameter combinations are presented in [94]. We note that from practical investigation it is known that $\alpha = 0.1$, $R_0 = 10^{-4}$ and $p = 1$ give the best performance for the matched 5-layers absorber. That is a matter of course, the thicker the absorbing layer and the flatter the profile are the better absorption will be achieved.

For one matched layer with j inhomogeneous materials the next expression is valid

$$\frac{\sigma_e^1}{\varepsilon^1} = \frac{\sigma_m^1}{\mu^1} = \cdots = \frac{\sigma_e^j}{\varepsilon^j} = \frac{\sigma_m^j}{\mu^j} \,. \tag{2.179}$$

Even though the matched layer absorbers can be placed relative close to a structure to be simulated, their implementation requires 5-15 additional cells at each wall that can cause significantly slower convergence of MOR-algorithms. Since the scattering and connection operations have to be performed in the absorbing layers as well, the size of the wave amplitude state vector is longer by the total number of additional cells

$$\underbrace{2N_{ML}(lm+mn+ln)}_{\text{plate regions}} + \underbrace{4N_{ML}^2(l+m+n)+8N_{ML}^3}_{\text{edge regions}} \tag{2.180}$$

multiplied by 12 or 18 wave amplitudes, e.g. we have to treat 126 000 variables with or 54 000 without edge regions more in case of 5 MLs for a $10 \times 10 \times 10$ discretized structure with losses. We note that edge regions are intersections of plate regions and require different parameters.

The TLM-ME Approach

The above-mentioned methods allow to put boundary conditions up to near field regions, however, these methods only yield approximate solutions of the radiating boundary condition problem and always require some distance between the structure to be modeled and the boundary of the computational domain. To model radiating electromagnetic structures a hybridization TLM-ME is introduced in [72]. A simulation region is divided into two sub-domains: inside of a sphere and outside of it. Within a spherical region complex electromagnetic structures are modeled with the TLM method. Outside the spherical region the field is expanded into spherical waves. At the boundaries of the spherical region the TLM solution is matched to the multi-pole expansion (ME) of the field in the outer region. This yields a potentially exact modeling of the radiating boundary conditions and allow to get closer to a structure to be analyzed, and even into the near field. To perform the TLM simulation inside the spherical computation domain the sphere is approximated by cubic TLM cells as shown in Fig. 2.3.

The advantage of this TLM-ME method is that it is potentially exact and that the basis functions of this expansion are orthogonal. This yields a compact lumped element equivalent circuit representation of the radiating modes where every radiating mode is described by a lumped element LC ladder network terminated with the radiation resistor. The connection of the outer boundaries of the TLM cells with the spherical modes is established by a circuit consisting of ideal transformers only.

The discretized TLM electromagnetic field at all nodes on the approximated sphere (Fig. 2.3) in the cartesian coordinate system can be transformed into transverse spherical field. We denote this transformed TLM electromagnetic field at time k as

$${}_k\boldsymbol{F}_t = [{}_k\boldsymbol{E}_\vartheta, {}_k\boldsymbol{E}_\varphi, {}_k\boldsymbol{H}_\vartheta, {}_k\boldsymbol{H}_\varphi]^T \,. \tag{2.181}$$

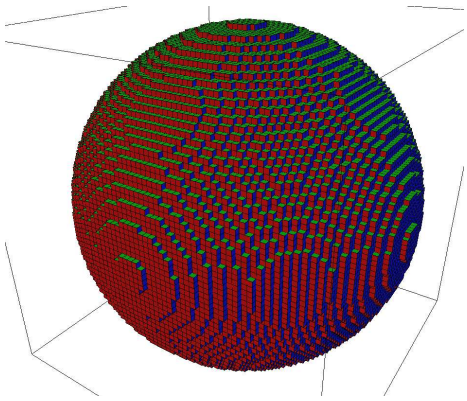


Figure 2.3: A spherical simulation domain.

The transverse electromagnetic field obtained by the TLM can be expressed at the sphere boundary as a linear combinations of the basis functions of the even (e) and odd (o) radiating modes according to

$$ {}_k\boldsymbol{F}_t = \sum_{n,m} {}_k\boldsymbol{\alpha}^e \, \boldsymbol{f}^e + \sum_{n',m'} {}_k\boldsymbol{\alpha}^o \, \boldsymbol{f}^o \,, \tag{2.182} $$

where

$$ \boldsymbol{f}^e = [(e_{mn})^e_\vartheta, (e_{mn})^e_\varphi, (h_{mn})^e_\vartheta, (h_{mn})^e_\varphi] \,, \tag{2.183} $$

$$ \boldsymbol{f}^o = [(e_{mn})^o_\vartheta, (e_{mn})^o_\varphi, (h_{mn})^o_\vartheta, (h_{mn})^o_\varphi] \,. \tag{2.184} $$

The basis functions of the TM radiating modes are given as

$$ \left\{ \begin{array}{c} (e_{mn})^e_\vartheta \\ (e_{mn})^o_\vartheta \end{array} \right\} = -c_{mn} \sin\vartheta \, P_n^{m'}(\cos\vartheta) \left\{ \begin{array}{c} \cos m\varphi \\ \sin m\varphi \end{array} \right\} , \tag{2.185} $$

$$ \left\{ \begin{array}{c} (e_{mn})^e_\varphi \\ (e_{mn})^o_\varphi \end{array} \right\} = c_{mn} \frac{1}{\sin\vartheta} \, m P_n^m(\cos\vartheta) \left\{ \begin{array}{c} -\sin m\varphi \\ \cos m\varphi \end{array} \right\} , \tag{2.186} $$

$$ Z^{TM}_{mn} \, (h_{mn})^e = \boldsymbol{r}_0 \times (e_{mn})^e \,, \tag{2.187} $$

$$ Z^{TM}_{mn} \, (h_{mn})^o = \boldsymbol{r}_0 \times (e_{mn})^o \,, \tag{2.188} $$

and the basis functions of the TE radiating modes are

$$ \left\{ \begin{array}{c} (e_{mn})^e_\vartheta \\ (e_{mn})^o_\vartheta \end{array} \right\} = -c_{mn} \frac{1}{\sin\vartheta} \, m P_n^m(\cos\vartheta) \left\{ \begin{array}{c} -\sin m\varphi \\ \cos m\varphi \end{array} \right\} , \tag{2.189} $$

$$ \left\{ \begin{array}{c} (e_{mn})^e_\varphi \\ (e_{mn})^o_\varphi \end{array} \right\} = -c_{mn} \sin\vartheta \, P_n^{m'}(\cos\vartheta) \left\{ \begin{array}{c} \cos m\varphi \\ \sin m\varphi \end{array} \right\} , \tag{2.190} $$

$$ Z^{TE}_{mn} \, (h_{mn})^e = \boldsymbol{r}_0 \times (e_{mn})^e \,, \tag{2.191} $$

$$ Z^{TE}_{mn} \, (h_{mn})^o = \boldsymbol{r}_0 \times (e_{mn})^o \,. \tag{2.192} $$

Here, c_{mn} are normalization coefficients depending on m and n; $P_n^m(\cos\vartheta)$ and $P_n^{m'}(\cos\vartheta)$ are Legendre polynomials and their derivations with respect to the argument $(\cos\vartheta)$; and r_0 is the unit radius.

Using the orthogonality of the radiating modes the coefficient matrices ${}_k\boldsymbol{\alpha}^e$ and ${}_k\boldsymbol{\alpha}^o$ at the time k can be found through the projection of the radiating modes (actually their basis functions) onto the discretized TLM transverse electromagnetic field ${}_k\boldsymbol{F}_t$ on the sphere.

The wave impedances of the TM radiation modes on the surface of the sphere are given by

$$Z_{nm}^{TM} = \jmath\,\eta \frac{(H_n^{(2)}(kr_0))'}{H_n^{(2)}(kr_0)}\,, \tag{2.193}$$

and of the TE radiation modes by

$$Z_{nm}^{TE} = -\jmath\,\eta \frac{H_n^{(2)}(kr_0)}{(H_n^{(2)}(kr_0))'} \tag{2.194}$$

where $\eta = \sqrt{\mu/\varepsilon}$ is the wave impedance of the plane wave; $H_n^{(2)}(kr_0) = kr_0 h_n^{(2)}(kr_0)$ with $h_n^{(2)}(kr_0)$ being the spherical HANKEL functions of the second kind; $(H_n^{(2)}(kr_0))'$ are derivations with respect to the (kr_0) argument.

Using the recurrence formula for the spherical HANKEL functions a continued fraction expansion of the wave impedances has been proposed in [26] for the outward propagating TM modes as

$$Z_{mn}^{TM} = \eta \left[\frac{n}{jkr} + \cfrac{1}{\frac{2n-1}{jkr} + \cfrac{1}{\frac{2n-3}{jkr} + \ddots + \cfrac{1}{\frac{3}{jkr} + \cfrac{1}{\frac{1}{jkr}+1}}}}} \right] \tag{2.195}$$

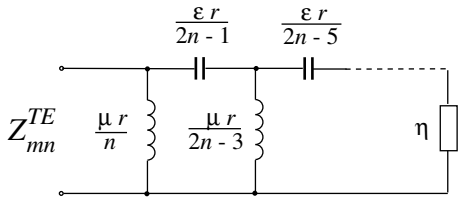
and for outward propagating TE modes as

$$Z_{mn}^{TE} = \eta \left[\cfrac{1}{\frac{n}{jkr} + \cfrac{1}{\frac{2n-1}{jkr} + \cfrac{1}{\frac{2n-3}{jkr} + \cfrac{1}{\frac{2n-5}{jkr} + \ddots + \cfrac{1}{\frac{3}{jkr} + \cfrac{1}{\frac{1}{jkr}+1}}}}}} \right] \tag{2.196}$$

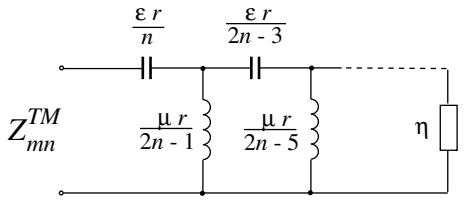
These continued fraction expansions represent the *Cauer canonic realizations* of the outward propagating TM modes (Fig. 2.4(a)) and TE modes (Fig. 2.4(b)). The equivalent circuit expansion of spherical waves is treated in the books of HARRINGTON [49] and FELSEN [42].

We note that the equivalent circuit representing the TE_{mn} mode is dual to the equivalent circuit representing the TM_{mn} mode. The equivalent circuits for the radiation modes exhibit high–pass character. For very low frequencies the wave impedance of the TM_{mn} mode is represented by a capacitor $C_{0n} = \varepsilon r/n$ and the characteristic impedance of the TE_{mn} mode is represented by an inductor $L_{0n} = \mu r/n$. For $f \to \infty$ we obtain Z_{mn}^{TM}, $Z_{mn}^{TE} \to \eta$.

A complete equivalent circuit of the connection between radiation modes and spherical TLM domain also can be synthesized. For this purpose we deal with



(a) An equivalent circuit for TE_{mn} modes



(b) An equivalent circuit for TM_{mn} modes

Figure 2.4: CAUER canonic realizations of the radiation modes.

wave amplitudes directly without computing the electromagnetic fields on the sphere in an explicit form. The spherical basis functions are transformed into cartesian coordinate system. The projection coefficients matrix is computed in time-domain directly in a cartesian coordinate system also. The back reflected to the TLM domain wave amplitudes on the approximated sphere are computed by a transformation network representing the projection operation and connected to the CAUER canonic realization of the mn-radiation mode. The corresponding global effective reflection coefficient in the outer region for the mn-mode is given as

$$(\rho_{mn})_{eff} = \frac{Z_{mn} - \eta}{Z_{mn} + \eta}\,. \tag{2.197}$$

Strictly speaking, the effective reflection coefficient is depending on the frequency and has to be transformed in to the time-domain, i.e.

$$(\rho_{mn})_{eff}(\omega) \longrightarrow (\rho_{mn})_{eff}(t)\,. \tag{2.198}$$

The local reflection coefficients for the inner TLM domain can be obtained through the scattering matrix of the transformation network. Actually, this scattering matrix describes the complete connection between incident and scattered wave amplitudes in two computational domains. As example, the transformer network for the connection of the mn–th radiation mode to the TLM domain for the second and forth transmission lines (for one of the six SCN arms) at the node j is shown in Fig. 2.5. Each arm of the j-th SCN adjacent to the sphere is represented by its two ports, which are linked with the ports of all cells approximating the sphere. For more details the reader is referred to [72] and [71].

The embedding of an electromagnetic structure to be simulated inside the sphere allows to reduce the total number of cells in the computational region and the number of boundary cells in comparison to the usage of reflection free walls as radiating boundary conditions. If a cubic box with the side length $d \to \infty$ is compared with the equal sphere of diameter $d \to \infty$ then the number

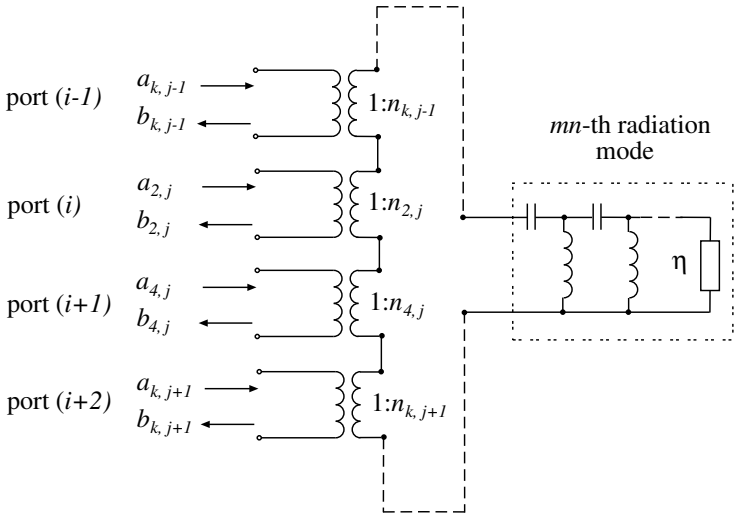


Figure 2.5: Transformer network for the connection of the n–th radiation mode to the TLM domain for the second and forth transmission lines at the node j.

of total and boundary two-dimensional cells is reduced by factor of $6/\pi$ with $\Delta l \to 0$. A comparison for a geometrical approximation for the $10 \times 10 \times 10$ box and equal sphere is given in Table 2.2.

Table 2.2: Comparison of the box and sphere domains.

Structure	Total 3-D cells	Boundary 2-D cells
Box	1 000	600
Sphere	480	456

Comparison of RBC Approximations

From the above-described techniques for modeling of the radiating boundary conditions (RBC) we summarize their important properties here. A brief comparison of the approximations is presented in TABLE 2.3. The influence of RBCs on the MOR convergence is given with reference to the one-sided reflection free wall.

Table 2.3: Comparison of RBC approximations.

Method	Approximation quality	Influence on MOR convergence
Wall	good	no
ML/PML	very good	negative
TLM-ME	very good	positive

2.7.3 Excitation

An electromagnetic structure to be simulated is excited normally by the normalized GAUSS pulse

$$g_g(t) = 2f_g \exp(-\pi(2f_g t)^2) \tag{2.199}$$

with

$$\int_{-\infty}^{\infty} g_n(t)dt = 1\,. \tag{2.200}$$

The spectrum of the normalized GAUSS pulse is given by its FOURIER transform

$$S_g(f) = \exp\left(-\frac{\pi f^2}{4f_g^2}\right)\,. \tag{2.201}$$

The GAUSS pulse shifted in time-domain is given as

$$g_g(t) = 2f_g \exp(-\pi(2f_g(t-t_0))^2)\,. \tag{2.202}$$

In order to shift the spectrum of the excitation signal the Gaussian-modulated sinusoidal can be used, i.e.

$$g_n(t) = 2f_g \exp(-\pi(2f_g(t-t_0))^2) \cdot \sin(2\pi f_c t)\,, \tag{2.203}$$

where f_c is the center frequency, and f_g is a parameter to control the bandwidth and the amplitude of the GAUSS signal.

In order to excite all possible modes, e.g. higher, spurious and intermesh modes, the DIRAC pulse can be applied

$$g_g(t) = \begin{cases} \delta, & \text{if} \quad t = t_0 \\ 0, & \text{if} \quad t \neq t_0 \end{cases}\,. \tag{2.204}$$

Actually, a structure is excited by the voltage and/or by the current pulses corresponding to the basic mode. For some structure like microstrip line and/or coplanar waveguide structures it is very simple to define the input-output voltages and currents corresponding to the fundamental and to several higher modes as well. These voltages and currents can be transformed to the electromagnetic fields and, consequently, to the wave amplitudes, and vice versa. For waveguide structures operating not only with the fundamental mode, the mapping between the input-output voltages (and currents) and the wave amplitudes is difficult to obtain. Since no eigenvalue problems are considered in the time-domain TLM method, the two-dimensional solvers from other frequency-domain-based numerical methods, e.g. Finite-Element or Mode-Matching methods (see Chapter 3), can be utilized in order to obtain the field pattern and the S-parameters (the voltages and currents simultaneously) for each separate mode. These two-dimensional solutions are imposed onto the electromagnetic fields obtained by TLM, and, therefore, the required voltages and currents for each mode can be computed by the integration of the projected electric and/or magnetic fields. For more details the reader is referred to the Chapter 3.

The second possibility to obtain the S-parameters is the use of a reference structure. This approach allows to get the S-parameters either by means of the voltages or via the currents only. The investigated and the reference structure are excited by the voltage (or by the current), and the S-parameters are gained

trough the output incident and scattered voltages (or the currents) in reference planes. For more in-depth details see Chapter 6.

Of course, the network parameters can be directly obtained from TLM simulations, if the field distribution profiles for the fundamental mode or a mode to be extracted can be clearly defined. In this case we have to be sure, that we deal with the currents and/or voltages according to this mode, without spurious and/or other modes have an influence.

In the present doctoral thesis the scattering parameters are extracted by the TLM method only. If desired, external two-dimensional solvers can be easily incorporated into the TLM simulation process and further-described MOR-techniques in order to obtain the needed network parameters for each mode.

Chapter 3

The Hybrid TLM-MM Method

The design of Microwave Monolithic Integrated Circuits (MMICs) requires the exact full-wave modeling of the electromagnetic field in the semiconductor and insulator regions as well as within the conductor layers. The transverse dimensions of the transmission structures in modern MMICs at frequencies above 10 GHz are comparable with the skin depth. Therefore, an accurate simulation of the electromagnetic field inside the conductors is necessary. The usage of conventional simulation tools is not always efficient.

In this chapter the effective *hybrid TLM-MM (Transmission Line Matrix Mode Matching) approach* is proposed for the full-wave analysis of transmission lines and discontinuities realized in damascene technology. The below described hybrid TLM-MM method can be easily involved in the macro-modeling process of MMICs, and the generation of a reduced-order macro-model can be simplified and additionally accelerated.

3.1 Introduction

The use of modern damascene technology with the application of copper as conductor material is a promising way for realization of highly integrated MMICs at frequencies of operation up to several 10 GHz. Compared to aluminium, copper exhibits significantly lower electro-migration, better electrical and thermal conductivity [53, 123]. The higher current density achievable in copper transmission lines and the application of dual damascene process allow a decrease in the cross-section of the conductors down to 0.2 μm width and 0.2 μm height. The thin copper conductors offer improvements in the size of MMICs and power consumption. The schematic view of an usual multilayered MMIC in damascene technology is shown in Fig. 3.1.

Usual design tools for planar circuit design are not applicable, as most of them do not consider finite thickness of the metallization and loss appropriately. The methods working in time-domain require very small discretization and time steps, therefore they are not efficient due to the high computational effort and the high memory consumption. In addition, an implementation of appropriate skin effect models is necessary [58].

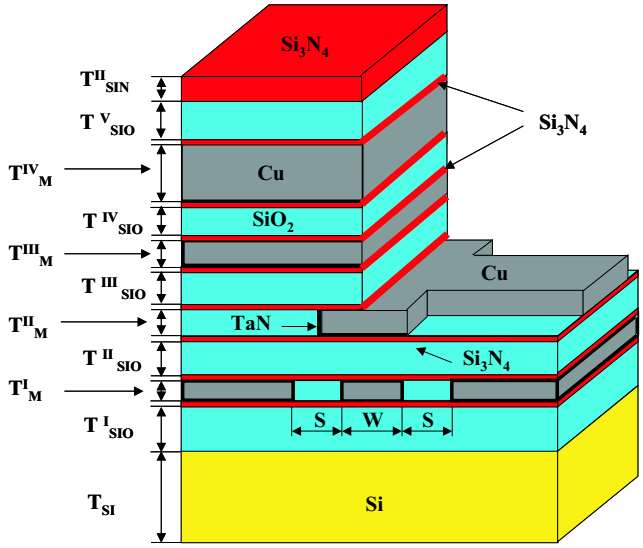


Figure 3.1: Schematic view of a multilayered MMIC in damascene technology.

A hybrid technique, i.e. a combination of different time/frequency-domain based numerical methods, in several cases yields an accurate and efficient solution of an electromagnetic problem. For example, the proposed couplings of Finite Element Method (FEM)-MM, FDTD-MoM, TLM-Integral Equation (TLM-IE) are powerful and flexible hybrid approaches [5, 30, 92, 91, 113].

In this chapter a novel hybrid time/frequency-domain technique, the combination of a frequency-domain Mode-Matching (MM) method for the modeling of homogeneous waveguide structures with the time-domain TLM method for the simulation of discontinuities, is presented. The MM approach allows an accurate and efficient computation of the loss and dispersion properties of the homogenous (in the wave propagation direction) multilayered structures ([63, 76, 106]), whereas the TLM method is advantageous for the modeling of discontinuities of arbitrary shape [23, 78, 99].

A MMIC in damascene technology exhibits a multilayered Si-based structure containing various transmission line structures, i.e. coplanar waveguides (CPWs), striplines, microstrip lines (MLs), conductor backed CPWs and diverse discontinuities, e.g. crossings, right angle bends and T-junctions (see Fig. 3.1). Such a damascene structure consists of many thin layers like diffusion barriers, inversion layers, channel stoppers, trenches and etc. Their influence and the influence of conductor losses, which are due to the skin effect, on propagation characteristics should be taken into account carefully.

One of the most efficient methods for numerical simulation of such homogeneous two-dimensional (2-D) structures is the frequency-domain based mode matching method. The MM method allows to compute a propagation constant and the corresponding electromagnetic field distribution for each mode in an arbitrary multilayered cross section using the matching of the tangential electromagnetic components at the layer/subdomain boundaries. The fields inside the conductors playing an important role and inside each of the thin layers are fully taken into consideration. However, only a few types of discontinuities can

be analyzed using the MM [106].

The transmission line matrix method is a well-established technique for solving electromagnetic problems in the time-domain (See Chapter 2). The versatility of the TLM method allows straightforward calculation of complex structures. However generally, if structures with highly nonuniform regions and/or curvilinear structures are under consideration, variable and curved meshes are required in order to ensure moderate computer run time and storage consumption. Usually, in a simulation by means of TLM the structure is excited with a GAUSS pulse. The frequency characteristics of the structures, for example S-parameters, can be determined by the Discrete FOURIER Transform (DFT) of the time response or system identification methods.

The goal of the hybrid time/frequency TLM-MM approach is to use the universality of TLM for the simulation of arbitrary shapes and the advantages of the MM method for accurate analysis of homogenous structures and at the same time to overcome the limitations of both methods. The hybrid TLM-MM technique is understood as a combination of the TLM and MM methods in the frequency-domain. The various homogenous transmission structures are simulated by the MM, for the analysis of the complex discontinuities and/or structures the TLM is used. It is shown, that this hybrid method allows to reduce the computation time and the memory requirement in relation to conventional 3-D methods. In addition, the TLM-MM approach takes into account the skin effect losses in the strip lines more exactly. In addition, this method can be easily extended to facilitate macro-modeling of MMICs, in the sense that through its application the generation of a reduced-order macro-model can be simplified and accelerated.

3.2 The Mode Matching Method

In this section a brief description of the Mode-Matching (MM) technique is presented. We start with the HELMHOLTZ equations for electric and magnetic longitudinal section waves in the x-direction (LSE_x and LSH_x) are

$$\begin{aligned} \Delta\Pi_x^e + \omega^2\mu\varepsilon\Pi_x^e &= 0\,, \\ \Delta\Pi_x^m + \omega^2\mu\varepsilon\Pi_x^m &= 0\,. \end{aligned} \tag{3.1}$$

For a wave propagating in the $+z$-direction the x-components of the electric and magnetic HERTZ vectors for the LSE_x and LSH_x waves are

$$\begin{aligned} \Pi_x^e = \sum_m^M &\left[(A_m^e \exp(\jmath k_{xm}^e x) + B_m^e \exp(-\jmath k_{xm}^e x)) \cdot \right. \\ &\left. \cdot (C_m^e \exp(\jmath k_{ym}^e y) + D_m^e \exp(-\jmath k_{ym}^e y))\right] \exp(\jmath(\omega t - k_z z))\,, \end{aligned} \tag{3.2}$$

$$\begin{aligned} \Pi_x^m = \sum_m^M &\left[(A_m^m \exp(\jmath k_{xm}^m x) + B_m^m \exp(-\jmath k_{xm}^m x)) \cdot \right. \\ &\left. \cdot (C_m^m \exp(\jmath k_{ym}^m y) + D_m^m \exp(-\jmath k_{ym}^m y))\right] \exp(\jmath(\omega t - k_z z))\,. \end{aligned} \tag{3.3}$$

And each field component can be given as superposition of the LSE_x and LSH_x waves, or to be precise

$$E_x = \frac{\partial^2 \Pi_x^e}{\partial x^2} + \omega^2 \mu \varepsilon \Pi_x^e \,, \tag{3.4}$$

$$E_y = \frac{\partial^2 \Pi_x^e}{\partial x \partial y} - \jmath\omega\mu \frac{\partial \Pi_x^m}{\partial z} \,, \tag{3.5}$$

$$E_z = \frac{\partial^2 \Pi_x^e}{\partial x \partial z} + \jmath\omega\mu \frac{\partial \Pi_x^m}{\partial y} \,, \tag{3.6}$$

$$H_x = \frac{\partial^2 \Pi_x^m}{\partial x^2} + \omega^2 \mu \varepsilon \Pi_x^m \,, \tag{3.7}$$

$$H_y = \frac{\partial^2 \Pi_x^m}{\partial x \partial y} + \jmath\omega\varepsilon \frac{\partial \Pi_x^e}{\partial z} \,, \tag{3.8}$$

$$H_z = \frac{\partial^2 \Pi_x^m}{\partial x \partial z} - \jmath\omega\varepsilon \frac{\partial \Pi_x^e}{\partial y} \,. \tag{3.9}$$

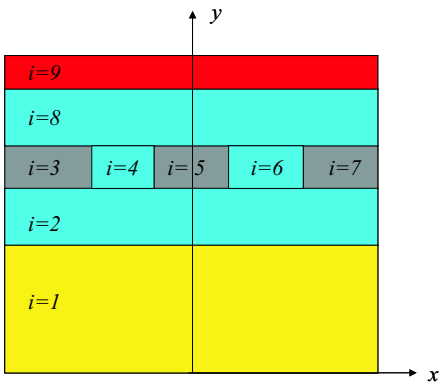


Figure 3.2: Schematic view of a multilayered cross section.

A structure under investigation is embedded in a rectangular box with ideal electric (and/or magnetic) walls and subdivided into homogenous blocs $1, 2 \ldots i \ldots$ as shown in Fig. 3.2. The electromagnetic fields in eq. (3.4) and eq. (3.7) in each sub-domain i are expanded in a finite sum of partial waves M with unknown amplitudes A_m, B_m, C_m, D_m and wave numbers k_{xm}, k_{ym}, k_z to be satisfied

$$k_i^2 = \omega^2 \mu_i \varepsilon_i = k_{ixm}^2 + k_{iym}^2 + k_z^2 \,. \tag{3.10}$$

For example, the E_y- component in the i-th domain may be expanded into m *partial waves* [63] as

$$\begin{aligned} E_y = \sum_m [& -k_{xm} k_{ym} (A_m^e \exp(\jmath k_{xm}^e x) - B_m^e \exp(-\jmath k_{xm}^e x)) \cdot \\ & \cdot (C_m^e \exp(\jmath k_{ym}^e y) - D_m^e \exp(-\jmath k_{ym}^e y)) - \\ & -\omega\mu k_z (A_m^m \exp(\jmath k_{xm}^m x) + B_m^m \exp(-\jmath k_{xm}^m x)) \cdot \\ & \cdot (C_m^m \exp(\jmath k_{ym}^m y) + D_m^m \exp(-\jmath k_{ym}^m y))] \exp(\omega t - k_z z). \end{aligned} \tag{3.11}$$

Utilizing the boundary conditions at the rectangular box walls and applying the Method of Moments (MoM) in order to match the tangential field components at the sub-domain i boundaries (e.g $i = 3$, $j = 4$ and $p = 8$) with normal vectors $\boldsymbol{n}_x$ and $\boldsymbol{n}_y$ in positive x- and y-direction respectively, i.e.

$$\boldsymbol{n}_x \times (\boldsymbol{E}_{ti} - \boldsymbol{E}_{tj}) = 0\,, \qquad \boldsymbol{n}_x \times (\boldsymbol{H}_{ti} - \boldsymbol{H}_{tj}) = 0\,, \tag{3.12}$$

$$\boldsymbol{n}_y \times (\boldsymbol{E}_{ti} - \boldsymbol{E}_{tp}) = 0\,, \qquad \boldsymbol{n}_y \times (\boldsymbol{H}_{ti} - \boldsymbol{H}_{tp}) = 0\,, \tag{3.13}$$

we obtain a system of homogenous equations like

$$\boldsymbol{G} \cdot \boldsymbol{X}_m = 0\,, \tag{3.14}$$

where the vector $\boldsymbol{X}_m$ contains all unknown partial wave amplitudes A_m, B_m, C_m and D_m, the matrix $\boldsymbol{G}$ depending on the unknown wave number k_z summarizes the matching coefficients. A non-trivial solution of the homogenous system of equations (3.14) is obtained by variation of k_z. The eigenvalues of this system of equations are the complex propagation constants k_z for each mode. From the propagation constant the phase constant β and the attenuation constant α can be obtained. After the propagation constant k_z is found we can compute all partial wave amplitudes contained in $\boldsymbol{X}_m$ and as following we have the complete field distribution in a cross section. Using the post-processing the characteristic wave impedance, S-parameters, equivalent network-oriented circuits and etc. can be obtained.

For more details of Mode-Matching technique the reader is referred to the doctoral theses [62, 63, 105, 106].

3.3 The TLM-MM Hybridization

In this section we consider the hybridization process of the TLM and MM methods. For this purpose we operate with electromagnetic fields in the frequency-domain. The transverse electromagnetic fields obtained by TLM method can be expanded into N eigenmodes of a transmission line in a specific section S as

$$\begin{aligned} \boldsymbol{E}_t^{TLM} &= \sum_{n=1}^{N} \gamma_n \boldsymbol{E}_{tn}^{MM}\,, \\ \boldsymbol{H}_t^{TLM} &= \sum_{n=1}^{N} \gamma_n \boldsymbol{H}_{tn}^{MM}\,. \end{aligned} \tag{3.15}$$

During the hybridization process the connection plane S is discretized. The transverse electromagnetic TLM fields in each cell are transformed from time-domain into frequency-domain by the FOURIER transform either for a needed frequency band by means of the FFT engine or for one single frequency by the numerical computation of the FOURIER integral. The second variant is attractive due to its low computational effort, if a few frequency points are to be considered.

It is conveniently to write the transverse field obtained by the MM in the form

$$\begin{aligned} \boldsymbol{E}_{tn}^{MM} &= \boldsymbol{e}_n \exp(-k_{zn} z)\,, \\ \boldsymbol{H}_{tn}^{MM} &= \boldsymbol{h}_n \exp(-k_{zn} z)\,, \end{aligned} \tag{3.16}$$

where $\boldsymbol{e}_n$ and $\boldsymbol{h}_n$ are transverse vector structure functions.

When there are no losses present in the transmission line these structure functions fulfill the orthogonality relation for two eigenmodes n_1 and n_2, i.e.

$$\int_S \int \boldsymbol{e}_{n1} \times \boldsymbol{h}^*_{n2}\, dS = 0\,, \tag{3.17}$$

where $\boldsymbol{h}^*_{n2}$ is the complex conjugate of $\boldsymbol{h}_{n2}$. For its proof see [28]. This relation shows that the total power flowing in a lossless guide is equal to the sum of the power carried by each mode individually. The orthogonality property (3.17) can be written in other form

$$\begin{aligned} \langle \boldsymbol{e}_{n1} \mid \boldsymbol{e}_{n2} \rangle &= \int_S \int \boldsymbol{e}_{n1} \cdot \boldsymbol{e}^*_{n2}\, dS = 0\,, \\ \langle \boldsymbol{h}_{n1} \mid \boldsymbol{h}_{n2} \rangle &= \int_S \int \boldsymbol{h}_{n1} \cdot \boldsymbol{h}^*_{n2}\, dS = 0\,. \end{aligned} \tag{3.18}$$

Therefore, the coefficients γ_n can be obtained by projection of the electromagnetic field of each n-th MM-mode onto the transverse field gained by means of the TLM, or to be precise

$$\gamma_n = \int_S \int \boldsymbol{E}^{TLM}_t \cdot (\boldsymbol{E}^{MM}_{tn})^*\, dS \sim \int_S \int \boldsymbol{H}^{TLM}_t \cdot (\boldsymbol{H}^{MM}_{tn})^*\, dS\,. \tag{3.19}$$

If losses have to be taken into consideration the situation is more complicated, because the desired orthogonality property is lost. In this case the MM-modes do not fulfill relations (3.18) and we can not compute the coefficients γ_n by eq. (3.19). If small losses preexist in a waveguide structure, so that the arising error is accepted, then γ_n can be obtained by eq. (3.19) as well. On the other hand, if we can recover in some way the orthogonality of the modes obtained by the MM, then eq. (3.19) can be directly used. In order to perform this orthogonalization we denote the discretized transverse electromagnetic fields of MM-modes in the section S of a lossy waveguide as vectors $\tilde{\boldsymbol{E}}^{MM}_{tn}$ which spann the vector space $\mathcal{C}^n$, or in other words, each vector $\tilde{\boldsymbol{E}}^{MM}_{tn}$ is in the span of a set of vectors, corresponding to the discretized transverse electromagnetic fields of the MM-eigenmodes.

Sequentially, we define the span of a set of the non-orthogonal linearly independent vectors, corresponding to the computed MM-eigenmodes as

$$\tilde{\boldsymbol{E}}^{MM}_t = \text{span}\,\{\tilde{\boldsymbol{E}}^{MM}_{t1}, \tilde{\boldsymbol{E}}^{MM}_{t2}, \tilde{\boldsymbol{E}}^{MM}_{t3}, \cdots, \tilde{\boldsymbol{E}}^{MM}_{tn}\}\,, \tag{3.20}$$

and the span of a set of the orthonormal vectors as

$$\boldsymbol{E}^{MM}_t = \text{span}\,\{\boldsymbol{E}^{MM}_{t1}, \boldsymbol{E}^{MM}_{t2}, \boldsymbol{E}^{MM}_{t3}, \cdots, \boldsymbol{E}^{MM}_{tn}\}\,. \tag{3.21}$$

Using the GRAM-SCHMIDT orthogonalization process (see APPENDIX B) we can establish the distinct relation between $\boldsymbol{E}^{MM}_t$ and $\tilde{\boldsymbol{E}}^{MM}_t$ as

$$\tilde{\boldsymbol{E}}^{MM}_t = \boldsymbol{E}^{MM}_t\, \boldsymbol{P}\,, \tag{3.22}$$

where $\boldsymbol{P}$ is $n \times n$ upper triangular matrix and invertible. Therefore we obtain the span of the required orthogonal (orthonormal) vectors as

$$\boldsymbol{E}_t^{MM} = \tilde{\boldsymbol{E}}_t^{MM}\, \boldsymbol{P}^{-1}\,. \tag{3.23}$$

In a similar way we can affiliate the span of the orthonormal vectors $\boldsymbol{H}_t^{MM}$ from the span of the vectors $\tilde{\boldsymbol{H}}_t^{MM}$ according with the MM-modes computed for a lossy waveguide, i.e.

$$\boldsymbol{H}_t^{MM} = \tilde{\boldsymbol{H}}_t^{MM}\, \boldsymbol{P}^{-1}\,. \tag{3.24}$$

Thus, $\boldsymbol{E}_t^{MM}$ and $\boldsymbol{H}_t^{MM}$ are fields in a lossless waveguide which are recovered from the appropriate electromagnetic fields in the waveguide with losses. The columns of the spans $\boldsymbol{E}_t^{MM}$ and $\boldsymbol{H}_t^{MM}$ represent the discretized transverse electromagnetic fields of the so-called orthonormalized MM-modes, that allows the expansion of the transverse electromagnetic TLM field into N eigenmodes in accordance with eq. (3.15), where the coefficients γ_n can be obtained by eq. (3.19).

3.4 The TLM-MM Approach

In the following we briefly describe the procedure of the TLM-MM hybrid approach:

1. A structure under investigation is divided into blocks; these consist of discontinuities and transmission lines homogenous in the propagation direction.

2. Multi-mode full-wave analysis of homogenous transmission lines is performed in an interesting frequency band using the MM method.

3. Various discontinuities are simulated by means of TLM with a GAUSS pulse as an excitation.

Of course, the parallel performance of sub-procedures (2) and (3) allows to increase the computational efficiency additionally.

4. The electromagnetic field distribution obtained by TLM method are expanded into N eigenmodes of the transmission line in a specific section

$$\boldsymbol{E}^{TLM} = \sum_{n=1}^{N} \gamma_n \boldsymbol{E}_n^{MM}\,,$$
$$\boldsymbol{H}^{TLM} = \sum_{n=1}^{N} \gamma_n \boldsymbol{H}_n^{MM}\,.$$

   In this section all modes propagating on the transmission line shall be extracted. (Actually, a typical transmission line exhibits the basic propagating mode only.)

5. Next, the S-parameters of the blocks for the propagating modes are extracted.

6. The generalized S-parameter matrix is formed for the initial structure.

Obviously, the accuracy and the computational effort of hybrid method are dependant on the number of eigenmodes taken into account and on the space discretization step in a connection plane. If necessary, the electromagnetic fields for each mode in a specific section can be obtained in time-domain as well.

3.5 Applications

Structures produced by means of the damascene technology with four metal layers have been investigated (see Fig. 3.1). Such a damascene structure consists of eleven main layers, eight thin passivation Si_3N_4-layers and an inversion metal sheet between the Si- and first SiO_2-layers.

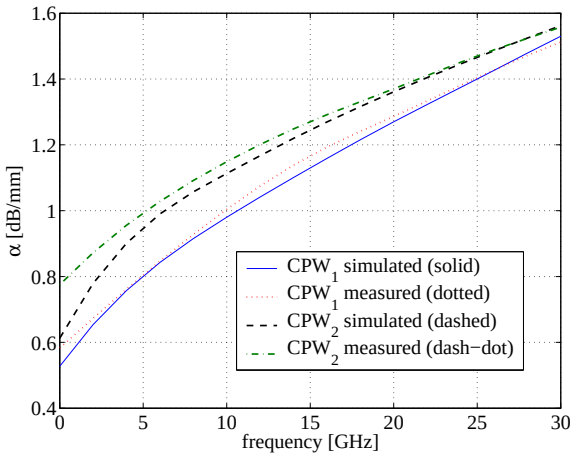
The simulations were performed using the following material parameters: $\varepsilon_{\mathrm{Si}} = 11.9$, $\sigma_{\mathrm{Si}} = 5.5$ S/m, $\varepsilon_{\mathrm{SiO_2}} = 3.9$, $\tan\delta_{\mathrm{SiO_2}} = 10^{-4}$, $\varepsilon_{\mathrm{Si_3N_4}} = 7$, $\tan\delta_{\mathrm{Si_3N_4}} = 10^{-3}$, $\sigma_{\mathrm{Cu}} = 5.9 \cdot 10^7$ S/m.

Vertical dimensions (thicknesses) of the layers are (see Fig. 3.1)) : $T_{\mathrm{Si}} = 380$ µm, $T^{\mathrm{I}}_{\mathrm{SiO_2}} = 1.8$ µm, $T^{\mathrm{I}}_{\mathrm{M}} = 0.4$ µm, $T^{\mathrm{II}}_{\mathrm{SiO_2}} = 0.9$ µm, $T^{\mathrm{II}}_{\mathrm{M}} = 0.6$ µm, $T^{\mathrm{III}}_{\mathrm{SiO_2}} = 0.9$ µm, $T^{\mathrm{III}}_{\mathrm{M}} = 0.6$ µm, $T^{\mathrm{IV}}_{\mathrm{SiO_2}} = 0.9$ µm, $T^{\mathrm{IV}}_{\mathrm{M}} = 2.5$ µm, $T^{\mathrm{V}}_{\mathrm{SiO_2}} = 0.75$ µm, $T^{\mathrm{II}}_{\mathrm{Si_3N_4}} = 0.55$ µm. The thickness of the intermediate passivation layer of Si_3N_4 is $T^{\mathrm{I}}_{\mathrm{Si_3N_4}} = 0.05$ µm.

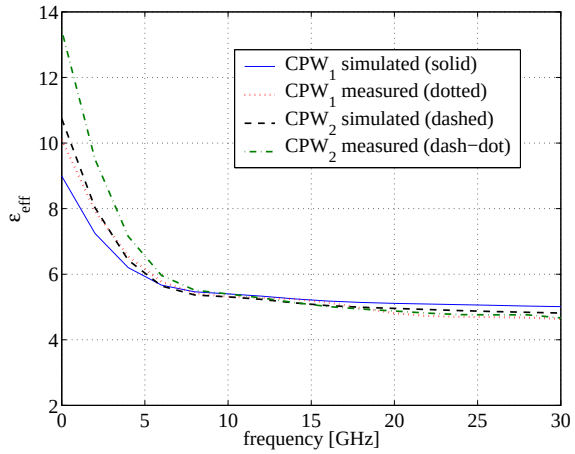
Next, we compare the MM-simulations with the measurements for two CPWs, which are placed in the third 0.6 µm-thick metallization layer. The first CPW has dimensions $W = 7.5$ µm and $S = 5$ µm (see Fig.3.1), the second structure has $W = 5$ µm and $S = 2.5$ µm. The attenuation constant and the effective permittivity of the CPWs are presented in Fig. 3.3. The measurement data are in a good agreement with the simulation results. The attenuation consists of the attenuation caused by the dielectric losses in SiO_2-substrate, the losses in Si-substrate and the ohmic losses in the conductors. The substrate attenuation is lower than 0.05 dB/mm, the attenuation in the conductors plays a dominant role. Fig. 3.3(b) illustrate, that the electromagnetic field concentrates more in the SiO_2 substrate with increasing frequencies. We note that the Si_3N_4 etching stoppers do not influence essentially the propagation characteristics.

The comparison of measurement with simulation results at frequency of 10 GHz for two microstrip lines (MLs) in the third (metallization thickness 0.6 µm, dielectric thickness 2.2 µm) and fourth (metallization thickness 2.5 µm, dielectric thickness 3.7 µm) metallization layers confirms the accuracy of the MM for different geometries. The first metallization layer with a thickness of 0.4 µm corresponds to the ground plane. We changed the width of the lines w from 1 µm to 10 µm and compared the attenuation calculated by MM with the measurement (see Fig. 3.4). The MM-simulation results are very accurate. The wave impedance Z_w of the MLs is shown in Fig. 3.5 as a function of the line width W at 10 GHz as well. Due to the losses in the conductors the imaginary part of the wave impedance can not be neglected (see Fig. 3.5(b)).

In order to gain waveguide parameters by means of the MM only $m = 50$ partial waves (see eq. 3.11) are taken into account, and therefore small computing power is required. At the frequencies up to 10 GHz it is enough to take into account 20 partial waves. We notice, in order to obtain accurate numerical solutions at higher frequencies (above 30 GHz) the number of partial waves should



(a) Attenuation constant



(b) Effective permittivity

Figure 3.3: Attenuation constant and effective permittivity of two CPWs with W=7.5 µm, S=5 µm and W=5 µm, S=2.5 µm, respectively.

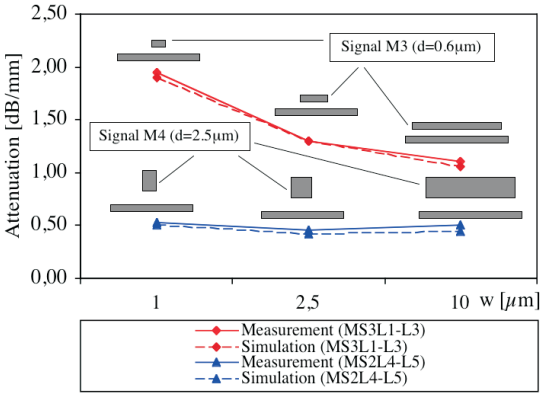
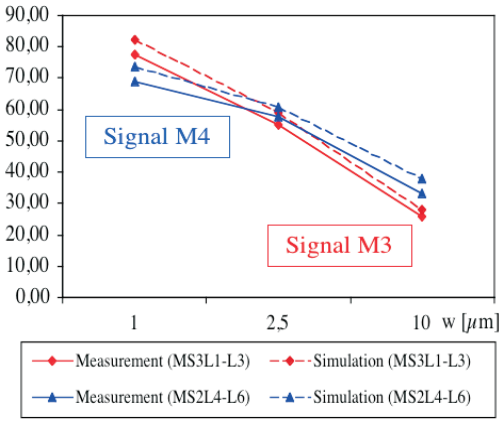
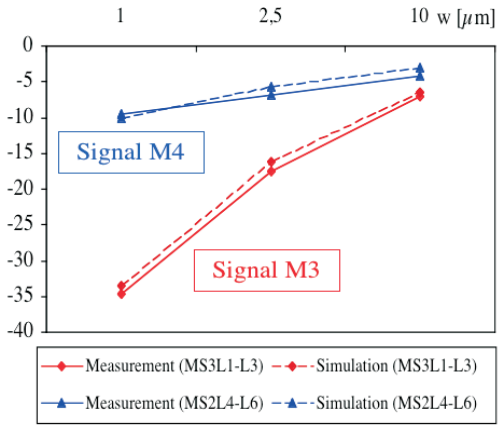


Figure 3.4: Attenuation of different microstrip lines with signal layers in the third and fourth metallization planes.



(a) Re $[Z_w]$, Ω



(b) Im $[Z_w]$, Ω

Figure 3.5: Wave impedance of different microstrip lines with signal layers in the third and fourth metallization planes.

be significantly increased to as many as 100, especially for the accuracy analysis of the attenuation. Due to the symmetry of the waveguides (see Fig. 3.2) we can put a magnetic wall in their center and only one half of a structure can be analyzed, and the computational time can be saved additionally.

Now, we discuss the simulation results for a CPW test structure analyzed by means of TLM and hybrid TLM-MM method consecutively. As a test damascene structure a bridge CPW T-junction is chosen. The layout of the T-junction is drawn in Fig. 3.6. Three bridges placed next to the discontinuities suppress the unwanted odd-mode ([109]). The distance from CPW metal layer to a bridge metal layer (the first metallization layer) is 2.2 μm. Four black rectangles in Fig. 3.6 denote vias, which have transverse (or cross) dimensions of 0.5 μm x 0.5 μm. The CPWs are dimensioned by MM method so that the higher order modes can not propagate, i.e. the CPW exhibits the quasi-TEM mode. First, the whole structure is embedded in a box (called TLM-box) and simulated by TLM only using a space discretization as coarse as 0.5 μm. Then we combine the TLM and MM methods for full-wave analysis of this structure. The structure is subdivided into one TLM-box and three lines (called MM-lines), which are simulated separately. The MM-lines have dimensions $W = 7.5$ μm and $S = 2.5$ μm (see Fig. 3.1). Both methods have been combined efficiently in frequency-domain by means of S-parameters for the fundamental CPW even-mode. Both even and odd propagating modes are taken into account, higher modes are not considered.

A comparison of calculated S-parameters using TLM and hybrid TLM-MM methods is shown in Fig. 3.7. It appears, that the frequency dependant skin effect losses are taken into consideration more accurately with the hybrid method than with the TLM method; in the TLM the conductor is modeled with only one cell of thickness 0.5 μm in order to ensure reasonable computational time. The results indicate very good agreement between the two methods.

Using the hybrid method we obtain accurate results with smaller numerical effort; for homogenous non-resonant structures only few frequency points are needed, an interpolation can be applied to obtain the characteristics in the whole frequency band; the TLM simulation of the small 110 μm x 110 μm box lasts much less than for the initial 255 μm x 185 μm TLM-box. In addition, the simulation of the TLM sub-domain and the MM calculation can be executed independently and, hence, can be run in parallel on different workstations/PCs. Using the hybrid method in parallel performance, the computation time is reduced by factor of two and memory requirements in this case are reduced by more than a factor of three. We note, that the computational time by the MM-method is neglected even in comparison with the TLM simulation of the small 110 μm x 110 μm TLM box.

3.6 Conclusions

Due to the small dimensions a careful investigation of various components based on damascene technology is required. The MM technique allows the accurate parameters computation of multilayered waveguides. For the full-wave analysis of structures in MMICs we combine successfully and efficiently the TLM and MM methods. Using the proposed hybrid TLM-MM approach, the time and memory requirements can been reduced several times, while moreover improving

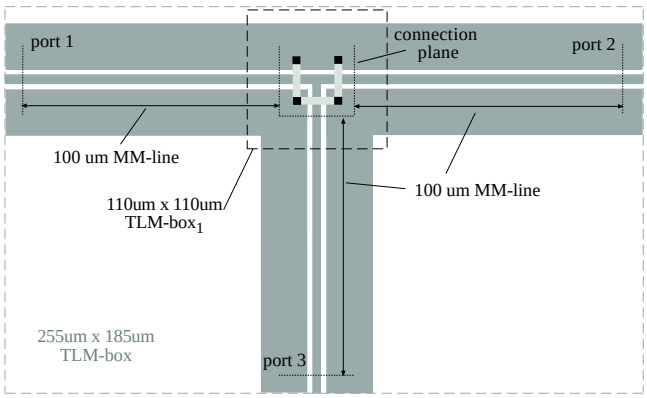
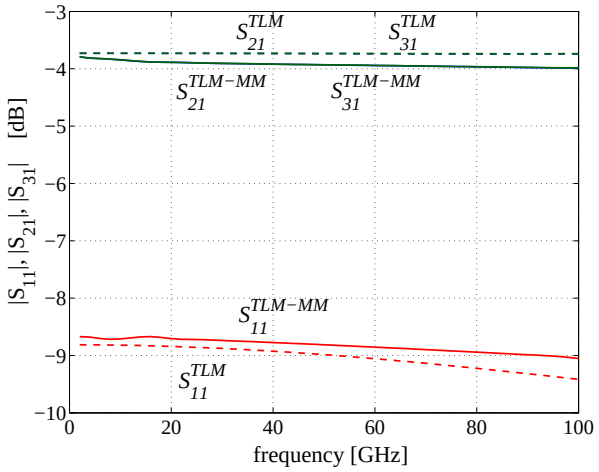
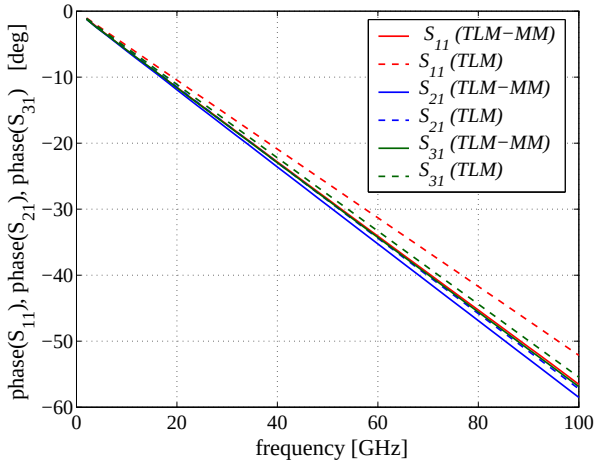


Figure 3.6: Layout of the bridge CPW-junction simulated by means of TLM and the hybrid TLM-MM approach.



(a) Magnitudes



(b) Phases

Figure 3.7: S-parameters for the test CPW structure calculated by means of TLM and hybrid approach.

accuracy. The accuracy of TLM-MM approach and its demand on computation resources are dependent on the number of eigenmodes taken into account and on the space discretization step in a connection cross section. The proposed hybrid technique can be recommended for a effective usage in CAD systems and for predesign of MMICs. The time/frequency hybrid technique based on MM can be applied to FDTD method in a similar manner.

In addition, the hybrid TLM-MM method can be easily extended to facilitate macro-modeling of MMICs, in the sense that through its application the generation of a reduced-order macro-model can be simplified and accelerated.

3.7 Remarks

We make a couple of notices concerning the application of the TLM-MM hybrid technique to TLM-MOR.

1. The proposed TLM-MM approach can be easily involved in the global macro-modeling process. ROMs of complex structures and discontinuities are obtained by TLM-MOR. Transmission lines homogenous in the propagation direction are simulated by means of the MM method. On the basis of numerical results the appropriate macro-models and/or the equivalent circuits can be generated. For example, the simple equivalent wide-band ladder networks for transmission lines are proposed by LUKASHEVICH in [74]. These macro-models (and/or network oriented models) can be easily incorporated in the network-oriented simulators together with macro-models, obtained by TLM-MOR and other MOR-techniques. Thus, the global macro-modeling process can be simplified and additionally accelerated.

2. The above-described kind of the hybridization allows to implicate external 2-D frequency-domain based solvers (i.e. FEM, Method of Lines (MoL) and etc.) for TLM-ROM since the time-domain TLM method does not have its own 2-D eigensolver. These external solvers allow us to extract a macro-model for an arbitrary mode by projection of the transverse electromagnetic field from an external 2-D solver onto the field distributions at the ports computed by TLM-MOR.

   For example, we consider a CPW structure with discontinuities like in Fig. 3.6. In order to gain a reduced-order model we have to excite the structure at the port 1 with the electromagnetic field corresponding to the fundamental CPW-mode. For this purpose we impose the transverse field distribution obtained by a 2-D solver (e.g. by MM or MoL) at the port 1 for the even-mode. After MOR-techniques have been applied, we project the 2-D solution onto the electromagnetic fields (at the ports 1, 2 and 3), which are obtained by MOR, and appropriate input-output voltages and currents for the even-mode can be computed. Thus, a reduced-order model can be generated for the pure fundamental even-mode. Evidently, the voltages and currents can be extracted also for all desired modes, and a generalized multi-mode macro-model can be obtained by interfacing the single-mode models. As a matter of course, the accuracy of this generalized macro-model increases by taking several tens or hundreds modes into account.

Actually, the 2-D distribution of the electromagnetic field can be obtained by TLM also if we simulate an infinitely long transmission line. Then, after some time the evanescent modes are attenuated and we have only the fundamental waveguide mode. The transverse field distribution in the frequency-domain can be obtained by the FOURIER transform. However, if several modes propagate on the transmission line, their separation by means of TLM is quite impossible. In this case the application of external 2-D frequency-domain-based solvers is strongly recommended.

Chapter 4

Krylov Subspace Methods

As it has been mentioned above, in recent years there has been a lot of interest in generating suitable reduced-order models by means of KRYLOV-subspace methods, such as the LANCZOS algorithm and the ARNOLDI process. In this chapter we briefly describe the use of KRYLOV subspace methods for the iterative solution of large linear eigensystems, such as the ones arising in circuit simulation and/or during the numerical simulations of discretized time-invariant linear systems. These iterative techniques, based on the orthogonal and/or oblique projection onto KRYLOV subspaces, allow to reduce a general matrix to be reduced to a condensed form. We also review variants of the ARNOLDI and LANCZOS algorithms for generating basis vectors for KRYLOV subspaces.

Generally, in this chapter we will answer two questions:

- How can a general operator be reduced to a condensed form using KRYLOV subspace methods?
- What must be done so that eigenvalues of a reduced matrix approximate the eigenvalue spectrum of a general operator?

4.1 Krylov Subspaces

Now, let the KRYLOV subspaces be induced by a $N \times N$ dimensional linear operator $\boldsymbol{A} \in \mathcal{H}^{N \times N}$ and a starting vector $|\boldsymbol{p}\rangle \in \mathcal{H}^N$. The *$n$-th* KRYLOV *subspace* is defined as

$$\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n) = \text{span } \{ \ |\boldsymbol{p}\rangle, \ \boldsymbol{A}|\boldsymbol{p}\rangle, \ \boldsymbol{A}^2|\boldsymbol{p}\rangle \ \ \boldsymbol{A}^3|\boldsymbol{p}\rangle, \ \ldots, \ \boldsymbol{A}^{n-1}|\boldsymbol{p}\rangle \ \} \, . \tag{4.1}$$

Let be introduced $n_0 \leq n$ as the largest possible integer n such that the vectors $\boldsymbol{A}^j|\boldsymbol{p}\rangle$, $1 \leq j \leq n-1$, are the linearly independent. In this case we have that

$$\begin{aligned} \text{rank } \{\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n)\} = \max_j \text{ rank } \{\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, j)\} = n_0 \equiv \\ \equiv \text{rank } \{\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n_0)\} \equiv \text{rank } \{\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, \infty)\} \, . \end{aligned} \tag{4.2}$$

Therefore, the KRYLOV sequence $|\boldsymbol{p}\rangle, \ \boldsymbol{A}|\boldsymbol{p}\rangle, \ \boldsymbol{A}^2|\boldsymbol{p}\rangle \ \ldots, \ \boldsymbol{A}^{n-1}|\boldsymbol{p}\rangle$ in (4.1) has the dimension of n if $1 \leq n \leq n_0$, and the dimension of n_0, if $n > n_0$, i.e.

$$\dim \ \{ \ \mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n) \ \} = \begin{cases} n, & 1 \leq n \leq n_0 \, , \\ n_0, & n > n_0 \, . \end{cases} \tag{4.3}$$

So, the *largest possible* KRYLOV *subspace*, induced by $\boldsymbol{A}$ and $|\boldsymbol{p}\rangle$ is clearly defined by $\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n_0)$. The subspace $\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n)$ is termed *exhausted* (see [43]), if $n > n_0$.

In general case, a *block* KRYLOV *sequence* is generated by a $N \times N$ dimensional operator $\boldsymbol{A} \in \mathcal{H}^{N\times N}$ and a $N \times m$-matrix of *m right starting vectors*

$$\boldsymbol{P} = [\ |\boldsymbol{p}_1\rangle,\ |\boldsymbol{p}_2\rangle,\ |\boldsymbol{p}_3\rangle,\ \ldots,\ |\boldsymbol{p}_m\rangle\ ] \in \mathcal{H}^{N\times m} \tag{4.4}$$

as following

$$\boldsymbol{P},\ \boldsymbol{AP},\ \boldsymbol{A}^2\boldsymbol{P},\ \boldsymbol{A}^3\boldsymbol{P},\ \ldots,\ \boldsymbol{A}^{j-1}\boldsymbol{P},\ \ldots\,. \tag{4.5}$$

We note that the operator $\boldsymbol{P}$ in (4.4) can be written in the DIRAC notation as following

$$\boldsymbol{P} = \sum_{i,j=1}^{N,m} |i\rangle\, p_j^i\, \langle j| = \sum_{j=1}^{m} |\boldsymbol{p}_j\rangle\, \langle j|\,, \tag{4.6}$$

however, the form (4.4) is preferred in the present doctoral thesis.

In the sequence (4.5) the j-th block, $\boldsymbol{A}^{j-1}\boldsymbol{P}$, contains the columns, which can be linearly dependent on columns to its left, i.e on columns of

$$\boldsymbol{P},\ \boldsymbol{AP},\ \boldsymbol{A}^2\boldsymbol{P},\ \ldots,\ \boldsymbol{A}^{j-2}\boldsymbol{P}. \tag{4.7}$$

Thus, the linear independence of the columns in (4.5) is generally lost . However, the existence of linear dependent columns in (4.5) does not indicate that the block sequence (4.5) is exhausted, and the vectors $\boldsymbol{A}^{j-1}\boldsymbol{P}$ need to be generated further. This problem of linear dependance can be easily solved, if we scan the columns in (4.5) from left to right and delete each column, which is linearly dependent on its left ones. Then, we obtain the *deflated block* KRYLOV *sequence* as

$$\boldsymbol{P}_1,\ \boldsymbol{AP}_2,\ \boldsymbol{A}^2\boldsymbol{P}_3,\ \boldsymbol{A}^3\boldsymbol{P}_4,\ \ldots,\ \boldsymbol{A}^{j-1}\boldsymbol{P}_j,\ \ldots,\ \boldsymbol{A}^{\hat{n}-1}\boldsymbol{P}_{\hat{n}}\,. \tag{4.8}$$

After the deflation, all columns in (4.8) are linearly independent, and we define the *n-th block* KRYLOV *subspace*, which is induced by the operator $\boldsymbol{A}$ and starting matrix $\boldsymbol{P}$ and spanned by the first n columns of the deflated block sequence, i.e.

$$\mathcal{K}(\boldsymbol{A}, \boldsymbol{P}, n) \ =\ \text{span}\,\{\boldsymbol{P}_1,\ \boldsymbol{AP}_2,\ \boldsymbol{A}^2\boldsymbol{P}_3,\ \boldsymbol{A}^3\boldsymbol{P}_4,\ \ldots,\ \boldsymbol{A}^{\hat{n}-1}\boldsymbol{P}_{\hat{n}}\}\,. \tag{4.9}$$

The number of the linear independent vectors in (4.9) is determined as $n = m_1 + m_2 + \cdots + m_j$, where m_j is the number of the columns in the j-th block $\boldsymbol{A}^{j-1}\boldsymbol{P}_j$ after the deflation.

For MOR-techniques, based on LANCZOS methods, we introduce the block KRYLOV supspace, generated by the transpose operator $\boldsymbol{A}^T \in H^{N\times N}$ and a matrix of *l starting vectors*

$$\boldsymbol{Q} = [\ |\boldsymbol{q}_1\rangle,\ |\boldsymbol{q}_2\rangle,\ |\boldsymbol{q}_3\rangle,\ \ldots,\ |\boldsymbol{q}_l\rangle\ ] \in \mathcal{H}^{N\times l}\,, \tag{4.10}$$

and spanned by the first n columns of the deflated block KRYLOV subspace

$$\boldsymbol{Q}_1,\ \boldsymbol{A}^T\boldsymbol{Q}_2,\ (\boldsymbol{A}^T)^2\boldsymbol{Q}_3,\ (\boldsymbol{A}^T)^3\boldsymbol{Q}_4,\ \ldots,\ (\boldsymbol{A}^T)^{\hat{n}-1}\boldsymbol{Q}_{\hat{n}} \tag{4.11}$$

as following

$$\mathcal{K}(\boldsymbol{A}^T, \boldsymbol{Q}, n) \;=\; \text{span}\,\{\boldsymbol{Q}_1,\; (\boldsymbol{A}^T)\boldsymbol{Q}_2,\; (\boldsymbol{A}^T)^2\boldsymbol{Q}_3,\; \ldots,\; (\boldsymbol{A}^T)^{\hat{n}-1}\boldsymbol{Q}_{\hat{n}}\}\,. \quad (4.12)$$

Here, $n = l_1 + l_2 + \cdots + l_j$, and l_j is the number of the columns in the j-th block $(\boldsymbol{A}^T)^{j-1}\boldsymbol{Q}_j$ after the deleting of linear dependence columns. In the case of $l = 1$ we have the usual n-th KRYLOV subspace induced by $\boldsymbol{A}^T$ and a single starting vector $|\boldsymbol{q}\rangle$

$$\mathcal{K}(\boldsymbol{A}^T, |\boldsymbol{q}\rangle, n) \;=\; \text{span}\,\{\,|\boldsymbol{q}\rangle,\; \boldsymbol{A}^T|\boldsymbol{q}\rangle,\; (\boldsymbol{A}^T)^2|\boldsymbol{q}\rangle,\; \ldots,\; (\boldsymbol{A}^T)^{n-1}|\boldsymbol{q}\rangle\,\}\,. \quad (4.13)$$

As we can see, each application of the general TLM operator ${}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}$ as matrix $\boldsymbol{A}$ for the construction of KRYLOV subspaces in (4.1) and (4.9) is equivalent to one time step in TLM time evolution scheme (2.123). This property makes the use of the TLM time evolution scheme (the TLM kernel) for MOR-techniques based on KRYLOV subspace methods attractive (see Chapter 5).

4.2 Basis Vectors

The columns of block KRYLOV sequences (4.9) and (4.12) and columns of KRYLOV sequences (4.1) and (4.13) induced by the single starting vectors $|\boldsymbol{p}\rangle$ in (4.1) and $|\boldsymbol{q}\rangle$ in (4.13) are linear-independent, and therefore, theoretically can be used in numerical computations as a basis. However, these columns tend to be almost linearly dependent even for moderate values of n. Therefore, their use is not recommended for numerical computational. Instead of that, we generate other basis vectors.

The following,

$$|\boldsymbol{v}_1\rangle, |\boldsymbol{v}_2\rangle, |\boldsymbol{v}_3\rangle, \cdots, |\boldsymbol{v}_n\rangle \in \mathcal{H}^N \quad (4.14)$$

denotes a set of basis vectors for $\mathcal{K}(\boldsymbol{A}, \boldsymbol{P}, n)$, i.e.

$$\text{span}\{\,|\boldsymbol{v}_1\rangle, |\boldsymbol{v}_2\rangle, |\boldsymbol{v}_3\rangle, \cdots, |\boldsymbol{v}_n\rangle\,\} = \mathcal{K}(\boldsymbol{A}, \boldsymbol{P}, n). \quad (4.15)$$

The $N \times n$ matrix

$$\boldsymbol{V}_n := [\,|\boldsymbol{v}_1\rangle, |\boldsymbol{v}_2\rangle, |\boldsymbol{v}_3\rangle, \cdots, |\boldsymbol{v}_n\rangle\,] \in \mathcal{H}^{N\times n} \quad (4.16)$$

is called a *basis matrix* for $\mathcal{K}(\boldsymbol{A}, \boldsymbol{P}, n)$.

Similarly,

$$|\boldsymbol{w}_1\rangle, |\boldsymbol{w}_2\rangle, |\boldsymbol{w}_3\rangle, \cdots, |\boldsymbol{w}_n\rangle \in \mathcal{H}^N \quad (4.17)$$

denotes a set of basis vectors for $\mathcal{K}(\boldsymbol{A}^T, |\boldsymbol{q}\rangle, n)$, i.e.

$$\text{span}\{|\boldsymbol{w}_1\rangle, |\boldsymbol{w}_2\rangle, |\boldsymbol{w}_3\rangle, \cdots, |\boldsymbol{w}_n\rangle\} = \mathcal{K}(\boldsymbol{A}^T, |\boldsymbol{q}\rangle, n). \quad (4.18)$$

And the $N \times n$ matrix

$$\boldsymbol{W}_n := [\,|\boldsymbol{w}_1\rangle, |\boldsymbol{w}_2\rangle, |\boldsymbol{w}_3\rangle, \cdots, |\boldsymbol{w}_n\rangle\,] \in \mathcal{H}^{N\times n} \quad (4.19)$$

is called a *basis matrix* for $\mathcal{K}(\boldsymbol{A}^T, |\boldsymbol{q}\rangle, n)$.

The matrices $\boldsymbol{V}_n$ and $\boldsymbol{W}_n$ can be also written in the DIRAC notation as following

$$\boldsymbol{V}_n = \sum_{i,j=1}^{N,m} |i\rangle\, v_j^i\, \langle j| = \sum_{j=1}^{m} |\boldsymbol{v}_j\rangle\, \langle j|\,, \quad (4.20)$$

$$\boldsymbol{W}_n = \sum_{i,j=1}^{N,m} |i\rangle\, w_j^i\, \langle j| = \sum_{j=1}^{m} |\boldsymbol{w}_j\rangle\, \langle j|\,, \tag{4.21}$$

with $\langle j_{k_1}|j_{k_2}\rangle = \delta_{k_1,k_2}$ and $\langle i_{l_1}|i_{l_2}\rangle = \delta_{l_1,l_2}$. The definitions of the vectors $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{w}_j\rangle$ follow from (4.20) and (4.21). Further we use the notation given through eq. (4.16) and (4.19).

4.3 Arnoldi Basis

The classical ARNOLDI process generates orthonormal basis vectors for the sequence of KRYLOV subspaces $\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n)$ induced by $\boldsymbol{A}$ and a single starting vector $|\boldsymbol{p}\rangle$ as in eq. (4.1), and belongs to *generalized* HESSENBERG *methods* (GHM), which reduce a general matrix to HESSENBERG form. The basis of all such methods is following.

We denote a set of n independent vectors $|\boldsymbol{x}_i\rangle$ as

$$\boldsymbol{X} := \mathrm{span}\{|\boldsymbol{x}_1\rangle, |\boldsymbol{x}_2\rangle, \cdots, |\boldsymbol{x}_i\rangle, \cdots, |\boldsymbol{x}_n\rangle\} \in \mathcal{H}^{N\times n}\,. \tag{4.22}$$

Starting on the arbitrary vector $|\boldsymbol{v}_1\rangle = |\boldsymbol{p}\rangle$ we form a set of modified KRYLOV vectors $|\boldsymbol{v}_2\rangle, |\boldsymbol{v}_3\rangle, \cdots, |\boldsymbol{v}_j\rangle, \cdots, |\boldsymbol{v}_{n+1}\rangle$ obtained by the next sequence

$$\beta_{j+1}|\boldsymbol{v}_{j+1}\rangle = \boldsymbol{A}|\boldsymbol{v}_j\rangle - \sum_{i=1}^{j} h_{ij}|\boldsymbol{v}_i\rangle\,, \tag{4.23}$$

where β_{j+1} are suitable normalizing coefficients, and h_{ij} are determined so that $|\boldsymbol{v}_{j+1}\rangle$ is orthogonal to $|\boldsymbol{x}_1\rangle, |\boldsymbol{x}_2\rangle, \cdots, |\boldsymbol{x}_j\rangle$.

The sequence (4.23) may be assembled into a single operator equation as

$$\begin{aligned}\boldsymbol{A}_{N\times N}\,\boldsymbol{V}_{N\times n} &= \boldsymbol{V}_{N\times n}\,\boldsymbol{H}_{n\times n} + \beta_{n+1}|\boldsymbol{v}_{n+1}\rangle\langle\boldsymbol{e}_n| = \\ &= V_{N\times(n+1)}\,\boldsymbol{H}_{(n+1)\times n}\,,\end{aligned} \tag{4.24}$$

where the matrix $\boldsymbol{H}$ is of the upper HESSENBERG form, i.e.

$$\boldsymbol{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1n} \\ \beta_2 & h_{22} & h_{23} & \cdots & h_{2n} \\ 0 & \beta_3 & h_{33} & \cdots & h_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \beta_n & h_{nn} \end{bmatrix}. \tag{4.25}$$

The matrix $\boldsymbol{H}$ represents a orthogonal projection of the original matrix $\boldsymbol{A}$ onto the KRYLOV subspace $\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n)$. The matrix $\boldsymbol{H}_{n\times n}$ represents a so-called *reduced form* of the original matrix $\boldsymbol{A}_{N\times N}$, because $n << N$.

The dyadic product $|\boldsymbol{v}_{n+1}\rangle\langle\boldsymbol{e}_n|$ presents a rank-one matrix, where $|\boldsymbol{e}_n\rangle = [0, 0, \ldots, 1]^T$.

A sketch of the matrix transformation (4.24) by GHM is presented in Fig. 4.1.

The orthogonality condition can be written in the form

$$\boldsymbol{X}^T_{n\times N}\,\boldsymbol{V}_{N\times n} = \boldsymbol{L}_{n\times n}\,, \tag{4.26}$$

where $\boldsymbol{L}$ is a lower triangular matrix.

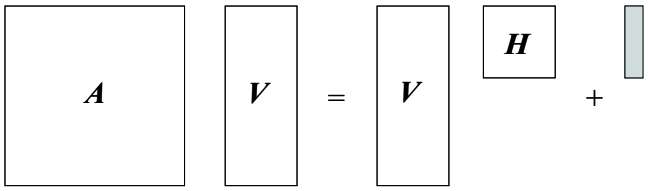


Figure 4.1: Matrix transformation by GHM.

If $\boldsymbol{V}$ is a basis matrix and, eq. (4.23) may be written in the form

$$\boldsymbol{V}^T \boldsymbol{A}\, \boldsymbol{V} = \boldsymbol{H}\,, \tag{4.27}$$

which represents a conformal mapping called *similarity transformation*, and matrices $\boldsymbol{A}$ and $\boldsymbol{H}$ are called *similar* [48]. The determinant of the similarity transformation of the matrix $\boldsymbol{V}^T \boldsymbol{A}\, \boldsymbol{V}$ is equal to the determinant of the original matrix $\boldsymbol{A}$. It is evident that the calculation of eigenvalues of $\boldsymbol{H}$ is significantly simpler than of $\boldsymbol{A}$ if $n << N$.

The GHM may break down for two different reasons [124]:

- A vector $|\boldsymbol{v}_{j+1}\rangle$, $(j < n)$ may be null (i.e. a vector with N components, each of them is 0). In this case we can replace the null $|\boldsymbol{v}_{j+1}\rangle$ by any vector $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ which is orthogonal to vectors $|\boldsymbol{x}_1\rangle, |\boldsymbol{x}_2\rangle, \cdots, |\boldsymbol{x}_j\rangle$.
- The elements of the matrix $\boldsymbol{H}$ are obtained by

$$h_{ij} = \frac{\langle \boldsymbol{x}_i |\, \boldsymbol{A}\, |\boldsymbol{v}_j\rangle}{\langle \boldsymbol{x}_i | \boldsymbol{v}_i\rangle}\,. \tag{4.28}$$

  Therefore, if $\langle \boldsymbol{x}_i | \boldsymbol{v}_i\rangle = 0$ then h_{ij} cannot be determined. It is evident that the vanishing of $\langle \boldsymbol{x}_i | \boldsymbol{v}_i\rangle$ is more serious (and occurs more often) than the vanishing of $|\boldsymbol{v}_i\rangle$ If the vector $|\boldsymbol{x}_i\rangle$ are not given in advance, we may be able to choose the $|\boldsymbol{x}_i\rangle$ that $\langle \boldsymbol{x}_i | \boldsymbol{v}_i\rangle \neq 0$. When the $|\boldsymbol{x}_i\rangle$ are given in advance we have to restart the generation of modified KRYLOV vectors by (4.24) with a different start vector $|\boldsymbol{v}_1\rangle = |\tilde{\boldsymbol{p}}\rangle$. If no $\langle \boldsymbol{x}_i | \boldsymbol{v}_i\rangle$ is zero, then $\boldsymbol{L}$ and $\boldsymbol{V}$ are non-singular.

The choice for $\boldsymbol{X}$ gives reasons for various modifications of GHM. So, one choice for $\boldsymbol{X}$ is the identity matrix. When $\boldsymbol{X} = \boldsymbol{I}$ the result is the *method of* HESSENBERG. Thus, the orthogonality property can be written in the form

$$\boldsymbol{L}_{n\times n} = \boldsymbol{X}^T_{n\times N}\, \boldsymbol{V}_{N\times n} = \boldsymbol{V}_{n\times n}\,, \tag{4.29}$$

where $\boldsymbol{V}_{n\times n}$ is a lower triangular matrix as well.

Another natural choice for the set of vectors $|\boldsymbol{x}_i\rangle$ in GHM is the set of the KRYLOV vectors $|\boldsymbol{v}_i\rangle$ themselves, that leads to the *method of* ARNOLDI. The basis generated by (4.23) satisfies

$$\boldsymbol{V}^T\, \boldsymbol{V} = \boldsymbol{D}\,, \tag{4.30}$$

where $\boldsymbol{D}_{n\times n}$ has diagonal form. We may choose the normalizing factors β_i so that

$$\boldsymbol{V}^T\, \boldsymbol{V} = \boldsymbol{I}\,, \tag{4.31}$$

and the GHM yields an orthonormal basis $\boldsymbol{V}$. The advantage of orthonormalization is that the maximum of each $\|\boldsymbol{v}_i\|_2 = 1$.

However, it is essential that the vectors $|\boldsymbol{v}_i\rangle$ remain strictly orthogonal with respect to the to working accuracy. We consider the vectors

$$\beta_{j+1}|\boldsymbol{v}_{j+1}\rangle = \boldsymbol{A}|\boldsymbol{v}_j\rangle - h_{jj}|\boldsymbol{v}_j\rangle - \ldots - h_{2j}|\boldsymbol{v}_2\rangle - h_{1j}|\boldsymbol{v}_1\rangle\,, \tag{4.32}$$

where the h_{ij} are chosen so that $\beta_{j+1}|\boldsymbol{v}_{j+1}\rangle$ is orthogonal to $|\boldsymbol{v}_i\rangle$. It may be that the components of $\|\beta_{j+1}\boldsymbol{v}_{j+1}\|_2$ are very small compared with $\|\boldsymbol{A}|\boldsymbol{v}_j\rangle\|_2$, and so extensive cancellation occurs. In this case $\beta_{j+1}|\boldsymbol{v}_{j+1}\rangle$ will not even be approximately orthogonal to $|\boldsymbol{v}_i\rangle$, or in other words, the orthogonality is lost. In order to obtain the vector $\beta_{j+1}|\boldsymbol{v}_{j+1}\rangle$ which is truly orthogonal to the $|\boldsymbol{v}_i\rangle$ we have to perform re-orthogonalization for $\beta_{j+1}|\boldsymbol{v}_{j+1}\rangle$. This re-orthogonalization may be based on the GRAM-SCHMIDT orthogonalization process (see APPENDIX B).

The ARNOLDI method can be essentially viewed as a modified GRAM-SCHMIDT process for building an orthogonal basis for the KRYLOV subspace $\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n)$. Other orthogonalization algorithms can be used in addition to the ones already mentioned. One of the most reliable orthogonalization techniques is the HOUSEHOLDER *algorithm* [48, 124], which has been implemented for the ARNOLDI procedure by WALKER in [121]. Even though the HOUSEHOLDER algorithm is assumed to be numerically more reliable than GRAM-SCHMIDT process, it is more expensive, having the some storage requirements as modified GRAM-SCHMIDT but about twice as many operations.

Now, let us consider the pseudocode of the basic ARNOLDI procedure.

Algorithm 4.3.1. The ARNOLDI algorithm.

(1) choose $|\boldsymbol{v}_1\rangle = \dfrac{|\boldsymbol{p}\rangle}{\|\boldsymbol{p}\|_2}$
(2) **for** $j = 1, 2, \ldots, n$ **do**
(3) $\quad|\tilde{\boldsymbol{v}}\rangle = \boldsymbol{A}|\boldsymbol{v}_j\rangle$
(4) $\quad$**for** $i = 1, 2, \ldots, j$ **do**
(5) $\quad\quad h_{ij} = \langle\tilde{\boldsymbol{v}}|\boldsymbol{v}_i\rangle$
(6) $\quad\quad|\tilde{\boldsymbol{v}}\rangle = |\tilde{\boldsymbol{v}}\rangle - h_{ij}|\boldsymbol{v}_i\rangle$
(7) $\quad$**end for**
(8) $\quad\beta_{j+1} = \|\tilde{\boldsymbol{v}}\|_2$
(9) $\quad$**if** $\beta_{j+1} = 0$ **then stop**
(10) $\quad|\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\beta_{j+1}}$
(11) **end for**

As it noted earlier, the algorithm breaks down when the norm of $\|\tilde{\boldsymbol{v}}\|_2$ on line (8) vanishes at the step j. It happens if, and only if, the starting vector $|\boldsymbol{p}\rangle$ is a linear combination of j basis vectors $|\boldsymbol{v}_i\rangle$.

The eigenvalue spectrum of the matrix $\boldsymbol{A}$ is approximated by eigenvalues of the HESSENBERG matrix $\boldsymbol{H}$, which are known as RITZ *values*. The eigenvector associated with a RITZ value $\lambda_i^{(n)}$ is called the RITZ *vector* and defined by

$$|u_i^{(n)}\rangle = \lambda_i^{(n)}\,|\boldsymbol{u}_i^{(n)}\rangle\,, \tag{4.33}$$

where $|\boldsymbol{u}_i^{(n)}\rangle$ is an eigenvector associated with the i-th RITZ value $\lambda_i^{(n)}$, obtained at the n-th step. The quality of the approximation will usually improve as n increases.

We have the residual vector as

$$|\boldsymbol{r}_{res}\rangle = (\boldsymbol{A} - \lambda_i^{(n)}\boldsymbol{I})|u_i^{(n)}\rangle = \beta_{n+1}\langle \boldsymbol{e}_n|\boldsymbol{u}_i^{(n)}\rangle\,|\boldsymbol{v}_{n+1}\rangle \tag{4.34}$$

and, therefore,

$$\|\boldsymbol{r}_{res}\|_2 = \|(\boldsymbol{A} - \lambda_i^{(n)}\boldsymbol{I})|u_i^{(n)}\rangle\|_2 = \beta_{n+1}\,|\langle \boldsymbol{e}_n|\boldsymbol{u}_i^{(n)}\rangle|\,. \tag{4.35}$$

Thus, the norm of the residual vector is equal to the absolute value of the last component of the eigenvector $|\boldsymbol{u}_i^{(n)}\rangle$ multiplied by the coefficient β_{n+1}. We note that the residual norms do not always indicate the actual errors, but can be used as stopping criteria if the convergence for needed eigenvalues is achieved.

One improvement would be to perform re-orthogonalization if necessary. Whenever the final vector $|\tilde{\boldsymbol{v}}\rangle$ obtained at the end of the second loop is computed, its norm is compared with the norm of initial $\boldsymbol{A}|\boldsymbol{v}_j\rangle$. If the vector reduction is below a certain threshold, the second orthogonalization is made. This change of the vector norm indicates that cancellation has occurred.

In order to reliably determine multiple and/or clustered eigenvalues, the block methods are used. In many cases the cost of computing few matrix-vector products is commensurate with that of one matrix-vector product; therefore, the computational overhead is neglectful. Although unblocked methods coupled with deflation strategy [7] may be used to compute multiple and/or clustered eigenvalues, they may prove inefficient for some eigenvalue problems because of computational costs. In this case (without deflation) the block ARNOLDI process forms the orthogonal basis $\boldsymbol{V}_{N\times nm}$ for KRYLOV subspace $\mathcal{K}(\boldsymbol{A}, \boldsymbol{P}, n)$, where

$$\boldsymbol{V}_{N\times \hat{n}m} = [\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_3, \ldots, \boldsymbol{V}_{\hat{n}}]\,. \tag{4.36}$$

For simplification we suppose all columns of $\boldsymbol{V}_{N\times \hat{n}m}$ to be linearly independent.

In this case the similarity transformation is given by the next relation

$$\boldsymbol{A}_{N\times N}\,\boldsymbol{V}_{N\times \hat{n}m} = \boldsymbol{V}_{N\times \hat{n}m}\,\boldsymbol{H}_{\hat{n}m\times \hat{n}m} + \boldsymbol{F}_{\hat{n}m\times \hat{n}m}\boldsymbol{E}_{\hat{n}m\times \hat{n}m}\,, \tag{4.37}$$

where m is the number of the vectors in the block $\boldsymbol{V}_j$; $\hat{n}$ denotes the total number of blocks; $n = \hat{n}m$ is the order of KRYLOV subspace (i.e. total number of basis vectors); $\boldsymbol{H}_{\hat{n}m\times \hat{n}m}$ is a band upper HESSENBERG matrix of order $m\hat{n}$ (an upper tridiagonal matrix with m subdiagonals).

The block ARNOLDI algorithm with the deflation can be found in [43]. For more details of ARNOLDI block algorithms the reader is referred to [103].

4.4 Lanczos Basis

The basic idea of the LANCZOS procedure is to replace the eigenvalue problem for a given general operator $\boldsymbol{A}$ by eigenvalue computations on one or more of the simpler tridiagonal LANCZOS matrices, which are assumed to be computationally optimal matrices. The classical LANCZOS method generates two sequences of vectors, which form a biorthogonal basis. The approximating tridiagonal LANCZOS matrix is the matrix representation of the biorthogonal projection of

the original matrix $\boldsymbol{A}$ onto KRYLOV subspaces induced by $\boldsymbol{A}$ and $\boldsymbol{A}^T$ and start vectors $|\boldsymbol{p}\rangle$ and $|\boldsymbol{q}\rangle$ respectively, i.e. onto the $\mathcal{K}(\boldsymbol{A}, |\boldsymbol{p}\rangle, n)$ and $\mathcal{K}(\boldsymbol{A}^T, |\boldsymbol{q}\rangle, n)$ subspaces respectively.

Let us point out that there are two equivalent formulations of the LANCZOS process for complex matrices - using either $\boldsymbol{A}^T$ or $\boldsymbol{A}^\dagger$. In this doctoral thesis the formulation with $\boldsymbol{A}^T$ is chosen for two reasons. First, it allows to avoid complex conjugated scalars in some of the recurrence relations; and second, the general TLM operator is presented by a real nonsymmetric matrix.

In order to derive the LANCZOS recurrences we return to GHM. There is one other possibility to choose the $|\boldsymbol{x}_i\rangle$ in GHM. We can derive the basis $\boldsymbol{W}$ using the transpose $\boldsymbol{A}^T$ in the same way as $|\boldsymbol{v}_i\rangle$ are obtained by $\boldsymbol{A}$. The relevant sequences are

$$\rho_{j+1}|\boldsymbol{v}_{j+1}\rangle = \boldsymbol{A}|\boldsymbol{v}_j\rangle - \sum_{i=1}^{j} h_{ij}|\boldsymbol{v}_i\rangle\,, \tag{4.38}$$

$$\eta_{j+1}|\boldsymbol{w}_{j+1}\rangle = \boldsymbol{A}^T|\boldsymbol{w}_j\rangle - \sum_{i=1}^{j} \hat{h}_{ij}|\boldsymbol{w}_i\rangle\,, \tag{4.39}$$

where h_{ij} and $\hat{h}_{ij}$ are chosen so that $|\boldsymbol{v}_{j+1}\rangle$ is orthogonal to $|\boldsymbol{w}_i\rangle$ and $|\boldsymbol{w}_{j+1}\rangle$ is orthogonal to $|\boldsymbol{v}_i\rangle$ respectively.

We may write these generation sequences in corresponding matrix forms as following

$$\boldsymbol{A}\,\boldsymbol{V} = \boldsymbol{V}\,\boldsymbol{H}\,, \tag{4.40}$$

$$\boldsymbol{A}^T\,\boldsymbol{W} = \boldsymbol{W}\,\hat{\boldsymbol{H}}\,, \tag{4.41}$$

where $\boldsymbol{H}$ and $\hat{\boldsymbol{H}}$ are of upper HESSENBERG form.

The sets $\boldsymbol{V}$ and $\boldsymbol{W}$ are connected with each other through

$$\boldsymbol{W}^T\,\boldsymbol{V} = \boldsymbol{L}\,, \tag{4.42}$$

$$\boldsymbol{V}^T\,\boldsymbol{W} = \hat{\boldsymbol{L}}\,, \tag{4.43}$$

where $\boldsymbol{L}$ and $\hat{\boldsymbol{L}}$ are lower triangular.

Following eq. (4.42) and eq. (4.43) is

$$\boldsymbol{L} = \hat{\boldsymbol{L}}^T\,, \tag{4.44}$$

and hence both have to be diagonal. We denote them as a diagonal matrix $\boldsymbol{D}$.

From eq. (4.41) we have

$$\boldsymbol{A} = (\boldsymbol{W}^{-1})^T\,\hat{\boldsymbol{H}}\,\boldsymbol{W}^T\,, \tag{4.45}$$

and from eq. (4.40) and eq. (4.45) we obtain

$$\boldsymbol{H} = \boldsymbol{V}^{-1}\,\boldsymbol{A}\,\boldsymbol{V} = \boldsymbol{D}^{-1}\,\hat{\boldsymbol{H}}^T\,\boldsymbol{D}. \tag{4.46}$$

The matrix $\boldsymbol{H}$ is upper HESSENBERG and the matrix $\boldsymbol{D}^{-1}\,\hat{\boldsymbol{H}}^T\,\boldsymbol{D}$ is upper HESSENBERG matrix. Therefore, both are tridiagonal, and in eq. (4.25) and eq. (4.38) we have

$$h_{i,j} = \hat{h}_{i,j} = 0\,, \quad \text{for } i = 1, 2, \ldots, n-2\,, j = i+2, \ldots, n\,. \tag{4.47}$$

As, as follows, equations in (4.38) are reduced to three-term sequences:

$$\rho_{j+1}|\boldsymbol{v}_{j+1}\rangle = \boldsymbol{A}|\boldsymbol{v}_j\rangle - \alpha_j|\boldsymbol{v}_j\rangle - \beta_j|\boldsymbol{v}_{j-1}\rangle\,, \tag{4.48}$$

$$\eta_{j+1}|\boldsymbol{w}_{j+1}\rangle = \boldsymbol{A}^T|\boldsymbol{w}_j\rangle - \alpha_j|\boldsymbol{w}_j\rangle - \gamma_j|\boldsymbol{w}_{j-1}\rangle\,, \tag{4.49}$$

where the coefficients α_j, β_j and γ_j are given by

$$\alpha_j = \frac{\langle \boldsymbol{w}_j|\boldsymbol{A}|\boldsymbol{v}_j\rangle}{\langle \boldsymbol{w}_j|\boldsymbol{v}_j\rangle} = \frac{\langle \boldsymbol{v}_j|\boldsymbol{A}^T|\boldsymbol{w}_j\rangle}{\langle \boldsymbol{v}_j|\boldsymbol{w}_j\rangle}\,, \tag{4.50}$$

$$\begin{aligned}\beta_j = \frac{\langle \boldsymbol{w}_{j-1}|\boldsymbol{A}|\boldsymbol{v}_j\rangle}{\langle \boldsymbol{w}_{j-1}|\boldsymbol{v}_{j-1}\rangle} = \frac{\langle \boldsymbol{v}_j|\boldsymbol{A}^T|\boldsymbol{w}_{j-1}\rangle}{\langle \boldsymbol{w}_{j-1}|\boldsymbol{v}_{j-1}\rangle} = \\ = \frac{\langle \boldsymbol{v}_j|\ (\eta_j|\boldsymbol{w}_j\rangle + \alpha_{j-1}|\boldsymbol{w}_{j-1}\rangle + \gamma_{j-1}|\boldsymbol{w}_{j-2}\rangle)}{\langle \boldsymbol{w}_{j-1}|\boldsymbol{v}_{j-1}\rangle} = \\ = \eta_j \frac{\langle \boldsymbol{v}_j|\boldsymbol{w}_j\rangle}{\langle \boldsymbol{w}_{j-1}|\boldsymbol{v}_{j-1}\rangle}\,.\end{aligned} \tag{4.51}$$

Similarly we obtain

$$\gamma_j = \rho_j \frac{\langle \boldsymbol{w}_j|\boldsymbol{v}_j\rangle}{\langle \boldsymbol{v}_{j-1}|\boldsymbol{w}_{j-1}\rangle}\,. \tag{4.52}$$

Thus, if $\boldsymbol{A}|\boldsymbol{v}_j\rangle$ is orthogonalized with respect to $|\boldsymbol{w}_j\rangle$ and $|\boldsymbol{w}_{j-1}\rangle$ it is automatically orthogonal with respect to all earlier $|\boldsymbol{w}_i\rangle$, and similarly, if $\boldsymbol{A}|\boldsymbol{w}_j\rangle$ is orthogonalized with respect to $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{v}_{j-1}\rangle$ it is automatically orthogonal with respect to all earlier $|\boldsymbol{v}_i\rangle$.

The LANCZOS algorithm may break down for two different reasons [124]:

- One or both of the vectors $|\boldsymbol{v}_{j+1}\rangle$ and $|\boldsymbol{w}_{j+1}\rangle$ may be null. In this case we take new vectors $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ and $|\tilde{\boldsymbol{w}}_{j+1}\rangle$ so that $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ is orthogonal to all $|\boldsymbol{w}_1\rangle, |\boldsymbol{w}_2\rangle, \ldots, |\boldsymbol{w}_j\rangle$ and $|\tilde{\boldsymbol{w}}_{j+1}\rangle$ is orthogonal to all previous $|\boldsymbol{v}_1\rangle, |\boldsymbol{v}_2\rangle, \ldots, |\boldsymbol{v}_j\rangle$, after that we can proceed the algorithm.
- On the other hand, if $\langle \boldsymbol{v}_j|\boldsymbol{w}_j\rangle = 0$ at one step the process breaks down completely, and we have to start the algorithm again with a new start vector (eventually with new both vectors), i.e. $|\boldsymbol{v}_1\rangle = |\tilde{\boldsymbol{p}}\rangle$ and/or $|\boldsymbol{w}_1\rangle = |\tilde{\boldsymbol{q}}\rangle$.

Algorithm 4.4.1. The nonsymmetric LANCZOS algorithm.

(1) choose $|\boldsymbol{v}_1\rangle = \dfrac{|\boldsymbol{p}\rangle}{\|\boldsymbol{p}\|_2}$ and $|\boldsymbol{w}_1\rangle = \dfrac{|\boldsymbol{q}\rangle}{\|\boldsymbol{q}\|_2}$
(2) set $|\boldsymbol{v}_0\rangle = \boldsymbol{0}$, $|\boldsymbol{w}_0\rangle = \boldsymbol{0}$ and $\rho_1 = \eta_1 = \delta_0 = 1$
(3) **for** $j = 1, 2, \ldots, n$ **do**
(4) $\quad \delta_j = \langle \boldsymbol{w}_j|\boldsymbol{v}_j\rangle$
(5) $\quad$ **if** $\delta_j = 0$ **then stop**
(6) $\quad \alpha_j = \dfrac{\langle \boldsymbol{w}_j|\boldsymbol{A}|\boldsymbol{v}_j\rangle}{\delta_j}$
(7) $\quad \beta_j = \eta_j\delta_j/\delta_{j-1}$
(8) $\quad \gamma_j = \rho_j\delta_j/\delta_{j-1}$
(9) $\quad |\tilde{\boldsymbol{v}}_{j+1}\rangle = \boldsymbol{A}\,|\boldsymbol{v}_j\rangle - \alpha_j|\boldsymbol{v}_j\rangle - \beta_j|\boldsymbol{v}_{j-1}\rangle$

(10) $|\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{A}^T |\boldsymbol{w}_j\rangle - \alpha_j |\boldsymbol{w}_j\rangle - \gamma_j |\boldsymbol{w}_{j-1}\rangle$
(11) $\rho_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(12) $\eta_{j+1} = \|\tilde{\boldsymbol{w}}_{j+1}\|_2$
(13) **if** $\rho_{j+1} = 0$ or $\eta_{j+1} = 0$ **then stop**
(14) $|\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\rho_{j+1}}$
(15) $|\boldsymbol{w}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{w}}_{j+1}\rangle}{\eta_{j+1}}$
(16) **end for**

The matrix $\boldsymbol{A}$ is represented by a $n \times n$ non-HERMITIAN tridiagonal matrix $\boldsymbol{T}$

$$\boldsymbol{W}^T \boldsymbol{A} \boldsymbol{V} = \boldsymbol{D}\,\boldsymbol{T}\,, \tag{4.53}$$

where

$$\boldsymbol{T} = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \rho_2 & \alpha_2 & \beta_3 & & \\ & \rho_3 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_n \\ & & & \rho_n & \alpha_n \end{bmatrix}, \tag{4.54}$$

and the matrix

$$\boldsymbol{D} = \mathrm{diag}\,(\delta_1, \delta_2, \ldots, \delta_j, \ldots \delta_n) \tag{4.55}$$

is determined by biorthogonality property

$$\boldsymbol{W}^T \boldsymbol{V} = \boldsymbol{D}\,. \tag{4.56}$$

In other words, the constructed LANCZOS vectors are biorthogonal. That is

$$\langle \boldsymbol{w}_i | \boldsymbol{v}_j \rangle = \begin{cases} \delta_j, & i = j\,, \\ 0, & i \neq j\,. \end{cases} \tag{4.57}$$

Strictly speaking, the governing relations of the LANCZOS factorization are given in a matrix form by

$$\boldsymbol{A}_{N\times N}\,\boldsymbol{V}_{N\times n} = \boldsymbol{V}_{N\times n}\,\boldsymbol{T}_{n\times n} + \beta_{n+1} |\boldsymbol{v}_{n+1}\rangle \langle \boldsymbol{e}_n|\,, \tag{4.58}$$

$$\boldsymbol{A}^T_{N\times N}\,\boldsymbol{W}_{N\times n} = \boldsymbol{W}_{N\times n}\,\boldsymbol{T}^T_{n\times n} + \gamma_{n+1} |\boldsymbol{w}_{n+1}\rangle \langle \boldsymbol{e}_n|\,. \tag{4.59}$$

Using eq. (4.46) we can write eq. (4.58), (4.59) in a slightly different form

$$\boldsymbol{A}_{N\times N}\,\boldsymbol{V}_{N\times n} = \boldsymbol{V}_{N\times (n+1)}\,\boldsymbol{T}_{(n+1)\times n}\,, \tag{4.60}$$

$$\boldsymbol{A}^T_{N\times N}\,\boldsymbol{W}_{N\times n} = \boldsymbol{W}_{N\times (n+1)}\,\boldsymbol{D}^{-1}_{(n+1)\times (n+1)}\,\boldsymbol{T}_{(n+1)\times n}\,\boldsymbol{D}_{n\times n}. \tag{4.61}$$

A sketch of the similarity transformation by the LANCZOS algorithm is presented in Fig. 4.2. The action of $\boldsymbol{A}$ on $\boldsymbol{V}$ gives $\boldsymbol{V}\boldsymbol{T}$ plus a rank-one matrix $\beta_{n+1}|\boldsymbol{v}_{n+1}\rangle\langle\boldsymbol{e}_n|$, and the action of $\boldsymbol{A}^T$ on $\boldsymbol{W}$ gives $\boldsymbol{W}\boldsymbol{T}^T$ plus a rank-one matrix $\gamma_{n+1}|\boldsymbol{w}_{n+1}\rangle\langle\boldsymbol{e}_n|$.

We note, that the strict bi-orthogonality of the two sequences $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{w}_j\rangle$ is usually soon lost. In order to avoid this effect the rebiorthogonalization can be performed at each stage. However, it cannot be too strongly emphasized

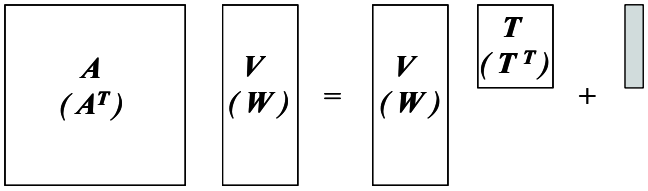


Figure 4.2: Similarity transformation by the LANCZOS algorithm.

that rebiorthogonalization does not produce the vectors which would have been obtained by more accurate computation and that this is no deleterious effect [124]. For this purpose the two-sided modified GRAM-SCHMIDT process may be utilized (see APPENDIX B).

The eigenvalues of $\boldsymbol{A}$ are approximated by the RITZ values $\lambda_i^{(n)}$, which are eigensolutions of $\boldsymbol{T}$ at the step n, i.e.

$$\boldsymbol{T}\,|\boldsymbol{u}_i^{(n)}\rangle = \lambda_i^{(n)}\,|\boldsymbol{u}_i^{(n)}\rangle\,, \tag{4.62}$$

$$\langle \boldsymbol{s}_i^{(n)}|\,\boldsymbol{T} = \langle \boldsymbol{s}_i^{(n)}|\,\lambda_i^{(n)}\,. \tag{4.63}$$

The right and left RITZ vectors, corresponding to the RITZ values $\lambda_i^{(n)}$ are given by

$$|u_i^{(n)}\rangle = \boldsymbol{V}\,|\boldsymbol{u}_i^{(n)}\rangle\,, \tag{4.64}$$

$$|u_i^{(n)}\rangle = \boldsymbol{W}\,|\boldsymbol{s}_i^{(n)}\rangle\,. \tag{4.65}$$

The convergence of the RITZ values and RITZ vectors to eigenvalues and eigenvectors of $\boldsymbol{A}$ can be evaluated by comparing the norms of the residuals

$$|\boldsymbol{r}_{res}\rangle = (\boldsymbol{A} - \lambda_i^{(n)}\boldsymbol{I})\,|u_i^{(n)}\rangle = \rho_{n+1}\{\langle \boldsymbol{e}_n\,|\,\boldsymbol{u}_i^{(n)}\rangle\}\;|\boldsymbol{v}_{n+1}\rangle\,, \tag{4.66}$$

$$\langle \boldsymbol{p}_{res}| = \langle u_i^{(n)}|\,(\boldsymbol{A} - \lambda_i^{(n)}\boldsymbol{I}) = \eta_{n+1}\langle \boldsymbol{w}_{n+1}|\;\{\langle \boldsymbol{s}_i^{(n)}\,|\,\boldsymbol{e}_n\rangle\}\,, \tag{4.67}$$

and, therefore,

$$\|\boldsymbol{r}_{res}\|_2 = \rho_{n+1}\;|\langle \boldsymbol{e}_n\,|\,\boldsymbol{u}_i^{(n)}\rangle|\,, \tag{4.68}$$

$$\|\boldsymbol{p}_{res}\|_2 = \eta_{n+1}\;|\langle \boldsymbol{s}_i^{(n)}\,|\,\boldsymbol{e}_n\rangle|\,. \tag{4.69}$$

In order to determine multiple and/or clustered eigenvalues the LANCZOS block and band methods can be introduced as well. The use of blocks of starting vectors is also beneficial whenever computing matrix-matrix products $\boldsymbol{AV}$ and $\boldsymbol{A}^T\boldsymbol{W}$, and is cheaper than sequentially computing matrix-vector products $\boldsymbol{A}|\boldsymbol{v}\rangle$ and $\boldsymbol{A}^T|\boldsymbol{w}\rangle$ for all the columns of $\boldsymbol{V}$ and $\boldsymbol{W}$. In the band variant in contrast to the block methods, the sizes of the left and the right blocks of LANCZOS vectors need not necessarily be the same. In particular, these band methods can be applied to MOR of linear dynamical systems, where the right and left starting blocks have different sizes.

With respect to TLM the utilization of block and/or band methods is very expensive in comparison to the use of single starting vectors (see Chapter 5). Therefore, the block and band algorithms are not described in detail here. The reader is referred to the works of FREUND [43]–[46].

Investigations and suggestions for efficient implementations of the nonsymmetric LANCZOS algorithm are presented in [33]. It has been shown that it is not necessary to rebiorthogonalize the LANCZOS vectors at each step to approximate the behavior of the algorithm in exact arithmetic.

4.4.1 The Hermitian (Symmetric) Lanczos Process

If $\boldsymbol{A}$ is HERMITIAN, or in the real symmetric case , and we take $|\boldsymbol{v}_1\rangle = |\boldsymbol{w}_1\rangle$, then the two sequences of vectors in eq. (4.48) and eq. (4.49) are identical. Because the second sequence in eq. (4.49) has not be generated, the HERMITIAN LANCZOS version is twice as cheap than its nonsymmetric variant. Finally, the symmetric algorithm is very attractive for dealing with implicit matrices, when the transpose matrix is not known and/or difficult to obtain.

Using the orthogonal basis $\boldsymbol{V}$ the operator $\boldsymbol{A}$ is presented by a real tridiagonal matrix

$$\boldsymbol{T} = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \beta_3 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_n \\ & & & \beta_n & \alpha_n \end{bmatrix} . \tag{4.70}$$

The pseudocode of the symmetric version is presented below. The algorithm carries out the orthogonal projection of the vector $\boldsymbol{A}|\boldsymbol{v}_j\rangle$ onto the two most recently generated LANCZOS vectors $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{v}_{j-1}\rangle$.

Algorithm 4.4.2. The symmetric LANCZOS algorithm.

(1) choose $|\boldsymbol{v}_1\rangle = \dfrac{|\boldsymbol{p}\rangle}{\|\boldsymbol{p}\|_2}$
(2) set $|\boldsymbol{v}_0\rangle = \boldsymbol{0}$ and $\beta_1 = 1$
(3) **for** $j = 1, 2, \ldots, n$ **do**
(4) $\quad \alpha_j = \langle \boldsymbol{v}_j | \boldsymbol{A} | \boldsymbol{v}_j \rangle$
(5) $\quad |\tilde{\boldsymbol{v}}_{j+1}\rangle = \boldsymbol{A}\,|\boldsymbol{v}_j\rangle - \alpha_j |\boldsymbol{v}_j\rangle - \beta_j |\boldsymbol{v}_{j-1}\rangle$
(6) $\quad \beta_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(7) $\quad$ **if** $\beta_{j+1} = 0$ **then stop**
(8) $\quad |\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\beta_{j+1}}$
(9) **end for**

For each j the next LANCZOS vector $|\boldsymbol{v}_{j+1}\rangle$ is determined by orthogonalizing the vector $\boldsymbol{A}\,|\boldsymbol{v}_j\rangle$ with respect to $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{v}_{j-1}\rangle$. The scalar coefficients α_j and β_{j+1} obtained in these orthogonalizations define the corresponding LANCZOS matrix $\boldsymbol{T}$.

Writing the similarity transformation in matrix form we obtain

$$\boldsymbol{A}_{N\times N}\,\boldsymbol{V}_{N\times n} \;=\; \boldsymbol{V}_{N\times n}\,\boldsymbol{T}_{n\times n} + \beta_{n+1} |\boldsymbol{v}_{n+1}\rangle\langle \boldsymbol{e}_n| \,. \tag{4.71}$$

The symmetric LANCZOS procedure transforms a general real HERMITIAN (and/or symmetric) matrix eigenvalues problem into a simple tridiagonal real

HERMITIAN (and/or symmetric) problem. Real symmetric tridiagonal matrices can be viewed as computationally optimal matrices. They have minimal storage requirements, as do the associated algorithms for eigenvalue computations and for solving trigiagonal systems of equations.

The use of the symmetric single-vector LANCZOS procedure for large matrix eigenvalue computations is thoroughly discussed in the book of CULLUM [32].

4.4.2 The Two-Term Lanczos Process

The above described LANCZOS algorithms utilize so-called three-term LANCZOS recursions. In this subsection we consider an alternative approach for constructing LANCZOS vectors. The basic idea is to exchange these three-term recurrences for coupled two-term recurrences by means of an additional suitable set of basis vectors for the underlying KRYLOV subspaces. In some cases the two-term procedure has better numerical properties than the original implementation based on three-term recurrences, e.g. for the quasi-minimal residual (QMR) algorithm. In addition, the two-term LANCZOS recurrences are effectively used with the classical biconjugate gradient (BCG) for solving systems of linear equations [46].

In order to obtain the coupled two-term version, we define a second set of basis vectors by P-Q sequence

$$\boldsymbol{P}_n : \quad = [\ |\boldsymbol{p}_1\rangle, |\boldsymbol{p}_2\rangle, |\boldsymbol{p}_3\rangle, \cdots, |\boldsymbol{p}_n\rangle\ ] \in \mathcal{H}^{N\times n}\,, \tag{4.72}$$

$$\boldsymbol{Q}_n : \quad = [|\boldsymbol{q}_1\rangle, |\boldsymbol{q}_2\rangle, |\boldsymbol{q}_3\rangle, \cdots, |\boldsymbol{q}_n\rangle] \in \mathcal{H}^{N\times n}\,, \tag{4.73}$$

and we refer to the LANCZOS vectors $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{w}_j\rangle$ as the V-W sequence.

We compute the second basis vectors by following recurrences

$$|\boldsymbol{p}_j\rangle \;=\; |\boldsymbol{v}_j\rangle - \sum_{i=1}^{j-1} u_{ij}|\boldsymbol{p}_i\rangle\,, \tag{4.74}$$

$$|\boldsymbol{q}_j\rangle \;=\; |\boldsymbol{w}_j\rangle - \sum_{i=1}^{j-1} u_{ij}\frac{\chi_j}{\chi_i}|\boldsymbol{q}_i\rangle\,, \tag{4.75}$$

with suitable coefficients $u_{ij} \in \mathcal{C}$. Similarly we can write that

$$\rho_{j+1}|\boldsymbol{v}_{j+1}\rangle \;=\; \boldsymbol{A}|\boldsymbol{p}_j\rangle - \sum_{i=1}^{j} l_{ij}|\boldsymbol{v}_i\rangle\,, \tag{4.76}$$

$$\eta_{j+1}|\boldsymbol{w}_{j+1}\rangle \;=\; \boldsymbol{A}^T|\boldsymbol{q}_j\rangle - \sum_{i=1}^{j} l_{ij}\frac{\chi_j}{\chi_i}|\boldsymbol{w}_i\rangle\,, \tag{4.77}$$

with suitable coefficients $l_{ij} \in \mathcal{C}$. Thus, the recurrences (4.74)–(4.77) are coupled for generating the P-Q and V-W sequences.

The recurrences (4.74)-(4.75) and (4.76)-(4.77) can be written compactly in

matrix form

$$\boldsymbol{V}_{N\times n} = \boldsymbol{P}_{N\times n}\boldsymbol{U}_{n\times n}\,, \tag{4.78}$$

$$\boldsymbol{A}_{N\times N}\boldsymbol{P}_{N\times n} = \boldsymbol{V}_{N\times(n+1)}\boldsymbol{L}_{(n+1)\times n}\,, \tag{4.79}$$

$$\boldsymbol{W}_{N\times n} = \boldsymbol{Q}_{N\times n}\boldsymbol{\chi}^{-1}_{n\times n}\boldsymbol{U}_{n\times n}\boldsymbol{\chi}_{n\times n}\,, \tag{4.80}$$

$$\boldsymbol{A}^{T}_{N\times N}\boldsymbol{Q}_{N\times n} = \boldsymbol{W}_{N\times(n+1)}\boldsymbol{\chi}^{-1}_{(n+1)\times(n+1)}\boldsymbol{L}_{(n+1)\times n}\boldsymbol{\chi}_{n\times n}\,, \tag{4.81}$$

where $\boldsymbol{\chi} = \text{diag}\,(\chi_1, \chi_2, \ldots, \chi_j, \ldots, \chi_n)$ is a scaling diagonal matrix.

Here $\boldsymbol{U}_n$ is an upper triangular matrix and $\boldsymbol{L}_n$ is an upper HESSENBERG matrix, which are given by

$$\boldsymbol{U}_n = \begin{bmatrix} 1 & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & 1 & u_{23} & \cdots & u_{2n} \\ 0 & 0 & 1 & \cdots & u_{3n} \\ \cdots & \cdots & \cdots & \cdots & u_{n-1,n} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}, \tag{4.82}$$

and

$$\boldsymbol{L}_n = \begin{bmatrix} l_{11} & l_{12} & h_{13} & \cdots & l_{1n} \\ \rho_2 & l_{22} & l_{23} & \cdots & l_{2n} \\ 0 & \rho_3 & l_{33} & \cdots & l_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \rho_n & l_{nn} \\ 0 & 0 & \cdots & 0 & \rho_{n+1} \end{bmatrix}. \tag{4.83}$$

Further we denote the matrices by their smaller size, e.g $\boldsymbol{V}_n := \boldsymbol{V}_{N\times n}$ and $\boldsymbol{L}_n := \boldsymbol{L}_{(n+1)\times n}$. (For convenience the large index N is skipped.)

From eq. (4.78)-(4.79) we obtain following relation by eliminating $\boldsymbol{P}_n$

$$\boldsymbol{A}\boldsymbol{V}_n = \boldsymbol{V}_{n+1}\boldsymbol{L}_n\boldsymbol{U}_n \tag{4.84}$$

and by comparison of eq. (4.40) and eq. (4.84) it follows that

$$\boldsymbol{H}_n = \boldsymbol{L}_n\boldsymbol{U}_n\,, \tag{4.85}$$

where $\boldsymbol{H}_n$ is a matrix having the upper HESSENBERG form.

The coefficients u_{ij} and l_{ij} are determined using the biorthogonality property in eq. (4.56). Then, in view of eq. (4.57) we have that

$$\boldsymbol{W}_n^T|\boldsymbol{v}_{n+1}\rangle = \boldsymbol{0}\,, \tag{4.86}$$

that is equivalent to

$$\boldsymbol{W}_n^T\boldsymbol{A}|\boldsymbol{p}_n\rangle - \sum_{i=1}^{n} l_{in}\boldsymbol{W}_n^T|\boldsymbol{v}_n\rangle = \boldsymbol{0}\,. \tag{4.87}$$

Now, we require that the vectors in the P-Q sequence be *A-biorthogonal*, or more precisely

$$\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{P} = \hat{\boldsymbol{D}}\,, \tag{4.88}$$

where

$$\hat{\boldsymbol{D}} = \text{diag}\,(\hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_j, \ldots, \hat{\delta}_n)\,. \tag{4.89}$$

The *A-biorthogonality* for two P-Q vectors can be expressed through

$$\langle \boldsymbol{q}_i|\boldsymbol{A}|\boldsymbol{p}_j\rangle = \begin{cases} \hat{\delta}_j, & i = j\,, \\ 0, & i \neq j\,. \end{cases} \tag{4.90}$$

Using the biorthogonality of the P-Q vectors (4.90) and taking the eq. (4.79), (4.83), (4.57) and finally the eq. (4.87) into account we obtain

$$l_{ij} = 0,\ i = 1, 2, \ldots, n-1,\ j = i+1, \ldots, n\,, \tag{4.91}$$

$$l_{ii} = \beta_i = \hat{\delta}_i/\delta_i. \tag{4.92}$$

Furthermore, by multiplying the eq. (4.74) by $\langle \boldsymbol{q}_j|\boldsymbol{A}$ and using the eq. (4.90) and (4.77) we obtain

$$\begin{aligned} \langle \boldsymbol{q}_i|\boldsymbol{A}|\boldsymbol{p}_j\rangle &= \langle \boldsymbol{q}_i|\boldsymbol{A}|\boldsymbol{v}_j\rangle - \hat{\delta}_i u_{ij} = \\ &= (\boldsymbol{A}^T|\boldsymbol{q}_i\rangle)^T|\boldsymbol{v}_j\rangle - \hat{\delta}_i u_{ij} = \\ &= \eta_{i+1}\langle \boldsymbol{w}_i|\boldsymbol{v}_j\rangle - \hat{\delta}_i u_{ij} = 0\,. \end{aligned} \tag{4.93}$$

With the eq. (4.57), (4.82) and (4.93) it follows that

$$u_{ij} = 0,\ i = 1, 2, \ldots, n-2,\ j = i+1, \ldots, n\,, \tag{4.94}$$

$$u_{i-1,i} = \eta_i\delta_i/\hat{\delta}_{i-1}. \tag{4.95}$$

In view of (4.91) and (4.94) the last terms in eq. (4.74)-(4.77) vanish and hence the eq. (4.74)-(4.77) reduces to a coupled two-term procedure generating the P-Q and V-W sequences.

Then, the LANCZOS triagonal matrix is given by

$$\boldsymbol{T}_{n\times n} = \boldsymbol{L}_{n\times n}\boldsymbol{U}_{n\times n}, \tag{4.96}$$

with

$$\boldsymbol{U}_{n\times n} = \begin{bmatrix} 1 & \beta_2 & & & \\ 0 & 1 & \beta_3 & & \\ & 0 & 1 & \ddots & \\ & & \ddots & \ddots & \beta_n \\ & & & 0 & 1 \end{bmatrix}, \tag{4.97}$$

and

$$\boldsymbol{L}_{n\times n} = \begin{bmatrix} \alpha_1 & 0 & & & \\ \rho_2 & \alpha_2 & 0 & & \\ & \rho_3 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & 0 \\ & & & \rho_n & \alpha_n \end{bmatrix}, \tag{4.98}$$

where

$$\alpha_j = \eta_j\delta_j/\hat{\delta}_{j-1}\,. \tag{4.99}$$

The pseudocode of the coupled two-term LANCZOS process is presented below. The two-term algorithm requires one additional pair of vectors to be stored in comparison to the three-term nonsymmetric version.

Algorithm 4.4.3. The coupled two-term LANCZOS algorithm

(1) choose $|\boldsymbol{v}_1\rangle = \dfrac{|\boldsymbol{p}\rangle}{\|\boldsymbol{p}\|_2}$ and $|\boldsymbol{w}_1\rangle = \dfrac{|\boldsymbol{q}\rangle}{\|\boldsymbol{q}\|_2}$
(2) set $|\boldsymbol{p}_0\rangle = \boldsymbol{0}$ and $|\boldsymbol{q}_0\rangle = \boldsymbol{0}$
(3) set $\rho_1 = \eta_1 = \hat{\delta}_0 = 1$
(4) **for** $j = 1, 2, \ldots, n$ **do**
(5) $\quad \delta_j = \langle \boldsymbol{w}_j | \boldsymbol{v}_j \rangle$
(6) $\quad$ **if** $\delta_j = 0$ **then stop**
(7) $\quad |\boldsymbol{p}_j\rangle = |\boldsymbol{v}_j\rangle - (\eta_j \delta_j / \hat{\delta}_{j-1}) |\boldsymbol{p}_{j-1}\rangle$
(8) $\quad |\boldsymbol{q}_j\rangle = |\boldsymbol{w}_j\rangle - (\rho_j \delta_j / \hat{\delta}_{j-1}) |\boldsymbol{q}_{j-1}\rangle$
(9) $\quad \hat{\delta}_j = \langle \boldsymbol{q}_j | \boldsymbol{A} | \boldsymbol{p}_j \rangle$
(10) $\quad$ **if** $\hat{\delta}_j = 0$ **then stop**
(11) $\quad \beta_j = \hat{\delta}_j / \delta_j$
(12) $\quad |\tilde{\boldsymbol{v}}_{j+1}\rangle = \boldsymbol{A}\,|\boldsymbol{p}_j\rangle - \beta_j |\boldsymbol{v}_j\rangle$
(13) $\quad |\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{A}^T\,|\boldsymbol{q}_j\rangle - \beta_j |\boldsymbol{w}_j\rangle$
(14) $\quad \rho_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(15) $\quad \eta_{j+1} = \|\tilde{\boldsymbol{w}}_{j+1}\|_2$
(16) $\quad$ **if** $\rho_{j+1} = 0$ or $\eta_{j+1} = 0$ **then stop**
(17) $\quad |\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\rho_{j+1}}$
(18) $\quad |\boldsymbol{w}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{w}}_{j+1}\rangle}{\eta_{j+1}}$
(19) **end for**

The coupled two-term version exhibits breakdowns like ones in the three-term recurrences. We note, if δ_j or $\hat{\delta}_{j-1}$ are nonzero, but small in some sense, the near-breakdowns provoking numerical instabilities may appear.

4.4.3 The Look-Ahead Lanczos Algorithm

Unfortunately in the classical LANCZOS algorithm, breakdowns cannot be completely excluded. Indeed, during the construction of the next vectors $|\boldsymbol{v}_{j+1}\rangle$ and $|\boldsymbol{w}_{j+1}\rangle$ an *exact breakdown* (if $\langle \boldsymbol{w}_j | \boldsymbol{v}_j \rangle = 0$) or a *near-breakdown* (if $\langle \boldsymbol{w}_j | \boldsymbol{v}_j \rangle$ is nonzero, but small) can occur. In order to avoid these breakdowns the first practical modification of the LANCZOS algorithm using look-ahead technique has been proposed in [89]. A different implementation of the look-ahead LANCZOS method is given in [44]. In this section a brief description of the look-ahead LANCZOS procedure is given.

As in the classical LANCZOS algorithm two sequences of LANCZOS vectors $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{w}_j\rangle$ are generated and collected in matrices $\boldsymbol{V}$ and $\boldsymbol{W}$. The matrix (4.56) can be presented as a block diagonal with l square blocks on the diagonal, or to be precise

$$\boldsymbol{D} = \text{diag}\,(\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_j, \ldots, \boldsymbol{D}_l)\,, \tag{4.100}$$

where

$$\boldsymbol{D}_j = \boldsymbol{W}_j^T \boldsymbol{V}_j \,. \tag{4.101}$$

The matrices $\boldsymbol{W}_j$ and $\boldsymbol{V}_j$ are defined by their partitioning into l blocks according to the look-ahead steps

$$\boldsymbol{V} = [\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_j, \ldots, \boldsymbol{V}_l] \,, \tag{4.102}$$

and

$$\boldsymbol{W} = [\boldsymbol{W}_1, \boldsymbol{W}_2, \ldots, \boldsymbol{W}_j, \ldots, \boldsymbol{W}_l] \,. \tag{4.103}$$

The matrices $\boldsymbol{W}_j$ and $\boldsymbol{V}_j$ are of size $N \times b_j$. Their columns are the LANCZOS vectors constructed in the j-th look-ahead step. The integer b_j is the length of the j-th look-ahead step and l is the number of look-ahead steps that have been performed during the first n steps of the LANCZOS process. For $j = 1, 2, \ldots, l$ we denote by n_j the index of the first vectors of the blocks $\boldsymbol{W}_j$ and $\boldsymbol{V}_j$, and as following we have

$$\boldsymbol{V}_j = [|\boldsymbol{v}_{n_j}\rangle, |\boldsymbol{v}_{n_j+1}\rangle, \ldots, |\boldsymbol{v}_{n_j+b_j}\rangle] \,, \tag{4.104}$$

and

$$\boldsymbol{W}_j = [|\boldsymbol{w}_{n_j}\rangle, |\boldsymbol{w}_{n_j+1}\rangle, \ldots, |\boldsymbol{w}_{n_j+b_j}\rangle] \,. \tag{4.105}$$

The indices n_j satisfy the following relations

$$n_1 < n_2 < \cdots < n_l \leq n < n_{l+1} \,. \tag{4.106}$$

The vectors $|\boldsymbol{v}_{n_j}\rangle$ and $|\boldsymbol{w}_{n_j}\rangle$ are called *regular* and the vectors in the blocks $\boldsymbol{W}_j$ and $\boldsymbol{V}_j$ are called *inner*. In view of (4.56) and (4.100) the regular vectors are biorthogonal to all previous LANCZOS vectors. That is

$$\langle \boldsymbol{w}_i | \boldsymbol{v}_{n_j} \rangle = \langle \boldsymbol{v}_{n_j} | \boldsymbol{w}_i \rangle = 0, \text{ for } i = 1, 2, \ldots, n_j - 1 \,, \tag{4.107}$$

while the inner vectors in the j-th blocks $\boldsymbol{W}_j$ and $\boldsymbol{V}_j$ are biorthogonal to all vectors from the previous blocks, but necessarily to the LANCZOS vectors in the j-th blocks.

In the look-ahead algorithm the LANCZOS vectors are again generated using only short vector recurrences, which involve vectors from the last two blocks $\boldsymbol{V}_j$, $\boldsymbol{V}_{j-1}$ and $\boldsymbol{W}_j$, $\boldsymbol{W}_{j-1}$ instead of just $|\boldsymbol{v}_j\rangle$, $|\boldsymbol{v}_{j-1}\rangle$ and $|\boldsymbol{w}_j\rangle$, $|\boldsymbol{w}_{j-1}\rangle$ as in the classical version, e.g. the next vector $|\boldsymbol{v}_{j+1}\rangle$ is computed by the next relation

$$\rho_{j+1} |\boldsymbol{v}_{j+1}\rangle = \boldsymbol{A} |\boldsymbol{v}_j\rangle - \boldsymbol{V}_i |\boldsymbol{\phi}_j\rangle - \boldsymbol{V}_{i-1} |\boldsymbol{\psi}_j\rangle \,, \tag{4.108}$$

where $|\boldsymbol{\phi}_j\rangle \in \mathcal{C}^{b_j}$ and $|\boldsymbol{\psi}_j\rangle \in \mathcal{C}^{b_{j-1}}$ are suitable coefficient vectors. The second LANCZOS vector is obtained in an analogous manner. As before, the look-ahead recurrences for the LANCZOS vectors can be summarized in a matrix form. Thus, the LANCZOS matrix T has a block diagonal form with l square blocks on the diagonal, where j-th block has dimension $b_i \times b_i$, $i = 1, 2, \ldots, l$.

We note that if look-ahead steps have a length of 1, i.e. $b_i = 1$, then blocks $\boldsymbol{V}_j$ and $\boldsymbol{W}_j$ consists of only single LANCZOS vectors (the look-ahead process is reduced to the classical algorithm); and true look-ahead steps, i.e. steps of size $b_i > 1$, are only used to avoid exact and near-breakdowns. Typically only a few true look-ahead steps are needed, and their size is mostly $b_i = 2$.

Although the look-ahead technique enhances the stability of LANCZOS methods, the computational costs increases, especially for cases with imlicit matrices. This doctoral thesis does not consider implementations with the look-ahead feature. For further details, properties and implementations of the look-ahead LANCZOS algorithm the reader is referred to [44].

4.5 Spectral Transformations

It is well known that the KRYLOV subspace methods provide approximations to well-separated extremal eigenvalues. The RITZ values tend to converge first to eigenvalues at the boundary of the convex hull of the eigenvalue spectrum. We note that if an original matrix $\boldsymbol{A}$ is HERMITIAN, the largest eigenvalues will be extracted first. If $\boldsymbol{A}$ has complex eigenvalues, the largest ones in magnitude ones will be computed at the beginning. However, in many cases the wanted eigenvalues cannot be well separated or located in the interior of the convex hull of eigenvalues. In these situations KRYLOV subspace methods require many steps to generate acceptable approximations if they converge at all. An alternative approach to expedite the convergence is to employ a *spectral transformation* of $\boldsymbol{A}$ so that needed (and eventually poorly separated) eigenvalues are transformed to extremal (and well-separated) eigenvalues. As a matter of course, the eigenvalues of the initial matrix can be easily recovered from the transformed eigenvalue spectrum.

For this purpose the *shift-inverse spectral transformation* (SI) is typically used. The basic idea is to replace an eigenvalue problem

$$\boldsymbol{A}|\boldsymbol{u}\rangle = \lambda_i|\boldsymbol{u}\rangle \tag{4.109}$$

with a shift-inverse eigenvalue problem

$$(\boldsymbol{A} - \sigma\boldsymbol{I})^{-1}|\boldsymbol{u}\rangle = \nu_i|\boldsymbol{u}\rangle\,, \tag{4.110}$$

where

$$\nu_i = \frac{1}{\lambda_i - \sigma}\,, \tag{4.111}$$

and σ is called a *shift*.

This SI-approach is very effective for finding of eigenvalues closed to σ, because the closest eigenvalue λ_i to σ corresponds to the eigenvalue, that is largest in magnitude for the $(\boldsymbol{A} - \sigma\boldsymbol{I})^{-1}$ matrix. Or in other words, $|\lambda_i - \sigma|^{-1}$ is the largest eigenvalue of $(\boldsymbol{A} - \sigma\boldsymbol{I})^{-1}$ in magnitude, and, therefore, the located λ_i is closest to the σ value. These transformed eigenvalues ν_i of the largest magnitude are precisely the eigenvalues that can be easily computed. If they are found, we can transform them back to eigenvalues of the initial eigenvalue problem by relation

$$\lambda_i = \sigma + \frac{1}{\nu_i}\,. \tag{4.112}$$

The eigenvector $|\boldsymbol{u}\rangle$ associated with ν_i in the transformed problem is the eigenvector corresponding to λ_i in the initial eigenvalue problem, i.e. no transformations of eigenvectors are needed.

The SI-approach is extremely effective in terms of iteration step if a few eigenvalues closest to σ have to be extracted. This is the case when interior eigenvalues are in demand, or when desired eigenvalues are significantly smaller in magnitude that the eigenvalues largest in magnitude, or when they are clustered. By choosing σ very close to a desired eigenvalue the inverse iteration can converge very quickly.

The drawback of the SI-transformation is that linear systems involving the inverse of $(\boldsymbol{A} - \sigma\boldsymbol{I})$ must be solved. This can be carried out via a matrix factorization (e.g. via LU-factorization) or by an iterative method, making

the SI-technique less attractive. Evidently, if we deal with implicit matrices (or implicit eigenvalue problems), only one second possibility may be taken into consideration. In this case the accuracy of the iterative solutions must be commensurate to the convergence tolerance used in the eigenvalue solver.

The next problem is specified for non-symmetric LANCZOS algorithms. We have to operate with $((\boldsymbol{A}-\sigma\boldsymbol{I})^{-1})^T$ in order to generate the second LANCZOS sequence in eq. (4.49), and linear systems involving the inverse of $(\boldsymbol{A}-\sigma\boldsymbol{I})^T$ must be solved as well. Thus, computational costs are radically increased by using implicit matrices.

The eigenvalue convergence of the SI-approach can be controlled by observation of the right and left residual vectors, e.g. for the right residual we have

$$\begin{aligned} |\boldsymbol{r}_{res}\rangle &= (\boldsymbol{A}-\sigma\boldsymbol{I})^{-1}\,|u_i^{(n)}\rangle - \lambda_i^{(n)}\boldsymbol{I}\,|u_i^{(n)}\rangle = \\ &= \rho_{n+1}\{\langle \boldsymbol{e}_n \,|\, \boldsymbol{u}_i^{(n)}\rangle\}\;|\boldsymbol{v}_{n+1}\rangle\,. \end{aligned} \tag{4.113}$$

The norms of the residual vectors are given by eq. (4.68)-(4.69). If $\|\boldsymbol{r}\|_2$ and $\|\boldsymbol{p}\|_2$ are less than a user-specified tolerance, then $\nu_i^{(n)}$ and $|\boldsymbol{u}_i^{(n)}\rangle$ is an approximating eigenpair, corresponding to $\lambda_i^{(n)}$ near σ.

When a spectral transformation is used, additional considerations can be made with respect to stopping criteria to take special advantage of the special nature of the transformed operator $(\boldsymbol{A}-\sigma\boldsymbol{I})^{-1}$. Moreover, the quality of the approximating eigenvectors can be improved with an additional post-processing. For more details see [39].

4.6 Numerical Examples

Let us consider an example. Eigenvalues of a randomly generated matrix $\boldsymbol{A}/\|\boldsymbol{A}\|_2$ are extracted using an implementation of the non-symmetric LANCZOS algorithm in MATLAB (The original code of the Algorithm 4.4.1 has been realized). The size of $\boldsymbol{A}$ is 50×50. The right and left vectors are generated randomly as well. In order to extract desired complex eigenvalues, 20 LANCZOS algorithm steps are performed for three cases:

- the operating matrix $\hat{\boldsymbol{A}}$ is equal to the original matrix $\boldsymbol{A}$;
- the operating matrix is inverse of the original matrix, i.e. $\hat{\boldsymbol{A}} = (\boldsymbol{A})^{-1}$;
- the operating matrix is shift-inverse of the original matrix with $\sigma = -0.385$, i.e.
$$\hat{\boldsymbol{A}} = (\boldsymbol{A} + 0.385\,\boldsymbol{I})^{-1}\;.$$

The original and approximated eigenvalue spectra in the complex z-plane are shown in Fig. 4.6. The original eigenvalues are obtained by the eigenvalue decomposition (EVD). As we can see in Fig. 4.6(a), the complex eigenvalues with the largest magnitude are extracted first, if we operate with the original matrix. Involving the inverse matrix we obtain the eigenvalues with the smallest magnitudes at the beginning as shown in Fig. 4.6(b). Introducing a shift $\sigma = -0.385$ we get the shifted spectrum with eigenvalues closed to σ at first (see Fig. 4.6(c)).

We note, that the choice of right and left start vectors has a slight influence on the convergence of eigenvalues.

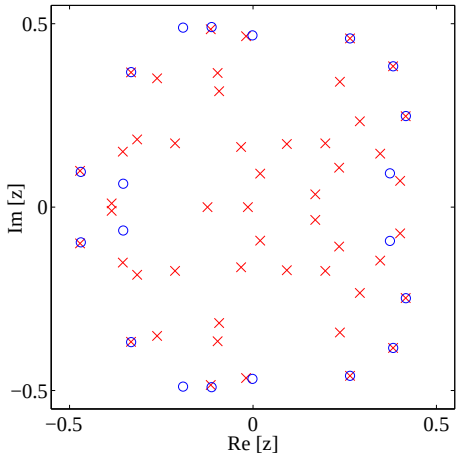
In conclusion, we note that needed eigenvalues may be extracted using KRYLOV subspace methods by spectral transformation with a small number of iterations.

4.7 Remarks

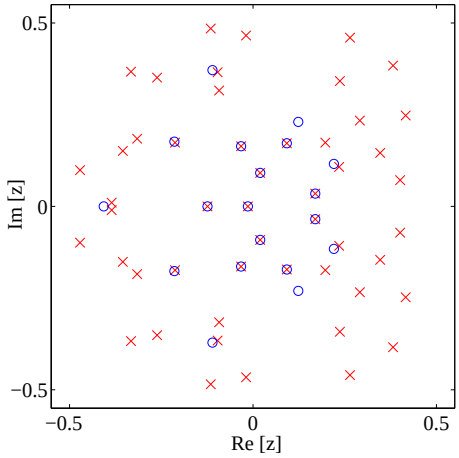
The application of the LANCZOS algorithm to practical electromagnetic problems appears to be computationally more efficient in contrast to the ARNOLDI procedure, because at each step in nonsymmetric LANCZOS recursions only few of the most recently-generated LANCZOS vectors must be kept in storage. In addition, when $\boldsymbol{A}$ is real and symmetric, then both the ARNOLDI and the nonsymmetric recursions reduce to the spare real symmetric recursion. However, there are many open questions regarding the behavior of KRYLOV subspace methods, e.g. what spectral properties of $\boldsymbol{A}$ control the convergence of these methods, or which of both procedures - LANCZOS or ARNOLDI - exhibits better convergence, or does the re(bi)orthogonalization improve the approximation accuracy. These questions should be answered for one particular problem. A comparison and a relationship between the ARNOLDI and the nonsymmetric LANCZOS algorithms as well as studies of their behaviors for several cases are given in [31].

The present doctoral work is focused, first of all, on the classical LANCZOS algorithms and their modifications due to their relatively small computational effort. Despite the economy in memory and small computational overhead, the Lanczos algorithms take a risk of (near)breakdowns and instability. We note, that, typically, the LANCZOS vectors do not remain biorthogonal, and any spurious eigenvalues appearing in the spectrum of any LANCZOS matrix $\boldsymbol{T}_n$ are caused by losses in the biorthogonality of the LANCZOS vectors and represent duplicates of converged eigenvalue approximations [31]. Despite these drawbacks, the Lanczos methods have received much attention and have been proven to be sufficient and useful for relevant problems.

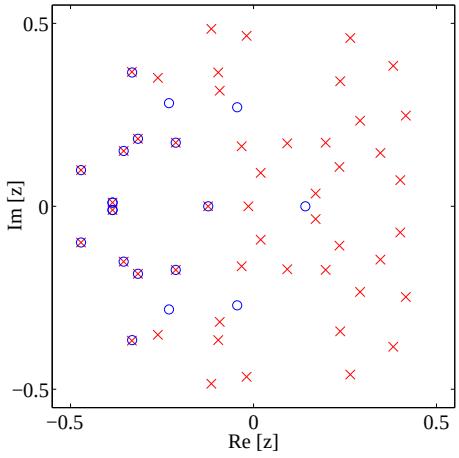
LANCZOS *phenomena* states that for large enough n all of the desired eigenvalues will appear in the eigenvalue spectrum $\boldsymbol{T}_n$ [31]. Using a spectral transformation we can expedite the convergence of desired eigenvalues, or in other words, we can first extract the desired eigenvalues at a small number of algorithm steps.



(a) $\hat{\boldsymbol{A}} = \boldsymbol{A}$



(b) $\hat{\boldsymbol{A}} = (\boldsymbol{A})^{-1}$



(c) $\hat{\boldsymbol{A}} = (\boldsymbol{A} + 0.385\,\boldsymbol{I})^{-1}$

Figure 4.3: Eigenvalue spectrum of a random generated 50×50 matrix $\boldsymbol{A}$ in the z-plane. The crosses denote original eigenvalues, while the circles denote eigenvalues approximated by 20 non-symmetric LANCZOS algorithm steps.

Chapter 5

Model Order Reduction in TLM

The discrete electromagnetic model resulting from the application of the TLM method is cast in a state-space matrix form involving wave amplitudes as state variables. For structures of practical interest the dimension of this matrix can be very large. Therefore, the use of MOR-techniques in the context of TLM-based electromagnetic modeling makes sense.

Specific methodologies for MOR suitable for the discretized electromagnetic TLM system are discussed in this chapter. The *implicit TLM time evolution scheme* is utilized in MOR process. The emphasis of the reduced-order modeling is on the nonsymmetric version of the LANCZOS algorithm and its modifications. Beneficial properties of the TLM matrix allow to modify the nonsymmetric LANCZOS procedure so that in each step the modified algorithm requires one TLM time evolution update and one additional application of the scattering matrix, whereas one LANCZOS step is equivalent to two time updates in the standard TLM scheme due to the imperative to deal with both the TLM operator and its transpose for the construction of a biorthogonal basis for KRYLOV subspaces.

Due to the implicit form of TLM operator an application of block and band KRYLOV algorithms is disadvantageous, because no matrix-vector multiplications are available at all and an impact of the TLM operator on the each vector in a starting block (or/and band) must be performed.

We begin with the general mathematical framework for MOR of discrete electromagnetic systems. This is followed by the development of TLM-MOR techniques and their associated algorithms. The reduced-order models in LAPLACE and z-domains are determined and their efficiency is considered. The opportunities for enhancing the computational efficiency of TLM through the application of MOR-techniques and possibilities for the generation of a macro-model with an optimal order are considered in detail. Thereafter, we give a comparison of the FDTD-, FIT- and TLM-based MOR-techniques.

5.1 State-Space Representation of Discrete EM-Systems

Let's assume that the media to be considered are linear, isotropic, passive and time-independent. We start with MAXWELL's curl equations

$$\nabla \times \boldsymbol{E} = -\jmath\omega\mu\boldsymbol{H} - \boldsymbol{J}_m \,, \tag{5.1}$$

$$\nabla \times \boldsymbol{H} = \jmath\omega\varepsilon\boldsymbol{E} + \sigma\boldsymbol{E} + \boldsymbol{J}_e \,. \tag{5.2}$$

Introducing the LAPLACE variable $s = \jmath\omega$, we write MAXWELL's curl equations in the LAPLACE domain

$$\nabla \times \boldsymbol{E} = -s\mu\boldsymbol{H} - \boldsymbol{J}_m^s \,, \tag{5.3}$$

$$\nabla \times \boldsymbol{H} = s\varepsilon\boldsymbol{E} + \sigma\boldsymbol{E} + \boldsymbol{J}_e^s \,. \tag{5.4}$$

The dependance of the electric and magnetic vectors $\boldsymbol{E}$ and $\boldsymbol{H}$ as well as current densities of imposed magnetic and electric sources $\boldsymbol{J}_m^s$ and $\boldsymbol{J}_e^s$ on the LAPLACE variable is suppressed for simplicity.

MAXWELL's curl equations (5.3)-(5.4) are usually approximated by a finite-difference scheme or by a finite-volume scheme. A discrete form of the system (5.3)-(5.4) can be obtained through the discretization of the curl operators using one of various finite methods as following:

$$\left(\begin{bmatrix} \boldsymbol{S}_e & \boldsymbol{D}_h \\ \boldsymbol{D}_e & \boldsymbol{0} \end{bmatrix} + s \begin{bmatrix} \boldsymbol{P}_e & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_m \end{bmatrix} \right) \begin{bmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{J}_e^s \\ \boldsymbol{J}_m^s \end{bmatrix} , \tag{5.5}$$

where the sparse matrices $\boldsymbol{D}_e$ and $\boldsymbol{D}_h$ denote the discrete forms of the curl operators; the matrices $\boldsymbol{S}_e$, $\boldsymbol{P}_e$ and $\boldsymbol{P}_m$, having diagonal form, represent the discrete form of the electric conductivity, electric permittivity and magnetic permeability respectively.

The discretization of the curl operators in MAXWELL's curl equations using the FDTD HILBERT space representation is presented in [100].

Now we write the system (5.5) in operator form using DIRAC notation. For this purpose we define the vector of state variables $|\boldsymbol{F}\rangle$, containing the discrete unknown electric and magnetic-field components, as

$$|\boldsymbol{F}\rangle = \sum_{l,m,n=-\infty}^{+\infty} \begin{bmatrix} E^x_{l,m,n} \\ E^y_{l,m,n} \\ E^z_{l,m,n} \\ H^x_{l,m,n} \\ H^y_{l,m,n} \\ H^z_{l,m,n} \end{bmatrix} |l,m,n\rangle \,. \tag{5.6}$$

Next, we introduce the vector $|\boldsymbol{J}\rangle$, which represents the LAPLACE transforms of the spatial distribution of the electric- and magnetic-current sources

$$|\boldsymbol{J}^s\rangle = \sum_{l,m,n=-\infty}^{+\infty} \begin{bmatrix} J^x_{e\ l,m,n} \\ J^y_{e\ l,m,n} \\ J^z_{e\ l,m,n} \\ J^x_{m\ l,m,n} \\ J^y_{m\ l,m,n} \\ J^z_{m\ l,m,n} \end{bmatrix} |l,m,n\rangle \,. \tag{5.7}$$

Now, we assume bounded media, and the dimension of $|\boldsymbol{F}\rangle$ is denoted as $N = 6 \times l \times m \times n$. Thus, $|\boldsymbol{F}\rangle \in \mathcal{H}^N$, and $|\boldsymbol{J}^s\rangle \in \mathcal{H}^N$, and the discrete curl operators $\boldsymbol{D}_e$, $\boldsymbol{D}_h$ and operators $\boldsymbol{S}_e$, $\boldsymbol{P}_e$ and $\boldsymbol{P}_m$ are in $\mathcal{H}^{N/2\times N/2}$.

Since our goal is to develop reduced-order models for multi-port electromagnetic systems, the grid source distributions must be presented through sources at ports. The excitation of the discrete system at the p_1 ports can be presented as

$$|\boldsymbol{J}^s\rangle = \boldsymbol{B}_1 \, |\boldsymbol{i}\rangle \, , \tag{5.8}$$

where the $N \times p_1$ selector operator $\boldsymbol{B}_1$ translates the port currents $|\boldsymbol{i}\rangle \in \mathcal{H}^{p_1}$ into grid current sources, or in other words, the operator $\boldsymbol{B}_1 \in \mathcal{H}^{N\times p_1}$ imposes a current pattern on a port plane.

With all these definitions the resulting compact operator form of (5.5) is given as

$$(\mathbf{G} + s\mathbf{C}) \, |\boldsymbol{F}\rangle = \boldsymbol{B}_1 \, |\boldsymbol{i}\rangle \ , \tag{5.9}$$

where

$$\boldsymbol{G} = \left[\begin{array}{cc} \boldsymbol{S}_e & \boldsymbol{D}_h \\ \boldsymbol{D}_e & \mathbf{0} \end{array} \right] \in \mathcal{H}^{N\times N} \text{ and} \tag{5.10}$$

$$\mathbf{C} = \left[\begin{array}{cc} \boldsymbol{P}_e & \mathbf{0} \\ \mathbf{0} & \boldsymbol{P}_m \end{array} \right] \in \mathcal{H}^{N\times N} \, . \tag{5.11}$$

We denote that due to the passivity of the media, the matrices $\boldsymbol{P}_e$, $\boldsymbol{P}_m$ are positive definite (permittivity and permeability matrices are symmetric and positive definite) and $\boldsymbol{S}_e$ is nonnegative definite. Additionally, each of them is symmetric. Hence, $\boldsymbol{C}$ is symmetric and positively defined. The spatial differential operator matrices $\boldsymbol{D}_e$ and $\boldsymbol{D}_m$ are symmetric, but $\boldsymbol{G}$ does not preserve the symmetry property. Moreover, the operator $\boldsymbol{G}$ is skew-symmetric if $\boldsymbol{S}_e = \mathbf{0}$. That is

$$\boldsymbol{G} = -\boldsymbol{G}^T = \left[\begin{array}{cc} \boldsymbol{S}_e & -\boldsymbol{D}_e \\ \boldsymbol{D}_e & \mathbf{0} \end{array} \right] \in \mathcal{H}^{N\times N} \, , \tag{5.12}$$

due to the property $\boldsymbol{D}_h = -\boldsymbol{D}_e$ followed from the system (5.3)-(5.4) and definition (5.5).

Introducing a desired output vector as an output voltage

$$|\boldsymbol{u}\rangle = \boldsymbol{B}_2 |\boldsymbol{F}\rangle \in \mathcal{H}^{p_2} \tag{5.13}$$

we obtain the linear state-space representation of the discretized electromagnetic system (5.5) as

$$\begin{aligned} (\mathbf{G} + s\mathbf{C}) \, |\boldsymbol{F}\rangle &= \boldsymbol{B}_1 \, |\boldsymbol{i}\rangle \ , \\ |\boldsymbol{u}\rangle &= \boldsymbol{B}_2 |\boldsymbol{F}\rangle \, , \end{aligned} \tag{5.14}$$

where $\boldsymbol{B}_2 \in \mathcal{H}^{N\times p_2}$ is a selector, extracting the the generalized voltage $|\boldsymbol{u}\rangle$ from the grid fields at the p_2 ports. We note that using the orthogonality of waveguide modes and proper normalizing, we can achieve the symmetry $\boldsymbol{B}_1 = \boldsymbol{B}_2^T$. As it has been mentioned above (see Chapter 2 and 3), the matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ can be obtained by a 2-D eigenvalue solution at the ports.

From (5.14) we obtain the output voltage vector $|\boldsymbol{u}\rangle$ as

$$|\boldsymbol{u}\rangle = \boldsymbol{B}_2 \left(\boldsymbol{G} + s\boldsymbol{C}\right)^{-1} \boldsymbol{B}_1 \, |\boldsymbol{i}\rangle \, . \tag{5.15}$$

Introducing a *system matrix*

$$\boldsymbol{A}_1 = -\boldsymbol{C}^{-1}\,\boldsymbol{G} = -\begin{bmatrix} \boldsymbol{P}_e & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_m \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{S}_e & -\boldsymbol{D}_e \\ \boldsymbol{D}_e & \boldsymbol{0} \end{bmatrix} \tag{5.16}$$

and modifying (or renaming) $\boldsymbol{B}_1$ as follows

$$\hat{\boldsymbol{B}}_1 = \boldsymbol{C}^{-1}\boldsymbol{B}_1 \in \mathcal{H}^{N\times N}\,, \tag{5.17}$$

we obtain from (5.15) the output voltage vector as

$$|\boldsymbol{u}\rangle = \boldsymbol{B}_2\,(s\boldsymbol{I} - \boldsymbol{A}_1)^{-1}\,\hat{\boldsymbol{B}}_1\,|\boldsymbol{i}\rangle\,, \tag{5.18}$$

and, hence, the multi-port *Generalized Impedance Matrix* (GIM) is given by

$$\boldsymbol{Z}_1(s) = \boldsymbol{B}_2\,(s\boldsymbol{I} - \boldsymbol{A}_1)^{-1}\,\hat{\boldsymbol{B}}_1\ \in \mathcal{H}^{p_1\times p_2}\,. \tag{5.19}$$

We emphasize, that the size of $\boldsymbol{Z}(s)$ is $p_1 \times p_2$, while the size of the *system operator* $\boldsymbol{A}_1$ is $N \times N$.

If the system (5.5) consists of no losses, and as following the system (5.14) is loss free as well, the *curl-curl* formulation allows us to obtain a second degree system [125] from (5.5) and (5.14), having only half of state variables, as only the electric field currents and voltages are relevant. In this case the GIM results in

$$\boldsymbol{Z}_2(s) = s\boldsymbol{B}_2'\,(s^2\boldsymbol{I} - \boldsymbol{A}_2)^{-1}\,\hat{\boldsymbol{B}}_1'\ \in \mathcal{H}^{p_1\times p_2}\,. \tag{5.20}$$

The system matrix $\boldsymbol{A}_2 \in \mathcal{H}^{N/2\times N/2}$ is real, semi-definite and symmetric and can be proven to be stable and passive [3].

It is obvious that the eigenvalues of loss-free skew-symmetric matrix $\boldsymbol{A}_1$ and, as following, the eigenvalues of $\boldsymbol{A}_2$ are imaginary and correspond to the operating frequencies by

$$\lambda_i = \jmath\,\omega_i\,. \tag{5.21}$$

5.1.1 Reduced-Order Models of Discrete Electromagnetic Systems

Now, we introduce the change of variables as

$$|\boldsymbol{F}\rangle = \boldsymbol{V}|\hat{\boldsymbol{F}}\rangle\,, \tag{5.22}$$

where $|\hat{\boldsymbol{F}}\rangle$ is the reduced state vector of the length n, whereas $n << N$. The matrix $\boldsymbol{V}$ of dimension $N \times n$ satisfies the relation

$$(\boldsymbol{W}^T)_{n\times N}\boldsymbol{V}_{N\times n} = \boldsymbol{I}_{n\times n}\,. \tag{5.23}$$

In other words, $\boldsymbol{W} \in \mathcal{H}_1^{N\times n}$ and $\boldsymbol{V} \in \mathcal{H}_2^{N\times n}$ present a biorthonormal basis.

Then by substitution of (5.22) into the system (5.14) and by multiplication on the left by $\boldsymbol{W}^T$ we have

$$\begin{aligned} \boldsymbol{W}^T\,(\mathbf{G} + s\mathbf{C})\,\boldsymbol{V}|\hat{\boldsymbol{F}}\rangle &= \boldsymbol{W}^T\boldsymbol{B}_1\,|\boldsymbol{i}\rangle\,, \\ |\boldsymbol{u}\rangle &= \boldsymbol{B}_2\boldsymbol{V}|\hat{\boldsymbol{F}}\rangle\,. \end{aligned} \tag{5.24}$$

By matrix manipulations we obtain the GIM from (5.19) and (5.24) as

$$\boldsymbol{Z}_{n_1}(s) = \boldsymbol{B}_2\boldsymbol{V}\,(s\boldsymbol{I} - \boldsymbol{W}^T\boldsymbol{A}_1\boldsymbol{V})^{-1}\,\boldsymbol{W}^T\hat{\boldsymbol{B}}_1\ \in \mathcal{H}^{p_1\times p_2}\,, \tag{5.25}$$

where $(\boldsymbol{W}^T)\boldsymbol{A}_1\boldsymbol{V}$ is of dimension $n \times n$ and called the *reduced system matrix* whose eigenvalues approximate the eigenvalue spectrum of $N \times N$-large system operator $\boldsymbol{A}_1$.

Inducing an orthonormal basis

$$(\boldsymbol{V}^T)_{n\times N}\boldsymbol{V}_{N\times n} = \boldsymbol{I}_{n\times n}\,, \tag{5.26}$$

we similarly obtain from (5.20)

$$\boldsymbol{Z}_{n_2}(s) = s\boldsymbol{B}_2'\,\boldsymbol{V}(s^2\boldsymbol{I} - \boldsymbol{V}^T\boldsymbol{A}_2\boldsymbol{V})^{-1}\,\boldsymbol{V}^T\hat{\boldsymbol{B}}_1' \in \mathcal{H}^{p_1\times p_2}\,, \tag{5.27}$$

where $(\boldsymbol{V}^T)\boldsymbol{A}_2\boldsymbol{V}$ of dimension $n \times n$ is much smaller than the initial matrix $\boldsymbol{A}_2$ of dimension $N/2 \times N/2$ and contains eigenfrequencies ω_i (5.21) of loss-free system (5.5) and, consequently, (5.14).

The GIMs $\boldsymbol{Z}_{n_1}$ (5.25) and $\boldsymbol{Z}_{n_2}$ (5.27) represent reduced-order models of the initial system (5.5) and, consequently, (5.14) in that sense they approximate the original GIMs given by (5.19) and (5.20) respectively.

A multi-port impulse response $\boldsymbol{\mathcal{H}}(s)$ of the systems (5.5) and (5.14) can be defined in a similar way and its reduced-order approximants $\tilde{\boldsymbol{\mathcal{H}}}_{n_1}(s)$ (in general case) and $\tilde{\boldsymbol{\mathcal{H}}}_{n_2}(s)$ (in lossless case) can be derived.

5.2 The TLM State-Space System

In order to obtain a state-space representation of the TLM in the frequency-domain we start with the TLM time evolution scheme

$$|_{k+1}\boldsymbol{a}\rangle = {}_k\boldsymbol{\Gamma}\,{}_k\boldsymbol{S}|_k\boldsymbol{a}\rangle\,. \tag{5.28}$$

We recall that in the general case of TLM we have that $N = 18 \times l \times m \times n$. We skip the left subscript k for the operators $\boldsymbol{\Gamma}$ and $\boldsymbol{S}$ because of linearity.

We consider the time evolution of the field in an interval from k_1 to k_2. We introduce the z-transform $|\tilde{\boldsymbol{a}}\rangle \in \mathcal{H}^N$ as

$$|\tilde{\boldsymbol{a}}\rangle \;=\; \sum_{k=k_1}^{k_2} \frac{1}{z^k}\,|_k\boldsymbol{a}\rangle\;. \tag{5.29}$$

Thus, one time-step in the time evolution scheme (5.28) accords to the multiplication with z-variable in the z-domain.

From (5.28) and (5.29) we obtain

$$z\,|\tilde{\boldsymbol{a}}\rangle = \boldsymbol{\Gamma}\,\boldsymbol{S}\,|\tilde{\boldsymbol{a}}\rangle\,, \tag{5.30}$$

and, as following, we obtain *the state equation of the TLM-system in z-domain*

$$(z - \boldsymbol{\Gamma}\,\boldsymbol{S})\,|\tilde{\boldsymbol{a}}\rangle = \boldsymbol{0}\,. \tag{5.31}$$

Substituting

$$z = e^{\jmath\,\Omega} \tag{5.32}$$

into (5.29) we obtain the state vector of wave amplitudes in the frequency-domain

$$|\tilde{\boldsymbol{a}}\rangle_\Omega \;=\; \sum_{k=k_1}^{k_2} e^{-\jmath\,k\Omega}\,|_k\boldsymbol{a}\rangle \in \mathcal{H}^N \tag{5.33}$$

with the normalized frequency

$$\Omega = \omega \Delta t = 2\pi \Delta t f \,, \tag{5.34}$$

where f is an operating frequency.

From (5.31), (5.32) and (5.33) we derive *the state equation of the TLM-system in the frequency-domain*

$$(e^{\jmath\,\Omega} - \boldsymbol{\Gamma}\,\boldsymbol{S})\;|\tilde{\boldsymbol{a}}\rangle_\Omega = \mathbf{0}\,. \tag{5.35}$$

The eigenvalues λ_i of $\boldsymbol{\Gamma}\,\boldsymbol{S}$ extracted by solving

$$\det\,(e^{\jmath\,\Omega} - \boldsymbol{\Gamma}\,\boldsymbol{S}) = \mathbf{0} \tag{5.36}$$

are related to operating eigenfrequencies via

$$\lambda_i = e^{\jmath\,\Omega_i}\,. \tag{5.37}$$

So, from $\lambda_1 = -1$ follows that $\Omega_1 = \pi$, i.e. to the eigenvalue $\lambda_1 = -1$ corresponds an evanescent nonphysical TLM-eigensolution with the frequency $f_1 = \frac{1}{2\Delta t}$. For $\lambda_2 = 1$ we have a static eigensolution, because of $\Omega_2 = 0$ and, as following, $f_2 = 0$. In [65] it has been shown that the frequency band during TLM simulation is bound by the lowest cutoff frequency $f_c = \frac{1}{6\Delta t}$. Therefore, the frequency band of a TLM-system is given by

$$-\frac{1}{6\Delta t} < f < \frac{1}{6\Delta t}\,. \tag{5.38}$$

Assuming the eqn. (5.5) and (5.35) to be used for modeling a same electromagnetic structure we may state, that the operator $\boldsymbol{\Gamma}\,\boldsymbol{S}$ in the scheme (5.35) is equivalent to the operator $e^{\boldsymbol{A}_1 \Delta t}$ from the system (5.14), or to be more precisely, the system (5.35) may be written as

$$e^{\jmath\,\Omega}\hat{\boldsymbol{M}}\;|\tilde{\boldsymbol{a}}\rangle_\Omega = e^{\boldsymbol{A}_1 \Delta t}\;\hat{\boldsymbol{M}}|\tilde{\boldsymbol{a}}\rangle_\Omega\,, \tag{5.39}$$

where the operator

$$\hat{\boldsymbol{M}} : \mathcal{H}^{18\times l\times m\times n} \longrightarrow \mathcal{H}^{6\times l\times m\times n} \tag{5.40}$$

maps the vector of wave amplitudes $|\tilde{\boldsymbol{a}}\rangle_\Omega$ onto the vector of electromagnetic fields $|\boldsymbol{F}\rangle$ defined by (5.6), i.e.

$$|\boldsymbol{F}\rangle = \hat{\boldsymbol{M}}\,|\tilde{\boldsymbol{a}}\rangle_\Omega\,. \tag{5.41}$$

5.2.1 Impulse Response of the TLM System

Introducing an excitation vector $|\tilde{\boldsymbol{a}}\rangle_{in} \in \mathcal{H}_F^{p_1}$ and a vector of port responses $|\tilde{\boldsymbol{a}}\rangle_{out} \in \mathcal{H}_F^{p_2}$, we obtain the TLM-state-space-system form in the z-domain from (5.28) and (5.31)

$$\begin{aligned} z\,|\tilde{\boldsymbol{a}}\rangle &= \boldsymbol{\Gamma}\boldsymbol{S}\,|\tilde{\boldsymbol{a}}\rangle + \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\,, \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\,|\tilde{\boldsymbol{a}}\rangle\,, \end{aligned} \tag{5.42}$$

where $\boldsymbol{R} \in \mathcal{H}^{N \times p_1}$ and $\boldsymbol{Q} \in \mathcal{H}^{p_2 \times N}$ are selector operators, associated with the non-zero wave amplitudes in the state-space vector $|\tilde{\boldsymbol{a}}\rangle$ corresponding to the p_1 and p_2 ports through which the structure is excited and observed.

From (5.42) we formally solve for the output vector as

$$|\tilde{\boldsymbol{a}}\rangle_{out} \quad = \quad \boldsymbol{Q}\,(z\boldsymbol{I} \;-\; \boldsymbol{\Gamma S})\;\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\;. \tag{5.43}$$

From this result, the $p_1 \times p_2$-port matrix impulse response is identified as the matrix rational function

$$\tilde{\boldsymbol{\mathcal{H}}}(z) = \boldsymbol{Q}(z\boldsymbol{I} - \boldsymbol{\Gamma S})^{-1}\,\boldsymbol{R} = \sum_{n=1}^{\infty} z^{-k}\boldsymbol{Q}\,(\boldsymbol{\Gamma S})^{k-1}\,\boldsymbol{R} \in \mathcal{H}^{p_1 \times p_2}\;. \tag{5.44}$$

The coefficients $(\boldsymbol{Q}\,(\boldsymbol{\Gamma S})^{k-1}\,\boldsymbol{R})$ in (5.44) are termed the *moments* of $\tilde{\boldsymbol{\mathcal{H}}}$ [43].

5.2.2 Generalized Impedance Matrix of the TLM System

Now, we write the system (5.42) in another form. That is

$$\begin{aligned} z\,|\tilde{\boldsymbol{a}}\rangle \quad &= \quad \boldsymbol{\Gamma S}\,|\tilde{\boldsymbol{a}}\rangle + \boldsymbol{R}_1\,|\boldsymbol{i}\rangle\;, \\ |\boldsymbol{u}\rangle \quad &= \quad \boldsymbol{Q}_1\,|\tilde{\boldsymbol{a}}\rangle\;, \end{aligned} \tag{5.45}$$

where the vector $|\boldsymbol{i}\rangle \in \mathcal{H}^{p_1}$ contains the expansion coefficients for the tangential magnetic field on a port; the vector $|\boldsymbol{u}\rangle \in \mathcal{H}^{p_2}$ contains the relevant expansion coefficients, corresponding to the tangential electric fields at ports.

The operators $\boldsymbol{R}_1$ and $\boldsymbol{Q}_1$ contain the field patterns $\boldsymbol{H}_{2D}$ and $\boldsymbol{E}_{2D}$ of the port planes and can be obtained by a computationally cheap 2-D external solver. The matrix $\boldsymbol{R}_1$ imposes a current pattern on the port plane which corresponds to $\boldsymbol{n} \times \boldsymbol{H}_{2D}$, and $\boldsymbol{Q}_1$ extracts the generalized voltage $|\boldsymbol{u}\rangle$ from the fields at the ports, based on the $\boldsymbol{E}_{2D} \times \boldsymbol{H}_{2D}$ orthogonality of waveguide modes. Using proper normalizing we can achieve the symmetry relation $\boldsymbol{R}_1 = (\boldsymbol{Q}_1)^T$. Or in other words, the operators $\boldsymbol{R}_1$ and $\boldsymbol{Q}_1$ translate the port currents $|\boldsymbol{i}\rangle$ and port voltages $|\boldsymbol{u}\rangle$ into grid incident and output wave amplitudes.

Therefore, we obtain the GIM for the multi-port TLM system as

$$\boldsymbol{Z}(z) = \boldsymbol{Q}_1(z\boldsymbol{I} - \boldsymbol{\Gamma S})^{-1}\,\boldsymbol{R}_1 = \sum_{k=1}^{\infty} z^{-k}\boldsymbol{Q}_1\,(\boldsymbol{\Gamma S})^{k-1}\,\boldsymbol{R}_1 \in \mathcal{H}^{p_1 \times p_2}\;, \tag{5.46}$$

and the vector of output voltages as

$$|\boldsymbol{u}\rangle = \boldsymbol{Z}|\boldsymbol{i}\rangle\,. \tag{5.47}$$

5.3 Reduced-Order Modeling in TLM

The basic idea of reduced-order modeling is to replace the original systems (5.42) and (5.45) by a system of the same type but of a much smaller dimension.

Although the order N of the original system is very high, a large number of its eigenstates are the result of the spatial oversampling of the fields in space, resulting from both the desire for accurate modeling of fine geometric features and the need for containing numerical dispersion, thus ensuring the accuracy of

the numerical modeling of the waves. Therefore, it is meaningful to generate a reduced-order model containing predominantly those eigenstates that correspond to the physical eigenfrequencies of the system and, in particular, those that impact the states involved in the excitation and observation of the system.

Thus, we seek a model of reduced order that approximates the impulse response (5.44) and GIM (5.46), in an equivalent sense, the discrete TLM-models of (5.42) and (5.45). This may be done using moment matching techniques based on KRYLOV subspace methods. The basic idea is to project the systems of (5.42) and (5.45) of dimension N onto a space of significantly lower dimension n such that only the first n moments in the LAURENT series expansions of the impulse response (5.44) and the GIM (5.46) of the original system, and the reduced system are matched. Some of the most commonly used approaches for this purpose are discussed next.

5.3.1 Orthogonal Projection

In order to derive a reduced-order model of the TLM-system (5.42) we induce an orthonormal basis

$$(\boldsymbol{V})^T_{n\times N}\boldsymbol{V}_{N\times n} = \boldsymbol{I}\,. \tag{5.48}$$

Using a projection operator

$$\boldsymbol{V} : \mathcal{H}^n \rightarrow \mathcal{H}^N \tag{5.49}$$

and its transpose

$$\boldsymbol{V}^T : \mathcal{H}^N \rightarrow \mathcal{H}^n \tag{5.50}$$

we obtain the reduced state vector of length $n << N$ as

$$|\tilde{\boldsymbol{a}_n}\rangle = \boldsymbol{V}^T|\tilde{\boldsymbol{a}}\rangle\,. \tag{5.51}$$

Consequently we have

$$|\tilde{\boldsymbol{a}}\rangle = \boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\ . \tag{5.52}$$

Then the system (5.42) can be written with the reduced state vector (5.51) as following

$$\begin{aligned} z\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{\Gamma S}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle + \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\ . \end{aligned} \tag{5.53}$$

By multiplication of the first row in (5.53) by $\boldsymbol{V}^T$ we obtain

$$\begin{aligned} z\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{V}^T\boldsymbol{\Gamma S}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle + \boldsymbol{V}^T\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\ , \end{aligned} \tag{5.54}$$

and by using matrix manipulations we have

$$\begin{aligned} (z\boldsymbol{I} - \boldsymbol{V}^T\boldsymbol{\Gamma S}\,\boldsymbol{V})\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{V}^T\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\ . \end{aligned} \tag{5.55}$$

From (5.55) we obtain the output vector

$$|\tilde{\boldsymbol{a}}\rangle_{out} = \boldsymbol{Q}\,\boldsymbol{V}(z\boldsymbol{I} - \boldsymbol{V}^T\boldsymbol{\Gamma S}\,\boldsymbol{V})^{-1}\boldsymbol{V}^T\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \tag{5.56}$$

and multi-port impulse response

$$\tilde{\boldsymbol{\mathcal{H}}}_{n_1}(z) = \boldsymbol{QV}(z\boldsymbol{I} - (\boldsymbol{\Gamma S})_n)^{-1}\boldsymbol{V}^T\boldsymbol{R} \in \mathcal{H}^{p_1 \times p_2}\,, \tag{5.57}$$

where

$$(\boldsymbol{\Gamma S})_n = \boldsymbol{V}^T\boldsymbol{\Gamma S}\,\boldsymbol{V} \tag{5.58}$$

is the *reduced TLM-operator* of dimension $n \times n$.

In order to obtain the orthogonal basis $\boldsymbol{V}$ the ARNOLDI algorithm can be utilized, wich projects the general TLM operator onto the KRYLOV subspace,which is induced by $\boldsymbol{\Gamma S}$ and a start vector being the excitation vector. Thus the reduced TLM operator (5.58) is presented by an upper HESSENBERG matrix $\boldsymbol{H}$.

Using the eigenvalue decomposition of $\boldsymbol{H}$, i.e.

$$\boldsymbol{H} = \boldsymbol{U\Lambda U}^{-1}\,, \tag{5.59}$$

we can write the multi-port impulse response in the next form

$$\tilde{\boldsymbol{\mathcal{H}}}_{n_1}(z) = \boldsymbol{QVU}(z\boldsymbol{I} - \boldsymbol{\Lambda})^{-1}\boldsymbol{U}^{-1}\boldsymbol{V}^T\boldsymbol{R}. \tag{5.60}$$

Let $|\boldsymbol{r}_j\rangle, j = 1, 2, \ldots, p_1$ be the j-th column of the selector matrix $\boldsymbol{R}$ and $\langle\boldsymbol{q}_i|, i = 1, 2, \ldots, p_2$ be the i-th row of the selector matrix $\boldsymbol{Q}$. Now we define the row vector $\langle\boldsymbol{\alpha}|^{(i)}$ and the column vector $|\boldsymbol{\beta}\rangle^{(j)}$ as following

$$\langle\boldsymbol{\alpha}|^{(i)} = \langle\boldsymbol{q}_i|\boldsymbol{VU}\,, \tag{5.61}$$

$$|\boldsymbol{\beta}\rangle^{(j)} = \boldsymbol{U}^{-1}\boldsymbol{V}^T|\boldsymbol{r}_j\rangle\,. \tag{5.62}$$

Then pole/residue distribution of the (i,j)-th element of the transfer function $\tilde{\boldsymbol{\mathcal{H}}}_{n_1}$ (from port j to port i) becomes

$$\tilde{\mathcal{H}}_{n_1}^{(i,j)} = \sum_{\hat{n}=1}^{n} \frac{\alpha_{\hat{n}}^{(i)}\beta_{\hat{n}}^{(j)}}{z - \lambda_{\hat{n}}} = \sum_{\hat{n}=1}^{n} \frac{c_{\hat{n}}^{(i,j)}}{z - \lambda_{\hat{n}}}\,. \tag{5.63}$$

In a similar way the reduced GIM may be obtained from (5.46) as

$$\boldsymbol{Z}_{n_1}(z) = \boldsymbol{Q_1VU}(z\boldsymbol{I} - \boldsymbol{\Lambda})^{-1}\boldsymbol{U}^{-1}\boldsymbol{V}^T\boldsymbol{R_1}\,. \tag{5.64}$$

The sketch of TLM-MOR based on the ARNOLDI procedure generating (5.60) (and/or (5.63)) is given below.

Algorithm 5.3.1. The TLM-ROM based on the ARNOLDI process.

(1) choose $|\tilde{\boldsymbol{v}}\rangle = \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}$
(2) $|\boldsymbol{v}_1\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\|\tilde{\boldsymbol{v}}\|_2}$
(3) **for** $j = 1, 2, \ldots, n$ **do**
(4) $\quad|\tilde{\boldsymbol{v}}\rangle = (\boldsymbol{\Gamma S})\,|\boldsymbol{v}_j\rangle$
(5) $\quad$**for** $i = 1, 2, \ldots, j$ **do**
(6) $\quad\quad h_{ij} = \langle\tilde{\boldsymbol{v}}|\boldsymbol{v}_i\rangle$
(7) $\quad\quad|\tilde{\boldsymbol{v}}\rangle = |\tilde{\boldsymbol{v}}\rangle - h_{ij}|\boldsymbol{v}_i\rangle$
(8) $\quad$**end for**
(9) $\quad$reorthogonalize if necessary
(10) $\quad\beta_{j+1} = \|\tilde{\boldsymbol{v}}\|_2$

(11) **if** $\beta_{j+1} = 0$ **then stop**
(12) $\boldsymbol{H}_j = \boldsymbol{U}\boldsymbol{\Lambda}_j\boldsymbol{U}^{-1}$
(13) test for convergence
(14) $|\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\beta_{j+1}}$
(15) **end for**
(16) $\boldsymbol{H}_n = \boldsymbol{U}\boldsymbol{\Lambda}_n\boldsymbol{U}^{-1}$
(17) compute impulse response $\tilde{\boldsymbol{\mathcal{H}}}_{n_1}$ (5.60) and/or GIM $\boldsymbol{Z}_{n_1}$ (5.64)
(18) compute $|\tilde{\boldsymbol{a}}\rangle_{out}$

The algorithm (5.3.1) yields the orthogonal basis $\boldsymbol{V}$ and the $n \times n$ reduced TLM operator (5.58) as an upper HESSENBERG matrix. The $N \times n$ matrix $\boldsymbol{V}$ form an orthonormal basis for the KRYLOV subspace (see Chapter 4) with a start vector as an excitation vector.

5.3.2 Oblique Projection

Next, we introduce a biorthogonal basis

$$(\boldsymbol{W})^T_{n\times N}\boldsymbol{V}_{N\times n} = \boldsymbol{D}_{n\times n}\,. \tag{5.65}$$

Using the projection operator

$$\boldsymbol{V} : \mathcal{H}^n \rightarrow \mathcal{H}^N \tag{5.66}$$

and its transpose

$$\boldsymbol{W}^T : \mathcal{H}^N \rightarrow \mathcal{H}^n \tag{5.67}$$

we obtain the reduced state vector as

$$|\tilde{\boldsymbol{a}}_n\rangle = \boldsymbol{W}^T\,|\tilde{\boldsymbol{a}}\rangle\,, \tag{5.68}$$

and the initial state vector as

$$|\tilde{\boldsymbol{a}}\rangle = \boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\,. \tag{5.69}$$

Then the system (5.42) can be written with the reduced state vector as

$$\begin{aligned} z\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{\Gamma}\boldsymbol{S}\;\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle + \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\;, \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\;\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\;, \end{aligned} \tag{5.70}$$

By multiplication of the first row in (5.70) by $\boldsymbol{W}^T$ and by using eq. (5.65) we obtain the next system

$$\begin{aligned} z\,\boldsymbol{D}\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{W}^T\,\boldsymbol{\Gamma}\boldsymbol{S}\;\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle + \boldsymbol{W}^T\,\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\;, \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\;\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\;. \end{aligned} \tag{5.71}$$

Applying the similarity transformation to the general TLM operator $\boldsymbol{\Gamma}\boldsymbol{S}$ by means of the LANCZOS algorithm we obtain the reduced TLM operator $(\boldsymbol{\Gamma}\boldsymbol{S})_n$

$$\boldsymbol{W}^T\,\boldsymbol{\Gamma}\boldsymbol{S}\;\boldsymbol{V} = \boldsymbol{D}\boldsymbol{T} \equiv (\boldsymbol{\Gamma}\boldsymbol{S})_n\,. \tag{5.72}$$

From (5.71) and (5.72) it follows that

$$\begin{aligned} z\,\boldsymbol{D}\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{DT}\,|\tilde{\boldsymbol{a}}_n\rangle + \boldsymbol{W}^T\,\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\ , \end{aligned} \tag{5.73}$$

$$\begin{aligned} \boldsymbol{D}(z\boldsymbol{I}-\boldsymbol{T})\,|\tilde{\boldsymbol{a}}_n\rangle &= \boldsymbol{W}^T\,\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in} \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\,\boldsymbol{V}\,|\tilde{\boldsymbol{a}}_n\rangle\ , \end{aligned} \tag{5.74}$$

From (5.74) we express the output state vector

$$|\tilde{\boldsymbol{a}}\rangle_{out} = \boldsymbol{Q}\,\boldsymbol{V}(z\boldsymbol{I}-\boldsymbol{T})^{-1}\boldsymbol{D}^{-1}\boldsymbol{W}^T\,\boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \tag{5.75}$$

and multi-port impulse response

$$\tilde{\boldsymbol{\mathcal{H}}}_{n_1}(z) = \boldsymbol{QV}(z\boldsymbol{I}-\boldsymbol{T})^{-1}\boldsymbol{D}^{-1}\boldsymbol{W}^T\boldsymbol{R} \in \mathcal{H}^{p_1\times p_2}\,, \tag{5.76}$$

where $\boldsymbol{DT}$ of dimension $n\times n$ is the biorthogonal (or oblique) projection of the general TLM operator $\boldsymbol{\Gamma S}$ onto KRYLOV subspaces.

Utilizing the eigendecomposition of $\boldsymbol{T}$

$$\boldsymbol{T} = \boldsymbol{U\Lambda U}^{-1}\,, \tag{5.77}$$

we can write the multi-port impulse response in the next form

$$\tilde{\boldsymbol{\mathcal{H}}}_{n_2}(z) = \boldsymbol{QVU}(z\boldsymbol{I}-\boldsymbol{\Lambda})^{-1}\boldsymbol{U}^{-1}\boldsymbol{D}^{-1}\boldsymbol{W}^T\boldsymbol{R}. \tag{5.78}$$

Introducing the row vector $\langle\boldsymbol{\alpha}^{(i)}|$ and the column vector $|\boldsymbol{\beta}^{(j)}\rangle$ as follows

$$\langle\boldsymbol{\alpha}^{(i)}| = \langle\boldsymbol{q}_i|\boldsymbol{VU}\,, \tag{5.79}$$

$$|\boldsymbol{\beta}^{(j)}\rangle = \boldsymbol{U}^{-1}\boldsymbol{D}^{-1}\boldsymbol{W}^T|\boldsymbol{r}_j\rangle\,, \tag{5.80}$$

we obtain the pole/residue distribution of the (i,j)-th element of the the transfer function $\tilde{\boldsymbol{\mathcal{H}}}_{n_2}$ (from port j to port i) as it has been given in the form of (5.63).

In a similar way the reduced GIM may be obtained from (5.46) as

$$\boldsymbol{Z}_{n_2}(z) = \boldsymbol{Q_1VU}(z\boldsymbol{I}-\boldsymbol{\Lambda})^{-1}\boldsymbol{U}^{-1}\boldsymbol{D}^{-1}\boldsymbol{W}^T\boldsymbol{R_1}\,. \tag{5.81}$$

A pseudo-code of the TLM-ROM based on the nonsymmetric LANCZOS algorithm is presented below.

Algorithm 5.3.2. The TLM-ROM based on the nonsymmetric LANCZOS algorithm.

(1) choose $|\tilde{\boldsymbol{v}}\rangle = \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}$ and $|\tilde{\boldsymbol{w}}\rangle =$ **random**

(2) $|\boldsymbol{v}_1\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\|\tilde{\boldsymbol{v}}\|_2}$ and $|\boldsymbol{w}_1\rangle = \dfrac{|\tilde{\boldsymbol{w}}\rangle}{\|\tilde{\boldsymbol{w}}\|_2}$

(3) set $|\boldsymbol{v}_0\rangle = \boldsymbol{0}$, $|\boldsymbol{w}_0\rangle = \boldsymbol{0}$ and $\rho_1 = \eta_1 = \delta_0 = 1$

(4) **for** $j = 1, 2, \ldots, n$ **do**

(5) $\quad\delta_j = \langle\boldsymbol{w}_j|\boldsymbol{v}_j\rangle$

(6) $\quad$**if** $\delta_j = 0$ **then stop**

(7) $\quad\alpha_j = \dfrac{\langle\boldsymbol{w}_j|\boldsymbol{\Gamma S}\,|\boldsymbol{v}_j\rangle}{\delta_j}$

(8) $\beta_j = \eta_j \delta_j / \delta_{j-1}$
(9) $\gamma_j = \rho_j \delta_j / \delta_{j-1}$
(10) $|\tilde{\boldsymbol{v}}_{j+1}\rangle = \boldsymbol{\Gamma S}\,|\boldsymbol{v}_j\rangle - \alpha_j |\boldsymbol{v}_j\rangle - \beta_j |\boldsymbol{v}_{j-1}\rangle$
(11) $|\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{S\Gamma}\,|\boldsymbol{w}_j\rangle - \alpha_j |\boldsymbol{w}_j\rangle - \gamma_j |\boldsymbol{w}_{j-1}\rangle$
(12) reorthogonalize if necessary
(13) $\rho_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(14) $\eta_{j+1} = \|\tilde{\boldsymbol{w}}_{j+1}\|_2$
(15) **if** $\rho_{j+1} = 0$ or $\eta_{j+1} = 0$ **then stop**
(16) $\boldsymbol{T}_j = \boldsymbol{U \Lambda}_j \boldsymbol{U}^{-1}$
(17) test for convergence
(18) $|\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\rho_{j+1}}$
(19) $|\boldsymbol{w}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{w}}_{j+1}\rangle}{\eta_{j+1}}$
(20) **end for**
(21) $\boldsymbol{T}_n = \boldsymbol{U \Lambda}_n \boldsymbol{U}^{-1}$
(22) compute impulse response $\tilde{\boldsymbol{\mathcal{H}}}_2$ and/or GIM $\boldsymbol{Z}_{n_2}$
(23) compute $|\tilde{\boldsymbol{a}}\rangle_{out}$

The algorithm 5.3.2 yields the biorthogonal $V - W$ basis for KRYLOV subspaces (see Chapter 4) with start vectors $|\boldsymbol{v}_1\rangle \in \mathcal{H}^N$ and $|\boldsymbol{w}_1\rangle \in \mathcal{H}^N$ and the reduced TLM operator as a triangular matrix $\boldsymbol{T}_n$.

5.3.3 The Scattering-Symmetric Lanczos Process

It has been observed that the nonsymmetric LANCZOS algorithm also simplifies for matrices that are J-HERMITIAN or J-symmetric with respect to inner products induced by general nonsingular, but not necessarily positive definite matrix $\boldsymbol{J}$. There are some special cases that allow use of this J-symmetry, e.g. HAMILTONIAN and TOEPLITZ matrices. However, the determination of a matrix $\boldsymbol{J}$ is in general very complicated and impractical, and the really useful cases of J-matrices are those for which $\boldsymbol{J}$ is known explicitly.

FREUND and ZHA showed in [45], [47] that for J-symmetric matrix $\boldsymbol{A}$,

$$\boldsymbol{A}^T \boldsymbol{J} = \boldsymbol{J A}\,, \tag{5.82}$$

the LANCZOS process can be simplified by choosing the second LANCZOS start vector properly. We recall that the second start vector in the nonsymmetric version is generated randomly.

For most real cases it is very difficult to find any sparse matrix $\boldsymbol{J}$ in any form in order to take advantage of the symmetry like in (5.82).

In this subsection we describe a practical simplified LANCZOS algorithm for TLM, which is derived using the J-symmetry. For this purpose we find a matrix $\boldsymbol{J}$ satisfying eq. (5.82) at first.

Theorem 5.3.1. The TLM-matrix $(\boldsymbol{\Gamma S})$ has a unique property; it is symmetric with respect to the scattering matrix $\boldsymbol{S}$. That is

$$(\boldsymbol{\Gamma S})^T \boldsymbol{S} = \boldsymbol{S}(\boldsymbol{\Gamma S})\,. \tag{5.83}$$

Proof. The right in eq. (5.83) is

$$\boldsymbol{S}(\boldsymbol{\Gamma S}) = \boldsymbol{S\Gamma S}\,, \tag{5.84}$$

and the left in eq. (5.83) is

$$(\boldsymbol{\Gamma S})^T\boldsymbol{S} = (\boldsymbol{S\Gamma})\boldsymbol{S} = \boldsymbol{S\Gamma S}\,. \tag{5.85}$$

□

Let us consider the actual operations involved in the LANCZOS process in some more detail. In order to generate the bi-orthogonal basis vectors, $|\boldsymbol{v}_n\rangle$ and $|\boldsymbol{w}_n\rangle$, for the two KRYLOV subspaces (4.1) and (4.13) the non-symmetric LANCZOS utilizes the sequences (10) and (11) in Algorithm 5.3.2.

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle = (\boldsymbol{\Gamma S})|\boldsymbol{v}_j\rangle - \alpha_n|\boldsymbol{v}_j\rangle - \beta_j|\boldsymbol{v}_{j-1}\rangle, \tag{5.86}$$

$$|\tilde{\boldsymbol{w}}_{j+1}\rangle = (\boldsymbol{S\Gamma})|\boldsymbol{w}_j\rangle - \alpha_j|\boldsymbol{w}_j\rangle - \gamma_j|\boldsymbol{w}_{j-1}\rangle\,. \tag{5.87}$$

We say that the TLM-operator is *scattering-symmetric*, and refer to it briefly as *S-symmetric*. On the basis of this property the second sequence (5.87) in the non-symmetric Lanczos algorithm can be replaced by the following one

$$|\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\,\rangle, \quad j = 1, 2, \ldots, n, \tag{5.88}$$

if we take the second start vector as

$$|\tilde{\boldsymbol{w}}_1\rangle = \boldsymbol{S}\,|\tilde{\boldsymbol{v}}_1\rangle\,. \tag{5.89}$$

In view of (5.88) a simplified version of the Lanczos process results, where only the right basis vectors $|\boldsymbol{v}_1\rangle, |\boldsymbol{v}_2\rangle, \ldots, |\boldsymbol{v}_n\rangle$ need be generated, while the left vectors $|\boldsymbol{w}_1\rangle, |\boldsymbol{w}_2\rangle, \ldots, |\boldsymbol{w}_n\rangle$ are obtained via (5.88). The bi-orthogonal basis is generated to satisfy

$$\boldsymbol{V}_n^T\boldsymbol{S}\,\boldsymbol{V}_n = \boldsymbol{D}_n\,. \tag{5.90}$$

Clearly, the resulting modified Lanczos process requires only one TLM time evolution step and one additional application of the scattering matrix. Hence, the process is faster and consumes less memory in comparison to the conventional non-symmetric Lanczos one.

An important attribute of the proposed modified process is that neither the operator $\boldsymbol{\Gamma S}$ nor its transpose need not be known explicitly, as only its impact on the sequence of Krylov vectors must be taken into account. Thus, the computation of the reduced model can be integrated in the usual iterative TLM process in a very efficient manner.

Below we give a pseudo-code of the TLM-MOR based on the S-symmetric version.

Algorithm 5.3.3. The TLM-MOR based on the S-symmetric LANCZOS algorithm.

(1) choose $|\tilde{\boldsymbol{v}}\rangle = \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}$ and $|\boldsymbol{v}_1\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\|\tilde{\boldsymbol{v}}\,\|_2}$

(2) $|\tilde{\boldsymbol{w}}\rangle = \boldsymbol{S}|\tilde{\boldsymbol{v}}\rangle$ and $|\boldsymbol{w}_1\rangle = \dfrac{|\tilde{\boldsymbol{w}}\rangle}{\|\tilde{\boldsymbol{w}}\,\|_2}$

(3) set $|\boldsymbol{v}_0\rangle = \mathbf{0}$, $|\boldsymbol{w}_0\rangle = \mathbf{0}$ and $\rho_1 = \eta_1 = \delta_0 = 1$
(4) **for** $j = 1, 2, \ldots, n$ **do**
(5) $\delta_j = \langle \boldsymbol{w}_j | \boldsymbol{v}_j \rangle$
(6) **if** $\delta_j = 0$ **then stop**
(7) $\alpha_j = \dfrac{\langle \boldsymbol{w}_j | \boldsymbol{\Gamma S} | \boldsymbol{v}_j \rangle}{\delta_j}$
(8) $\beta_j = \eta_j \delta_j / \delta_{j-1}$
(9) $|\tilde{\boldsymbol{v}}_{j+1}\rangle = \boldsymbol{\Gamma S} |\boldsymbol{v}_j\rangle - \alpha_j |\boldsymbol{v}_j\rangle - \beta_j |\boldsymbol{v}_{j-1}\rangle$
(10) $|\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{S} |\tilde{\boldsymbol{v}}_{j+1}\rangle$
(11) reorthogonalize if necessary
(12) $\rho_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(13) $\eta_{j+1} = \|\tilde{\boldsymbol{w}}_{j+1}\|_2$
(14) **if** $\rho_{j+1} = 0$ or $\eta_{j+1} = 0$ **then stop**
(15) $\boldsymbol{T}_j = \boldsymbol{U} \boldsymbol{\Lambda}_j \boldsymbol{U}^{-1}$
(16) test for convergence
(17) $|\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\rho_{j+1}}$
(18) $|\boldsymbol{w}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{w}}_{j+1}\rangle}{\eta_{j+1}}$
(19) **end for**
(20) $\boldsymbol{T}_n = \boldsymbol{U} \boldsymbol{\Lambda}_n \boldsymbol{U}^{-1}$
(21) compute impulse response $\tilde{\boldsymbol{\mathcal{H}}}_2$ and/or GIM $\boldsymbol{Z}_{n_2}$
(22) compute $|\tilde{\boldsymbol{a}}\rangle_{out}$

5.3.4 Matrix-Padé-Approximations

The concept of transfer function allows the definition of the reduced-order models by means of PADÉ or PADÉ-type approximation in the case of TLM.

Let $z_0 \in \mathcal{C}$ be any point such that z_0 is not a pole of the transfer function $\tilde{\boldsymbol{\mathcal{H}}}$. In practice, the point z_0 is chosen such that it is in some sense close to the frequency band of interest, e.g. in the middle of a frequency range. Then the transfer function $\tilde{\boldsymbol{\mathcal{H}}}$ admits the TAYLOR expansion about z_0 as

$$\tilde{\boldsymbol{\mathcal{H}}}(z) = \boldsymbol{M}_0 + \boldsymbol{M}_1 (z - z_0) + \boldsymbol{M}_2 (z - z_0)^2 + \ldots + \boldsymbol{M}_j (z - z_0)^j + \ldots = \\ = \sum_{j=1}^{\infty} \boldsymbol{M}_j (z')^j \,, \quad (5.91)$$

where $\boldsymbol{M}_j$ are, in general, $p_1 \times p_2$ matrices, which are called the *moments* of $\tilde{\boldsymbol{\mathcal{H}}}$ about z_0, and $z' = z - z_0$ is called a *shift*.

A reduced model (5.53) and (5.70) of the state-space dimension n is called the n-th *Padé model* at the expansion point z_0 of the original system (5.42), if the TEYLOR expansions about z_0 of the transfer function $\tilde{\boldsymbol{\mathcal{H}}}$ in (5.42) of the original system and the transfer functions $\tilde{\boldsymbol{\mathcal{H}}}_{n_1}$ in (5.57) and $\tilde{\boldsymbol{\mathcal{H}}}_{n_2}$ in (5.76) of the reduced-order models agree in as many leading terms as possible

$$\tilde{\boldsymbol{\mathcal{H}}} = \tilde{\boldsymbol{\mathcal{H}}}_n + \mathcal{O}\left((z - z_0)^{\phi(n)} \right) \,, \quad (5.92)$$

where $\phi(n)$ is as large as possible. For introduction to PADÉ approximation we refer the reader to [61].

In order to obtain a matrix PADÉ approximant about a point z_0 we perform the following matrix manipulations

$$\begin{aligned}\tilde{\boldsymbol{\mathcal{H}}}(z') = \boldsymbol{Q}(z\boldsymbol{I} - \boldsymbol{\Gamma S})^{-1}\,\boldsymbol{R} = \boldsymbol{Q}\,(z_0\boldsymbol{I} + (z-z_0)\boldsymbol{I} - \boldsymbol{\Gamma S})^{-1}\,\boldsymbol{R} = \\ = -\boldsymbol{Q}\,(\boldsymbol{\Gamma S} - z_0\boldsymbol{I} - (z-z_0)\boldsymbol{I})^{-1}\,\boldsymbol{R} = \\ = \boldsymbol{Q}\,(\boldsymbol{I} - (z-z_0)\hat{\boldsymbol{A}})^{-1}\,\hat{\boldsymbol{R}} = \\ = \boldsymbol{Q}\,(\boldsymbol{I} - z'\hat{\boldsymbol{A}})^{-1}\,\hat{\boldsymbol{R}} = \sum_{j=1}^{\infty}(z')^j \boldsymbol{Q}\,\hat{\boldsymbol{A}}^{j-1}\,\hat{\boldsymbol{R}}, \qquad (5.93)\end{aligned}$$

where

$$\hat{\boldsymbol{A}} = (\boldsymbol{\Gamma S} - z_0\boldsymbol{I})^{-1}\,, \qquad (5.94)$$

and

$$\hat{\boldsymbol{R}} = -(\boldsymbol{\Gamma S} - z_0\boldsymbol{I})^{-1}\boldsymbol{R}. \qquad (5.95)$$

Unless we have $p_1 = p_2 = 1$, the transfer functions $\boldsymbol{\mathcal{H}}$ and, consequently, $\boldsymbol{\mathcal{H}}_n$ are matrix valued, then the PADÉ and PADÉ-type models are so-called *matrix*-PADÉ and *matrix*-PADÉ-*type approximants* in general.

By comparison of (5.91) and (5.93) we obtain the PADÉ-moments

$$\boldsymbol{M}_j = \boldsymbol{Q}\,\hat{\boldsymbol{A}}^j\,\hat{\boldsymbol{R}}. \qquad (5.96)$$

For simplification we restrict to the single-input single-output case (a SISO system) , i.e $p_1 = p_2 = 1$. Thus, from (5.93) we have a scalar-valued rational transfer function

$$\tilde{\mathcal{H}}(z) = \langle \boldsymbol{q}|\,(\boldsymbol{I} - z'\hat{\boldsymbol{A}})^{-1}\,|\hat{\boldsymbol{r}}\rangle\,. \qquad (5.97)$$

5.3.5 Explicit Moment Matching

Instead of eq. (5.57) and eq. (5.76) we present the scalar-valued rational function $\tilde{\mathcal{H}}_n$ of a reduced SISO-system in terms of these polynomials, or more precisely

$$\tilde{\mathcal{H}}_n(z) = \frac{\varphi_{n-1}(z')}{\psi_n(z')} = \frac{b_{n_1}(z')^{n_1} + \cdots + b_1 z' + b_0}{a_{n_2}(z')^{n_2} + \cdots + a_1 z' + 1}\,. \qquad (5.98)$$

Typically $n_2 = n$ and $n_1 = n_2 - 1 = n - 1$. There are $n_1 + n_2 + 1 = 2n$ free coefficients of the polynomials φ_{n-1} and ψ_n in (5.98). These free parameters can be chosen such that in (5.92) the first $2n$ moments will be matched

$$\tilde{\mathcal{H}} = \tilde{\mathcal{H}}_n + \mathcal{O}\left((z-z_0)^{2n}\right) = \sum_{j=0}^{2n-1} M_j (z-z_0)^j + \mathcal{O}\left((z-z_0)^{2n}\right)\,, \qquad (5.99)$$

where M_j are *moments*, which can be expressed according to (5.96) and (5.97) as following:

$$M_j = \langle \boldsymbol{q}|\,(\hat{\boldsymbol{A}})^j|\hat{\boldsymbol{r}}\rangle\,, \quad j = 0, 1, 2, \ldots, (2n-1). \qquad (5.100)$$

We refer to $\tilde{\mathcal{H}}_n$ at the n-th PADÉ approximant to the impulse response $\tilde{\mathcal{H}}$ in (5.97) . By means of partial fraction expansion we can write $\tilde{\mathcal{H}}_n$ in the form

$$\tilde{\mathcal{H}}_n = c_\infty + \sum_{j=1}^{\hat{n}} \frac{c_j}{z' - p_j}\,. \qquad (5.101)$$

The standard approach to computing $\mathcal{H}_n$ is based on the representation (5.98) and on *explicit moment generation* by (5.100).

- First, the leading $2n$ moments $M_0, M_1, M_2, \ldots, M_{2n-1}$ of $\tilde{\mathcal{H}}$ are computed by (5.100);
- then, the coefficients of the polynomials $\varphi_{n-1}(z')$ and $\psi_n(z')$ are generated by the solving of linear equations with a $n \times n$ HANKEL matrix (a square matrix with constant skew diagonals), whose entries are the moments (5.100);
- the poles p_j in (5.101) are then obtained as the roots of the equation

$$a_n(z')^n + a_{n-1}(z')^{n-1} + \cdots + a_1 z' + 1 = 0\,; \tag{5.102}$$

- the constant c_∞ and the residues c_j in (5.101) are computed by solving another linear system of the order n;

  For more details see [22], [93].

This standard approach to computing $\mathcal{H}_n$ is employed in the *Asymptotic Waveform Evaluation* (AWE) method [22], [93].

However, computing PADÉ approximants using explicit moment computations is numerically unstable, and indeed, this approach can be employed in practice for very moderate values of n, such as $n < 10$ (see [41]). The problem is that the vectors $\hat{\boldsymbol{A}}^j|\hat{\boldsymbol{r}}\rangle$ quickly converge to an eigenvector corresponding to a dominant eigenvalue of $\hat{\boldsymbol{A}}$, and, therefore, the moments M_j computed by (5.100), even for moderate values of j, contain only information about this dominant eigenvalue. Thus, the PADÉ approximant $\mathcal{H}_n$ generated from the moments only contains information about a part of the spectrum of $\hat{\boldsymbol{A}}$, and the computed $\mathcal{H}_n$ does not converge to the transfer function $\mathcal{H}$. To obtain the broadband frequency response, more expansion points are required. The multi-point expansion is used for the fast frequency sweep (FFS). The moment matching technique is not considered in this work.

5.3.6 The Padé-via-Lanczos Algorithm

Together with employing explicit moment matching techniques, the use of the KRYLOV subspace techniques to compute the PADÉ model is possible.

In this subsection we describe the *Padé-via-Lanczos* (PVL) algorithm which exploits the connection between PADÉ approximation and the LANCZOS process to obtain the direct computation of the moments. The connection of the LANCZOS process to the PADÉ approximation is described in detail in [41]. The derivation of PVL showed that the pole/residue representation (5.101) of the PADÉ approximant $\mathcal{H}_n$ can be obtained by running the LANCZOS algorithm and by computing an eigendecomposition of the LANCZOS matrix $\boldsymbol{T}_n$. The resulting computational procedure is the PVL algorithm, the sketch of which is presented below.

Algorithm 5.3.4. The sketch of the TLM-MOR based on the PVL algorithm.

(1) Run n steps of the LANCZOS algorithm (see Algorithm 4.4.1) with the matrix $\hat{\boldsymbol{A}}$ (5.94) to obtain the tridiagonal matrix $\boldsymbol{T}_n$.

(2) Compute an eigenvalue decomposition

$$\boldsymbol{T}_n = \boldsymbol{U}\boldsymbol{\Lambda}_n\boldsymbol{U}^{-1}\,. \tag{5.103}$$

(3) Compute the vectors

$$\langle \boldsymbol{f}| = \boldsymbol{U}^T \langle \boldsymbol{e}_1| \quad \text{and} \quad |\boldsymbol{g}\rangle = \boldsymbol{U}^{-1}|\boldsymbol{e}_1\rangle\,. \tag{5.104}$$

(4) Compute the poles and residues of $\mathcal{H}_n$ (5.101) by setting

$$p_j = 1/\lambda_j\,, \tag{5.105}$$

$$c_j = \langle \boldsymbol{q}|\hat{\boldsymbol{r}}\rangle \frac{f_j g_j}{\lambda_j} \quad j = 1, 2, \ldots, n \quad \text{with} \quad \lambda_j \neq 0\,, \tag{5.106}$$

$$c_\infty = \sum_{j=0}^{n} f_j g_j \, \langle \boldsymbol{q}|\hat{\boldsymbol{r}}\rangle \quad \text{with} \quad \lambda_j = 0\,. \tag{5.107}$$

We note that the moments (5.100) in the PVL method are computed using the biorthogonal basis $V - W$ for right and left n-th KRYLOV subspaces

$$\mathcal{K}(|\hat{\boldsymbol{r}}\rangle, \hat{\boldsymbol{A}}, n) = \text{span}\,\{\,|\hat{\boldsymbol{r}}\rangle,\, \hat{\boldsymbol{A}}|\hat{\boldsymbol{r}}\rangle,\, \hat{\boldsymbol{A}}^2|\hat{\boldsymbol{r}}\rangle,\, \ldots,\, \hat{\boldsymbol{A}}^{n-1}|\hat{\boldsymbol{r}}\rangle\,\} \tag{5.108}$$

and

$$\mathcal{K}(|\boldsymbol{q}\rangle, \hat{\boldsymbol{A}}^T, n) = \text{span}\,\{\,|\boldsymbol{q}\rangle,\, \hat{\boldsymbol{A}}^T|\boldsymbol{q}\rangle,\, (\hat{\boldsymbol{A}}^T)^2|\boldsymbol{q}\rangle,\, \ldots,\, (\hat{\boldsymbol{A}}^T)^{n-1}|\boldsymbol{q}\rangle\,\} \tag{5.109}$$

respectively. Using the basis vectors $|\boldsymbol{v}_j\rangle$ and $|\boldsymbol{w}_j\rangle$, the explicit moment computations are avoided, and instead of moments (5.100) we now calculate modified moments

$$\langle \boldsymbol{w}_j|\boldsymbol{v}_j\rangle \quad \text{and} \quad \langle \boldsymbol{w}_j|\hat{\boldsymbol{A}}|\boldsymbol{v}_j\rangle,\ j = 1, 2, \ldots, n\,. \tag{5.110}$$

The modified moments (5.110) contain the same information as the moments (5.100), and for each $j = 0, 1, 2, \ldots, (2n-1)$, the j-th moment can be expressed as a suitable combination of the numbers (5.110).

While PADÉ models often provide very good approximations, in general they do not preserve the stability and passivity of the original system. However, a reduced-order model can be generated to be stable and passive. In [41] it has been shown, that the PADÉ approximation via a LANCZOS process produces more accurate and higher order approximants and has significantly superior numerical stability than explicit moment matching, i.e. AWE.

In a similar way the *Padé-via-Arnoldi* (PVA) algorithm can be derived.

5.3.7 The Shift-Inverse Approach

In order to extract eigenvalues of the general TLM operator, which are closed to a point z_0 in the complex plain (Fig. 5.1), at first we involve the *sift-inverse approach.* (see Chapter 4).

For this purpose we generate the KRYLOV subspaces by means of a matrix $\hat{\boldsymbol{A}}$, which has been defined by (5.94), or in other words, a shift-inverse general TLM-operator is utilized to generate a biorthogonal (and/or an orthogonal) basis for KRYLOV subspaces. Moreover, we prove the possibility to use the S-symmetric version of the LANCZOS methods. We would like to point out, that the shift-inverse method via Arnoldi process can be derived in a similar way.

Theorem 5.3.2. The shift-inverse matrix $(\boldsymbol{\Gamma S} - z_0 \boldsymbol{I})^{-1}$ has a unique property; it is symmetric with respect to the scattering matrix $\boldsymbol{S}$. That is

$$\left((\boldsymbol{\Gamma S} - z_0 \boldsymbol{I})^{-1}\right)^T \boldsymbol{S} = \boldsymbol{S}\,(\boldsymbol{\Gamma S} - z_0 \boldsymbol{I}))^{-1}\,. \tag{5.111}$$

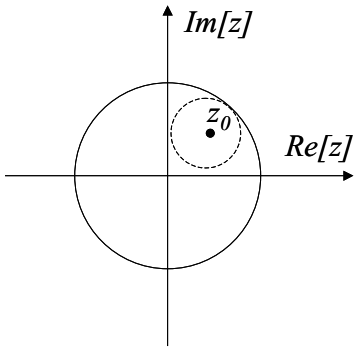


Figure 5.1: Expansion point z_0 in the complex Z-plane.

Proof. The right in eq. (5.111) is

$$\boldsymbol{S}\ (\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I}))^{-1} = \left((\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})\boldsymbol{S}^{-1}\right)^{-1} = \\ = \left((\boldsymbol{\Gamma} - z_0\boldsymbol{S}^{-1})\right)^{-1}, \tag{5.112}$$

and the left in eq. (5.111) is

$$\left((\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})^{-1}\right)^T \boldsymbol{S} = \left(\boldsymbol{S}\,(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})^{-1}\right)^T. \tag{5.113}$$

Using the relation (5.112) we derive that

$$\left((\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})^{-1}\right)^T \boldsymbol{S} = \left(\left((\boldsymbol{\Gamma} - z_0\boldsymbol{S}^{-1}\right)^{-1}\right)^T = \\ = \left(\left((\boldsymbol{\Gamma} - z_0\boldsymbol{S}^{-1}\right)^T\right)^{-1} = \left((\boldsymbol{\Gamma} - z_0\boldsymbol{S}^{-1})\right)^{-1}. \tag{5.114}$$

We note that the last equation in (5.114) is obtained using the symmetry of $\boldsymbol{S}$ and $\boldsymbol{\Gamma}$ operators. □

The theorem 5.3.2 allows the S-symmetric LANCZOS algorithm (Algorithm 5.3.3) to be involved in the sift-inverse approach that can save half of work. The pseudo-code of the sift-inverse based MOR technique is given below.

Algorithm 5.3.5. The TLM-MOR based on the shift-inverse approach using the S-symmetric LANCZOS algorithm.

(1) choose $|\tilde{\boldsymbol{v}}\rangle = \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}$ and $|\boldsymbol{v}_1\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\|\tilde{\boldsymbol{v}}\|_2}$

(2) $|\tilde{\boldsymbol{w}}\rangle = \boldsymbol{S}|\tilde{\boldsymbol{v}}\rangle$ and $|\boldsymbol{w}_1\rangle = \dfrac{|\tilde{\boldsymbol{w}}\rangle}{\|\tilde{\boldsymbol{w}}\|_2}$

(3) set $|\boldsymbol{v}_0\rangle = \boldsymbol{0}$, $|\boldsymbol{w}_0\rangle = \boldsymbol{0}$ and $\rho_1 = \eta_1 = \delta_0 = 1$

(4) **for** $j = 1, 2, \ldots, n$ **do**

(5) $\quad \delta_j = \langle \boldsymbol{w}_j | \boldsymbol{v}_j \rangle$

(6) $\quad$ **if** $\delta_j = 0$ **then stop**

(7) $\quad$ solve $(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\boldsymbol{v}_j\rangle$ for $|\tilde{\boldsymbol{v}}_{j+1}\rangle$

(8) $\quad \alpha_j = \dfrac{\langle \boldsymbol{w}_j | \tilde{\boldsymbol{v}}_{j+1} \rangle}{\delta_j}$

(9) $\beta_j = \eta_j \delta_j / \delta_{j-1}$
(10) $|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\tilde{\boldsymbol{v}}_{j+1}\rangle - \alpha_j |\boldsymbol{v}_j\rangle - \beta_j |\boldsymbol{v}_{j-1}\rangle$
(11) $|\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle$
(12) rebiorthogonalize if necessary
(13) $\rho_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(14) $\eta_{j+1} = \|\tilde{\boldsymbol{w}}_{j+1}\|_2$
(15) **if** $\rho_{j+1} = 0$ or $\eta_{j+1} = 0$ **then stop**
(16) $\tilde{\boldsymbol{T}}_j = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}_j\boldsymbol{U}^{-1}$
(17) test for convergence
(18) $|\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\rho_{j+1}}$
(19) $|\boldsymbol{w}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{w}}_{j+1}\rangle}{\eta_{j+1}}$
(20) **end for**
(21) $\tilde{\boldsymbol{T}}_n = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}_n\boldsymbol{U}^{-1}$
(22) $\boldsymbol{\Lambda}_n = z_0 + (\tilde{\boldsymbol{\Lambda}}_n)^{-1}$
(23) compute impulse response $\tilde{\boldsymbol{\mathcal{H}}}_2$ and/or GIM $\boldsymbol{Z}_{n_2}$
(24) compute $|\tilde{\boldsymbol{a}}\rangle_{out}$

As we can see, the main problem in the Algorithm 5.3.5 is to solve the equation

$$(\boldsymbol{\Gamma S} - z_0\boldsymbol{I})|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\boldsymbol{v}_j\rangle \tag{5.115}$$

for $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ at each step j. We recall that the shift of the general TLM-operator $(\boldsymbol{\Gamma S} - z_0\boldsymbol{I})$ is not given in an explicit form. There are two general numerical solutions: conventional iterative (we rename them to iterative methods) and projection iterative (we rename them to projection methods).

We note, if we know how to operate efficiently with the shift-inverse operator $\hat{\boldsymbol{A}}$ (defined by (5.94)) without it forming we will be able to proceed the direct application of LANCZOS and ARNOLDI algorithms, and we can exchange the sequence (7) in the Algorithm 5.3.5 for

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle = \hat{\boldsymbol{A}}|\boldsymbol{v}_j\rangle\,. \tag{5.116}$$

However, first, the inversion of the explicit operator $(\boldsymbol{\Gamma S} - z_0\boldsymbol{I})$ (if we had known $\boldsymbol{\Gamma S}$ in explicit form) requires a high computational effort due to its large dimension. Second, it seems to be nearly impossible to gain $\hat{\boldsymbol{A}}$ in an implicit form trough any considerations.

The same problem (how to obtain $\hat{\boldsymbol{A}}$ and how to operate with it) we have by the use of the PVL and PVA algorithms, because PADÉ approximations and shift-inverse approaches are equivalent, that follows by direct comparing the Algorithms 5.3.4 and 5.3.5. Both of them approximate an original system about the point z_0 and both of them deal with the shift inverse TLM matrix. In order to obtain the next LANCZOS vector $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ an impact of shift-inverse $\hat{\boldsymbol{A}}$ onto the vector $|\boldsymbol{v}_j\rangle$ in (5.116) can be replaced by solution of (5.115) for $|\tilde{\boldsymbol{v}}_{j+1}\rangle$.

5.3.8 Iterative Solvers for the Shift-Inverse Approach

In this subsection we describe one of the simplest stationary iterative solvers for linear equation systems in the SI-approach.

For convenience we write the system (5.115) to be solved as follows

$$(\boldsymbol{\Gamma S} - z_0 \boldsymbol{I})|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\boldsymbol{v}_j\rangle \tag{5.117}$$

where $|\boldsymbol{v}_j\rangle$ is known state vector, and $|\boldsymbol{v}_{j+1}\rangle$ is a state vector to be computed by an iterative way.

For this purpose we express the equation (5.117) in a special form

$$\boldsymbol{x}^{(k)} = \boldsymbol{\alpha}\boldsymbol{x}^{(k-1)} + \boldsymbol{\beta}\,,\ k = 1, 2, \ldots . \tag{5.118}$$

Through straightforward manipulations in (5.117) we have

$$(\boldsymbol{I} - \frac{1}{z_0}\,\boldsymbol{\Gamma S})|\tilde{\boldsymbol{v}}_{j+1}\rangle = -\frac{1}{z_0}\,|\boldsymbol{v}_j\rangle \tag{5.119}$$

and, therefore, we obtain the stationary iterative scheme

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle^{(k)} = \frac{1}{z_0}\,\boldsymbol{\Gamma S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle^{(k-1)} - \frac{1}{z_0}\,|\boldsymbol{v}_j\rangle\,. \tag{5.120}$$

We note, the scheme (5.120) is called stationary, because neither the TLM operator $\boldsymbol{\Gamma S}$ nor $|\boldsymbol{v}_j\rangle$ depend upon the iteration count k (i.e. the iterative scheme is kept at each iteration k).

Proposition 5.3.1. The iterative scheme (5.120) is a convergent scheme if, only if, an expansion point z_0 is located outside the complex unit circle.

Proof. It was shown in [34], that the iterative scheme given in (5.118) is a convergent scheme only if a norm of the matrix $\boldsymbol{\alpha}$ is less than one, or more precisely,

$$\|\boldsymbol{\alpha}\| < 1\,. \tag{5.121}$$

By comparing (5.118) and (5.120), we see that

$$\boldsymbol{\alpha} = \frac{1}{z_0}\,\boldsymbol{\Gamma S} \quad \text{and} \quad \boldsymbol{\beta} = -\frac{1}{z_0}\,|\boldsymbol{v}_j\rangle\,, \tag{5.122}$$

and from (5.121) we draw a conclusion that the iterative scheme (5.120) is a convergent scheme only if

$$\left\| \frac{1}{z_0}\,\boldsymbol{\Gamma S} \right\| < 1\,. \tag{5.123}$$

Since $\|\boldsymbol{\Gamma S}\|_2 \leq 1$, we require that

$$\left| \frac{1}{z_0} \right| < 1\,, \tag{5.124}$$

and, as follows,

$$|z_0| > 1\,. \tag{5.125}$$

□

According to (5.120) we start to generate approximations with $|\tilde{\boldsymbol{v}}_{j+1}\rangle^{(0)} = |\boldsymbol{v}_j\rangle$

$$\begin{aligned} |\tilde{\boldsymbol{v}}_{j+1}\rangle^{(1)} &= \frac{1}{z_0}\boldsymbol{\Gamma}\boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle^{(0)} - \frac{1}{z_0}\,|\boldsymbol{v}_j\rangle\ , \\ |\tilde{\boldsymbol{v}}_{j+1}\rangle^{(2)} &= \frac{1}{z_0}\boldsymbol{\Gamma}\boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle^{(1)} - \frac{1}{z_0}\,|\boldsymbol{v}_j\rangle\ , \\ \cdots \quad &\cdots \quad \cdots \\ |\tilde{\boldsymbol{v}}_{j+1}\rangle^{(k)} &= \frac{1}{z_0}\boldsymbol{\Gamma}\boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle^{(k-1)} - \frac{1}{z_0}\,|\boldsymbol{v}_j\rangle\,, \end{aligned} \tag{5.126}$$

until

$$\left\| |\tilde{\boldsymbol{v}}_{j+1}\rangle^{(k)} - |\tilde{\boldsymbol{v}}_{j+1}\rangle^{(k-1)} \right\|_2 > \epsilon\,. \tag{5.127}$$

We note that one of the more complicated and, eventually, better converged stationary iterative methods, e.g. the JACOBI method, the GAUSS-SEIDEL method, the Successive Overrelaxation method and the Symmetric Successive Overrelaxation method, can be involved to solve (5.117). For their detailed descriptions and implementations the reader is referred to [8], [34]. One other possibility is the use of iterative projection methods.

5.3.9 Projection Solvers for the Shift-Inverse Approach

Most of the existing practical iterative techniques for solving large linear systems of equations utilize a projection process, which represent a canonical way for extracting an approximation of the solution of a linear system from a subspace.

Now, we need to solve the linear system (5.120) as well . If all of the right hand sides in (5.120) were available simultaneously, the block LANCZOS methods could be used. However, in the case of TLM the right hand sides are not available at the same time, since a right-side vector $|\boldsymbol{v}_j\rangle$ depends on the solutions of $|\boldsymbol{v}_i\rangle$, $i = 1, \ldots, j-1$, of the previous linear systems. PARLETT [88] proposed to use the LANCZOS algorithms for solving the first system and to save the LANCZOS basis in order to provide good approximate solutions for the subsequent systems. An approximate solution for the second system (and for subsequent systems also) can be obtained by using projection techniques onto KRYLOV subspaces generated during the solving of the first linear system. We note that an *orthogonal* and *oblique* projections are possible.

Let $|\tilde{\boldsymbol{v}}_{j+1}\rangle_0$ be an initial guess for the solution of the initial system

$$(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\boldsymbol{v}_j\rangle\,, \tag{5.128}$$

and $|\boldsymbol{r}\rangle_0$ is the initial residual vector

$$|\boldsymbol{r}\rangle_0 = |\boldsymbol{v}_j\rangle - \boldsymbol{\Gamma}\boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle_0\,. \tag{5.129}$$

Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be two n-dimensional subspaces in $\mathcal{H}^N$. A goal of a projection technique onto the subspace $\mathcal{K}_1$ and orthogonal to $\mathcal{K}_2$ is to find a approximate solution $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ by imposing the conditions that $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ belongs to $\mathcal{K}_1$ and that the new residual vector be orthogonal to $\mathcal{K}_2$. That is

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle \in |\tilde{\boldsymbol{v}}_{j+1}\rangle_0 + \mathcal{K}_1, \tag{5.130}$$

$$|\boldsymbol{v}_j\rangle - \boldsymbol{\Gamma}\boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle \perp \mathcal{K}_2\,. \tag{5.131}$$

Orthogonal projection methods correspond to the case when the two subspaces $\mathcal{K}_1$ and $\mathcal{K}_2$ are identical, otherwise the projection is called *oblique*.

Let $\boldsymbol{V}$ be an $N \times n$ matrix whose column vectors form a basis for $\mathcal{K}_1$ and, similarly, $\boldsymbol{W}$ be an $N \times n$ matrix whose column vectors form a basis for $\mathcal{K}_2$, so that

$$\boldsymbol{W}^T \boldsymbol{V} = \boldsymbol{D} \,. \tag{5.132}$$

Then, according to (5.130) the approximate solution can be written as

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\tilde{\boldsymbol{v}}_{j+1}\rangle_0 + \boldsymbol{V}|\hat{\boldsymbol{v}}\rangle \,, \;\; |\hat{\boldsymbol{v}}\rangle \in \mathcal{K}_1 \,. \tag{5.133}$$

For the variable $|\hat{\boldsymbol{v}}\rangle$ the condition (5.131) translates into the next condition

$$\boldsymbol{W}^T(|\boldsymbol{r}\rangle_0 - \boldsymbol{\Gamma}\boldsymbol{S}\,\boldsymbol{V}|\hat{\boldsymbol{v}}\rangle) = \boldsymbol{0} \,, \tag{5.134}$$

from (5.134) we obtain

$$|\hat{\boldsymbol{v}}\rangle = (\boldsymbol{W}^T \boldsymbol{\Gamma}\boldsymbol{S}\,\boldsymbol{V})^{-1} \boldsymbol{W}^T |\boldsymbol{r}\rangle_0 \,. \tag{5.135}$$

From (5.133) and (5.135) the following expression for the approximate solution $|\tilde{\boldsymbol{v}}_{j+1}\rangle$ results in

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\tilde{\boldsymbol{v}}_{j+1}\rangle_0 + \boldsymbol{V}(\boldsymbol{W}^T \boldsymbol{\Gamma}\boldsymbol{S}\,\boldsymbol{V})^{-1} \boldsymbol{W}^T |\boldsymbol{r}\rangle_0 \,. \tag{5.136}$$

Using the similarity transformation given by the LANCZOS algorithm (5.72) we obtain

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle = |\tilde{\boldsymbol{v}}_{j+1}\rangle_0 + \boldsymbol{V}\boldsymbol{T}^{-1}\boldsymbol{D}^{-1}\boldsymbol{W}^T |\boldsymbol{r}\rangle_0 \,. \tag{5.137}$$

An iterative projection scheme for solving (5.128) belongs to KRYLOV subspace methods if, for any choice of $|\tilde{\boldsymbol{v}}_{j+1}\rangle$, it produce approximate solutions of the form

$$|\tilde{\boldsymbol{v}}_{j+1}\rangle \in |\tilde{\boldsymbol{v}}_{j+1}\rangle_0 + \mathcal{K}(|\boldsymbol{r}\rangle_0, \boldsymbol{\Gamma}\boldsymbol{S}, n) \,, \;\; j = 1, 2, \ldots, n. \tag{5.138}$$

Thus, the design of a KRYLOV subspace algorithm consists of two main parts: the construction of suitable vectors for the KRYLOV subspaces $\mathcal{K}(|\boldsymbol{r}\rangle_0, \boldsymbol{\Gamma}\boldsymbol{S}, n)$ and choice of the actual iterates $|\tilde{\boldsymbol{v}}_{j+1}\rangle$. The Quasi-Minimal Residual (QMR) method is an example of a KRYLOV subspace iteration where the basis vectors are generated by means of the nonsymmetric LANCZOS method, and the iterates are characterized by a quasi minimal residual properties.

The QMR method

When the three-term LANCZOS method is used, the n-th QMR iterate $|\tilde{\boldsymbol{v}}_n\rangle$ is defined as follows

$$|\tilde{\boldsymbol{v}}_n\rangle = |\tilde{\boldsymbol{v}}_n\rangle_0 + \boldsymbol{V}|\hat{\boldsymbol{v}}_n\rangle \,, \tag{5.139}$$

where $|\hat{\boldsymbol{v}}_n\rangle$ is the unique solution of the least-square problem

$$\| \, |\boldsymbol{f}_{n+1}\rangle - \boldsymbol{\chi}_{n+1}\boldsymbol{T}_n |\hat{\boldsymbol{v}}_n\rangle \|_2 = \min \| \, |\boldsymbol{f}_{n+1}\rangle - \boldsymbol{\chi}_{n+1}\boldsymbol{T}_n |\hat{\boldsymbol{v}}_n\rangle \|_2 \,. \tag{5.140}$$

Here

$$\boldsymbol{\chi}_{n+1} = \mathrm{diag}(\|\tilde{\boldsymbol{v}}_1\|_2, \|\tilde{\boldsymbol{v}}_2\|_2, \ldots, \|\tilde{\boldsymbol{v}}_{n+1}\|_2) \tag{5.141}$$

and

$$|\boldsymbol{f}_{n+1}\rangle = [\, |\boldsymbol{r}\rangle_0, 0, \ldots, 0\,]^T \in \mathcal{H}^{n+1} \,. \tag{5.142}$$

The residual vector $|\boldsymbol{r}\rangle_n$ satisfies the relation

$$|\boldsymbol{r}\rangle_n = \left(\boldsymbol{V}_{n+1}\boldsymbol{\chi}_{n+1}^{-1}\right)\left(|\boldsymbol{f}_{n+1}\rangle - \boldsymbol{\chi}_{n+1}\boldsymbol{T}_n|\hat{\boldsymbol{v}}_n\rangle\right) . \tag{5.143}$$

Thus the choice of $|\hat{\boldsymbol{v}}_n\rangle$ means a quasi-minimization of the residual norm, i.e. the second factor in (5.143) is minimal.

QMR-from-BCG

The QMR-from-BCG algorithm was first proposed by FREUND and SZETO [46], and is the simplest possible implementation of the QMR method. The algorithm uses two-term recurrences (see Chapter 4) of the classical *biconjugate gradient* (BCG) algorithm to generate the vectors $|\tilde{\boldsymbol{v}}_j\rangle$ and $|\tilde{\boldsymbol{p}}_j\rangle$. By adding updates for the QMR iterates and the QMR search directions and some minor scalar operation to these BCG recursions, we obtain an algorithm that is only marginally more expensive than BCG [8],[103]. In the case of TLM the proposed in [46] QMR-from-BCG algorithm may be simplified due to the S-symmetry of the general TLM operator. In case the system

$$(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})|\boldsymbol{v}_2\rangle = |\boldsymbol{v}_1\rangle \tag{5.144}$$

to be solved for $|\boldsymbol{v}_2\rangle$ using the Algorithm 5.3.5, the simplified QMR-from-BCG algorithm takes on the following form:

Algorithm 5.3.6. The S-symmetric QMR-from-BCG algorithm.

(1) choose $|\boldsymbol{v}_2\rangle_0 = |\boldsymbol{v}_1\rangle$
(2) $|\tilde{\boldsymbol{v}}_1\rangle = |\boldsymbol{v}_1\rangle - (\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})\,|\boldsymbol{v}_2\rangle_0$
(3) $|\tilde{\boldsymbol{p}}_1\rangle = |\tilde{\boldsymbol{v}}_1\rangle$ and $|\tilde{\boldsymbol{q}}_1\rangle = \boldsymbol{S}\,|\tilde{\boldsymbol{p}}_1\rangle$
(4) $\tau_1 = \langle\tilde{\boldsymbol{v}}_1|\tilde{\boldsymbol{v}}_1\rangle$ and $\rho_1 = \langle\tilde{\boldsymbol{q}}_1|\tilde{\boldsymbol{v}}_1\rangle$
(5) set $|\boldsymbol{d}_0\rangle = \boldsymbol{0}$ and $\vartheta_0 = 0$
(6) **for** $j = 1, 2, \ldots, n$ **do**
(7) $\quad|\boldsymbol{t}_j\rangle = (\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})|\tilde{\boldsymbol{p}}_j\rangle$
(8) $\quad\sigma_j = \langle\tilde{\boldsymbol{q}}_j|\boldsymbol{t}_j\rangle$
(9) $\quad$**if** $\sigma_j = 0$ **then stop**
(10) $\quad\zeta_n = \dfrac{\xi_n}{\sigma_n}$
(11) $\quad|\tilde{\boldsymbol{v}}_{n+1}\rangle = |\tilde{\boldsymbol{v}}_n\rangle - \zeta_n|\boldsymbol{t}_j\rangle$
(12) $\quad\vartheta_n = \dfrac{\langle\tilde{\boldsymbol{v}}_n|\tilde{\boldsymbol{v}}_n\rangle}{\tau_n}$, $\psi_n = \dfrac{1}{1+\vartheta_n}$ and $\tau_{n+1} = \tau_n\vartheta_n\psi_n$
(13) $\quad|\boldsymbol{d}_j\rangle = \psi_n\vartheta_{n-1}|\boldsymbol{d}_{j-1}\rangle + \psi_n\alpha_n|\tilde{\boldsymbol{p}}_j\rangle$
(14) $\quad|\boldsymbol{v}_2\rangle_j = |\boldsymbol{v}_2\rangle_{j-1} + |\boldsymbol{d}_j\rangle$
(15) $\quad$**if** $|\boldsymbol{v}_2\rangle_j$ has converged **then stop**
(16) $\quad$**if** $\xi_n = 0$ **then stop**
(17) $\quad|\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{S}\,|\tilde{\boldsymbol{w}}_j\rangle$
(18) $\quad\xi_{j+1} = \langle\tilde{\boldsymbol{w}}_{j+1}|\tilde{\boldsymbol{v}}_{j+1}\rangle$ and $\iota_j = \dfrac{\xi_{j+1}}{\xi_j}$
(19) $\quad\tilde{\boldsymbol{p}}_{j+1} = \tilde{\boldsymbol{v}}_{j+1} + \iota_j\tilde{\boldsymbol{p}}_j$
(20) $\quad\tilde{\boldsymbol{q}}_{j+1} = \tilde{\boldsymbol{w}}_{j+1} + \iota_j\tilde{\boldsymbol{q}}_j$
(21) **end for**

The algorithm yields the solution of the first linear system (5.144) and the biorthogonal basis $\tilde{\boldsymbol{V}}$ and $\tilde{\boldsymbol{W}}$, and the tridiagonal matrix $\tilde{\boldsymbol{T}}$

$$\tilde{\boldsymbol{T}} = \boldsymbol{L}\boldsymbol{U}\,, \tag{5.145}$$

where

$$\boldsymbol{L}_{n\times n} = \begin{bmatrix} \zeta_1^{-1} & 0 & & & \\ -\zeta_1^{-1} & \zeta_2^{-1} & 0 & & \\ & -\zeta_2^{-1} & \zeta_3^{-1} & \ddots & \\ & & \ddots & \ddots & 0 \\ & & & -\zeta_{n-1}^{-1} & \zeta_n^{-1} \end{bmatrix}, \tag{5.146}$$

and

$$\boldsymbol{U}_{n\times n} = \begin{bmatrix} 1 & -\iota_2 & & & \\ 0 & 1 & -\iota_3 & & \\ & 0 & 1 & \ddots & \\ & & \ddots & \ddots & -\iota_n \\ & & & 0 & 1 \end{bmatrix}. \tag{5.147}$$

Now, we can use the information gathered during the solution of the first system in order to provide an approximation of the second system (the second iteration in the Algorithm 5.3.5)

$$(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})|\boldsymbol{v}_3\rangle = |\boldsymbol{v}_2\rangle\,. \tag{5.148}$$

For this purpose, the basis $\tilde{\boldsymbol{V}}$ and $\tilde{\boldsymbol{W}}$ and the tridiagonal matrix $\tilde{\boldsymbol{T}}$ must be saved. Supposing the initial guess for $|\boldsymbol{v}_3\rangle$ to be $|\boldsymbol{v}_3\rangle_0 = |\boldsymbol{v}_2\rangle$ we can write the corresponding residual vector as

$$|\boldsymbol{r}_2\rangle_0 = |\boldsymbol{v}_2\rangle - (\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})\,|\boldsymbol{v}_3\rangle_0\,. \tag{5.149}$$

A natural way of improving the $|\boldsymbol{v}_3\rangle_0$ is the use of a projection onto KRYLOV subspace $\mathcal{K}\left(|\boldsymbol{r}_1\rangle_0, (\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})\,, n\right)$ generated for the solution of the first system. According to eq. (5.137) we obtain the solution of the second system as

$$|\boldsymbol{v}_3\rangle = |\boldsymbol{v}_3\rangle_0 + \tilde{\boldsymbol{V}}\tilde{\boldsymbol{T}}^{-1}\tilde{\boldsymbol{D}}^{-1}\tilde{\boldsymbol{W}}^T|\boldsymbol{r}_2\rangle_0\,, \tag{5.150}$$

where

$$\tilde{\boldsymbol{W}}^T(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})\,\tilde{\boldsymbol{V}} = \tilde{\boldsymbol{D}}\tilde{\boldsymbol{T}}\,. \tag{5.151}$$

And the solution of the third system (the third iteration in the Algorithm 5.3.5)

$$(\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})|\boldsymbol{v}_4\rangle = |\boldsymbol{v}_3\rangle \tag{5.152}$$

is approximated by the next expression

$$|\boldsymbol{v}_4\rangle = |\boldsymbol{v}_4\rangle_0 + \tilde{\boldsymbol{V}}\tilde{\boldsymbol{T}}^{-1}\tilde{\boldsymbol{D}}^{-1}\tilde{\boldsymbol{W}}^T|\boldsymbol{r}_3\rangle_0\,, \tag{5.153}$$

where $|\boldsymbol{v}_4\rangle_0 = |\boldsymbol{v}_3\rangle$ and the $|\boldsymbol{r}_3\rangle_0$ is the third residual vector given by

$$|\boldsymbol{r}_3\rangle_0 = |\boldsymbol{v}_3\rangle - (\boldsymbol{\Gamma}\boldsymbol{S} - z_0\boldsymbol{I})\,|\boldsymbol{v}_4\rangle_0\,. \tag{5.154}$$

The goal of this projection procedure is that only the first system (5.144) must be solved iteratively by some iterative method (e.g. QMR, QMR-from-BCG), and the second, third, and following linear systems are solved by the projection onto KRYLOV subspace generated during the solution of the first system. As we can see, the operator $\boldsymbol{\Gamma S} - z_0 \boldsymbol{I}$ is involved only when computing the residuals $|\boldsymbol{r}_2\rangle_0, |\boldsymbol{r}_3\rangle_0, \ldots, |\boldsymbol{r}_n\rangle_0$. Thus, the use of iterative projection techniques is expected to be less time consuming than the utilization of conventional iterative algorithms; however, it consumes much more memory, because the full biorthogonal basis $\tilde{\boldsymbol{V}}$ and $\tilde{\boldsymbol{W}}$ as well as the reduced matrix $\tilde{\boldsymbol{T}}$ must be stored. An important question regarding the approximation accuracy of the projection process is examined in [102].

The sketch of TLM-MOR based on the shift-inverse approach and iterative projection techniques is given below.

Algorithm 5.3.7. TLM-ROM based on the sift-inverse approach using the S-symmetric LANCZOS algorithm and iterative projection techniques.

(1) choose $|\tilde{\boldsymbol{v}}\rangle = \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}$ and $|\boldsymbol{v}_1\rangle = \dfrac{|\tilde{\boldsymbol{v}}\rangle}{\|\tilde{\boldsymbol{v}}\|_2}$
(2) $|\tilde{\boldsymbol{w}}\rangle = \boldsymbol{S}|\tilde{\boldsymbol{v}}\rangle$ and $|\boldsymbol{w}_1\rangle = \dfrac{|\tilde{\boldsymbol{w}}\rangle}{\|\tilde{\boldsymbol{w}}\|_2}$
(3) set $|\boldsymbol{v}_0\rangle = \boldsymbol{0}$, $|\boldsymbol{w}_0\rangle = \boldsymbol{0}$ and $\rho_1 = \eta_1 = \delta_0 = 1$
(4) **for** $j = 1, 2, \ldots, n$ **do**
(5) $\quad \delta_j = \langle \boldsymbol{w}_j | \boldsymbol{v}_j \rangle$
(6) $\quad$ **if** $\delta_j = 0$ **then stop**
(7) $\quad$ **if** j = 1 **then do**
(8) $\quad\quad$ solve eq. (5.144) by QMR-from-BCG Algorithm 5.3.6
(9) $\quad\quad$ store $\tilde{\boldsymbol{V}}$, $\tilde{\boldsymbol{W}}$ and $\tilde{\boldsymbol{T}}$
(10) $\quad$ **else**
(11) $\quad\quad |\boldsymbol{r}_j\rangle_0 = |\boldsymbol{v}_j\rangle - (\boldsymbol{\Gamma S} - z_0 \boldsymbol{I})\,|\boldsymbol{v}_j\rangle$
(12) $\quad\quad |\tilde{\boldsymbol{v}}_{j+1}\rangle = |\boldsymbol{v}_j\rangle + \tilde{\boldsymbol{V}}\tilde{\boldsymbol{T}}^{-1}\tilde{\boldsymbol{D}}^{-1}\tilde{\boldsymbol{W}}^T |\boldsymbol{r}_j\rangle_0$
(13) $\quad$ **end if**
(14) $\quad \alpha_j = \dfrac{\langle \boldsymbol{w}_j | \tilde{\boldsymbol{v}} \rangle}{\delta_j}$
(15) $\quad \beta_j = \eta_j \delta_j / \delta_{j-1}$
(16) $\quad |\tilde{\boldsymbol{v}}_{j+1}\rangle = |\tilde{\boldsymbol{v}}_{j+1}\rangle - \alpha_j |\boldsymbol{v}_j\rangle - \beta_j |\boldsymbol{v}_{j-1}\rangle$
(17) $\quad |\tilde{\boldsymbol{w}}_{j+1}\rangle = \boldsymbol{S}\,|\tilde{\boldsymbol{v}}_{j+1}\rangle$
(18) $\quad$ reorthogonalize if necessary
(19) $\quad \rho_{j+1} = \|\tilde{\boldsymbol{v}}_{j+1}\|_2$
(20) $\quad \eta_{j+1} = \|\tilde{\boldsymbol{w}}_{j+1}\|_2$
(21) $\quad$ **if** $\rho_{j+1} = 0$ or $\eta_{j+1} = 0$ **then stop**
(22) $\quad \tilde{\boldsymbol{T}}_j = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}_j \boldsymbol{U}^{-1}$
(23) $\quad$ test for convergence
(24) $\quad |\boldsymbol{v}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{v}}_{j+1}\rangle}{\rho_{j+1}}$
(25) $\quad |\boldsymbol{w}_{j+1}\rangle = \dfrac{|\tilde{\boldsymbol{w}}_{j+1}\rangle}{\eta_{j+1}}$
(26) **end for**
(27) $\tilde{\boldsymbol{T}}_n = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}_n \boldsymbol{U}^{-1}$
(28) $\boldsymbol{\Lambda}_n = z_0 + (\tilde{\boldsymbol{\Lambda}}_n)^{-1}$

(29) compute impulse response $\tilde{\mathcal{H}}$
(30) compute $|\tilde{\boldsymbol{a}}\rangle_{out}$

In a similar way a projection solver via ARNOLDI process can be derived, which, evidently, is more expensive than Algorithm 5.3.7.

5.4 Passivity and Stability of the TLM system

In this subsection, we discuss the concepts of stability and passivity of a discretized TLM system.

An important property of linear dynamical systems is stability. A physical system needs to be stable in order to function properly. If the TLM systems (5.42) and (5.45) are used as a description of such a physical system, then they should be stable. Moreover, if (5.42) and (5.45) are replaced by reduced-order models (5.53) and (5.70), which are used in time-domain analysis, the reduced-order model also needs to be stable. A stable system is one for which every bounded input produces a bounded output (BIBO system). Next to being stable, a TLM system should also be passive. Roughly speaking, a system is passive if it does not generate energy. We note that passivity is stronger than stability, and that the stability of the overall simulation can be only guaranteed if the reduced-order models preserve the passivity of the original TLM-system. Therefore, it is important to have techniques to check the passivity of a extracted reduced-order model. If we can show that the reduced systems (5.53) and (5.70) are passive, they are proven to be stable as well.

From eq. (2.130) we recall that

$$|_k\boldsymbol{a}\rangle = (\boldsymbol{\Gamma S})^{k-1}|_0\boldsymbol{a}\rangle\,, k = 1, 2, \ldots, \infty\,, \tag{5.155}$$

and

$$\langle_k\boldsymbol{a}| = \left((\boldsymbol{\Gamma S})^T\right)^{k-1}\langle_0\boldsymbol{a}|\,, \tag{5.156}$$

due to reality of $\boldsymbol{\Gamma S}$ operator.

Using the z-transform (5.29) we obtain that

$$|\tilde{\boldsymbol{a}}\rangle_{out} = \sum_{k=1}^{\infty} z^{-k}(\boldsymbol{\Gamma S})^{k-1}\,|\tilde{\boldsymbol{a}}\rangle_{in}\,, \tag{5.157}$$

and

$$\langle\tilde{\boldsymbol{a}}|_{out} = \sum_{k=1}^{\infty}(z^*)^{-k}\left((\boldsymbol{\Gamma S})^T\right)^{k-1}\,\langle\tilde{\boldsymbol{a}}|_{in}\,. \tag{5.158}$$

In order to ensure the passivity we require

$$\Delta P = P_{in} - P_{out} = \langle\tilde{\boldsymbol{a}}|\tilde{\boldsymbol{a}}\rangle_{in} - \langle\tilde{\boldsymbol{a}}|\tilde{\boldsymbol{a}}\rangle_{out} \geq 0\,, \tag{5.159}$$

which is equivalent to the relation

$$\langle\tilde{\boldsymbol{a}}|\tilde{\boldsymbol{a}}\rangle_{out} \leq \langle\tilde{\boldsymbol{a}}|\tilde{\boldsymbol{a}}\rangle_{in}\,. \tag{5.160}$$

By taking eq. (5.157) and eq. (5.158) into account it follows that from (5.160) comes that

$$\sum_{k=1}^{\infty}(z^*z)^{-k}\left((\boldsymbol{\Gamma S})^T\boldsymbol{\Gamma S}\right)^{k-1} \leq 1 \tag{5.161}$$

Since eq. (5.161) is identical to the condition

$$\sum_{k=1}^{\infty}\left((\boldsymbol{\Gamma S})^T\boldsymbol{\Gamma S}\right)^{k-1} \leq 1 \tag{5.162}$$

for $|z| = 1$, the condition for passivity is equivalent to

$$\|\boldsymbol{\Gamma S}\| \leq \mathbf{1}\,, \tag{5.163}$$

and to the following relation

$$\lambda_{max}(\boldsymbol{\Gamma S}) \leq \mathbf{1}\,, \tag{5.164}$$

where λ_{max} the eigenvalue of the $\boldsymbol{\Gamma S}$-operator with maximum module. In this case we are dealing with a *marginally stable* system, i.e. there are no eigenvalues outside the unit circle, and there are no repeated roots on the unit circle. Or in other words, eq. (5.157) describes a system, which in the absence of input produces a steady non-zero output or a bounded oscillatory output over time in response to non-zero input conditions. We note that the system (5.157) with the condition (5.163) is not a BIBO system with a definition of stability, which says that a system is stable if a bounded input produces a bounded output, whereas a marginally stable system remains bounded if and only if the input response is zero. The BIBO-system demands, that all $|\lambda_i| < \mathbf{1}$. Further we refer to a system with eigenvalues satisfying eq. (5.164) as being stable.

Proposition 5.4.1. The system (5.42) having the impulse response (5.44) is passive and stable.

Proof. By appropriate normalizing of $\boldsymbol{R}$ and $\boldsymbol{Q}$ so that $\|\boldsymbol{R}\| = 1$ and $\|\boldsymbol{Q}\| = 1$, and by taking (5.160) and (5.163) into consideration, we see that the system (5.42) is passive and stable. (We remember, that $\boldsymbol{R}$ and $\boldsymbol{Q}$ are real diagonal matrices). □

The poles of the transfer function (5.44) are located inside the unit circle (including the unit circle itself), i.e. the Region of Convergence (ROC) includes the unit circle (Fig. 5.2).

That is essential that the generated reduced-order model preserve the passivity and stability of the initial system (5.42). For example, we consider the reduced-order model (5.53) with the impulse response given by (5.57). Now we show the reduced-order model (5.53) to be passive and stable.

Proposition 5.4.2. The reduced-order model (5.53) with the impulse response given by (5.57) is passive and stable.

Proof. The reduced TLM-operator $(\boldsymbol{\Gamma S})_n$ in (5.57) is obtained by applying the similarity transformation (5.58). It is well-known that the similarity transformation does not change the eigenvalues of an initial matrix (e.g. see [128]). That is

$$\det(\boldsymbol{\Gamma S}) = \det\left((\boldsymbol{\Gamma S})_n\right) \tag{5.165}$$

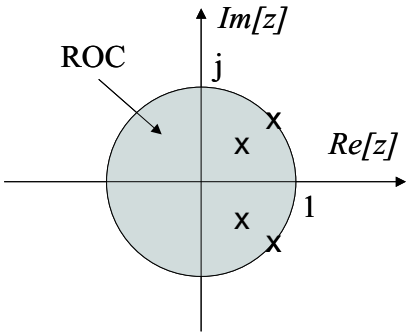


Figure 5.2: ROC for the TLM-system in the complex Z-plane.

and, as follows, the eigenvalues of the reduced TLM-operator $(\boldsymbol{\Gamma S})_n$ satisfy the condition

$$\lambda_{max}\left((\boldsymbol{\Gamma S})_n\right) \leq \mathbf{1}\,. \tag{5.166}$$

In the context of Proposition 5.4.1 we conclude that the system (5.53) preserves the passivity and stability of the initial system (5.42). □

Next, we consider the reduced system (5.70) with the impulse response (5.76).

Proposition 5.4.3. The reduced system (5.70) with the impulse response given by (5.76) is passive and stable.

Proof. The proof follows in context of Proposition 5.4.1 and Proposition 5.4.2 using the similarity transformation (5.72) performed via the LANCZOS algorithm. □

Corollary 5.4.1. The reduced-order models (5.53) and (5.70) with impulse transfer functions (5.57) and (5.76), respectively preserve the passivity and stability of the initial TLM-system (5.42), if only if the extracted approximating eigenvalues of reduced TLM-operators (5.58) and (5.72) are placed inside the unit circle, including the unit circle itself in the complex z-plain (Fig. 5.2).

Corollary 5.4.2. Under the assumption of passivity of the original discrete TLM-system, the passivity of the appropriate reduced-order systems is guaranteed using congruence transformations.

5.5 Comparison of FDTD-, FIT-, and TLM-based MOR

In this section a brief comparison is provided of the FDTD-, FIT- and TLM-based MOR-techniques. For the case of lossless media use is made of the curl-curl formulation of FDTD and FIT [125] to halve the number of state variables and to utilize the symmetric LANCZOS procedure for the generation of the reduced-order model. This reduction in the number of unknowns translates to increased efficiency in the generation of the reduced-order model (TABLE 5.1). However, the presence of loss destroys the symmetry of the FDTD and

the FIT state-space system. Thus, despite the fact that the length of the state vector of the FDTD/FIT model is smaller than that of the TLM, the use of a non-symmetric LANCZOS process results in increased computational complexity of the model order reduction process. Nevertheless, despite the computational advantage of the S-symmetric TLM-MOR process, the larger dimension of the TLM state vector (three times that of the FDTD/FIT model for the case of lossy media) is an issue, since in some cases, excessively large dimensions of TLM-based state-space systems are responsible for slow convergence of the model order reduction process.

Equivalence of a finite-difference scheme and the TLM time evolution scheme is given by eq. (5.39).

Table 5.1: Comparison of FDTD, FIT and TLM-MOR.

Method	lossless case	with losses
FDTD, FIT	$3 \times N_{cells}$ [1]	$6 \times N_{cells}$
TLM	$12 \times N_{cells}$	$18 \times N_{cells}$
FDTD, FIT	symmetric LANCZOS	non-symmetric LANCZOS
TLM	S-symmetric LANCZOS	S-symmetric LANCZOS

[1] N_{cells} is the total number of discretization cells.

5.5.1 The TLM Reduced-Order Model in the Laplace Domain

To obtain the s-domain statement of the TLM discretized system we need to utilize the bilinear transformation that maps the unit circle on the z-plane to the left half plane of the s-plane [86]. This transformation is

$$z = \frac{1+s}{1-s}\,. \tag{5.167}$$

Utilizing the state-space representation of the TLM system (5.30), from (5.167) we obtain

$$s(\boldsymbol{I}+\boldsymbol{\Gamma S})|\tilde{\boldsymbol{a}}\rangle_s + (\boldsymbol{I}-\boldsymbol{\Gamma S})|\tilde{\boldsymbol{a}}\rangle_s = \boldsymbol{0}\,, \tag{5.168}$$

which is *state equation of the TLM-system in Laplace domain.*

Taking eq. (5.168) into account, we present eq. (5.45) in the form of (5.14). That is

$$\begin{aligned} (\boldsymbol{G}+s\boldsymbol{C})\,|\tilde{\boldsymbol{a}}\rangle_s &= \boldsymbol{R}_1\,|\boldsymbol{i}\rangle\,, \\ |\boldsymbol{u}\rangle &= \boldsymbol{Q}_1\,|\tilde{\boldsymbol{a}}\rangle_s\,, \end{aligned} \tag{5.169}$$

where

$$\boldsymbol{C} = \boldsymbol{I}+\boldsymbol{\Gamma S} \;\text{ and }\; \boldsymbol{G} = \boldsymbol{I}-\boldsymbol{\Gamma S}\,. \tag{5.170}$$

The GIM of the TLM-system results in the s-domain as

$$\boldsymbol{Z}(s) = \boldsymbol{Q}_1(\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{R}_1\,. \tag{5.171}$$

and the impulse response of the TLM-system (5.42) in the s-domain

$$\begin{aligned} (\boldsymbol{G}+s\boldsymbol{C})\ |\tilde{\boldsymbol{a}}\rangle_s &= \boldsymbol{R}\,|\tilde{\boldsymbol{a}}\rangle_{in}\ , \\ |\tilde{\boldsymbol{a}}\rangle_{out} &= \boldsymbol{Q}\ |\tilde{\boldsymbol{a}}\rangle_s\ , \end{aligned} \tag{5.172}$$

is given by

$$\tilde{\boldsymbol{\mathcal{H}}}(s) = \boldsymbol{Q}(\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{R}. \tag{5.173}$$

Introducing

$$\boldsymbol{A}_1 = -\boldsymbol{C}^{-1}\boldsymbol{G} = -(\boldsymbol{I}+\boldsymbol{\Gamma S})^{-1}(\boldsymbol{I}-\boldsymbol{\Gamma S}) \tag{5.174}$$

we can write (5.171) and (5.173) in the forms

$$\boldsymbol{Z}(s) = \boldsymbol{Q}_1(s\boldsymbol{I}-\boldsymbol{A}_1)^{-1}(\boldsymbol{I}+\boldsymbol{\Gamma S})^{-1}\boldsymbol{R}_1\,, \tag{5.175}$$

and

$$\tilde{\boldsymbol{\mathcal{H}}}(s) = \boldsymbol{Q}(s\boldsymbol{I}-\boldsymbol{A}_1)^{-1}(\boldsymbol{I}+\boldsymbol{\Gamma S})^{-1}\boldsymbol{R}. \tag{5.176}$$

From (5.175) and (5.176) we can see that that the generation of reduced-order models in the LAPLACE domain by the KRYLOV subspace methods already requires the inverse of the implicit general TLM-operator.

Next, we examine the possibility to apply the PLV or/and SI approaches for generating TLM-reduced-order models in the s-domain. For this purpose we consider the system (5.172) with the impulse response (5.173), which can be modified in the following way:

$$\begin{aligned} \tilde{\boldsymbol{\mathcal{H}}}(s') = \boldsymbol{Q}(\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{R} &= \boldsymbol{Q}(\boldsymbol{G}+s_0\boldsymbol{C}+(s-s_0)\boldsymbol{C})^{-1}\boldsymbol{R} = \\ &= \boldsymbol{Q}\left(\boldsymbol{I}+(s-s_0)(\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{C}\right)^{-1}(\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{R} = \\ &= \boldsymbol{Q}(\boldsymbol{I}-s'\boldsymbol{A}_2)^{-1}\hat{\boldsymbol{R}}, \end{aligned} \tag{5.177}$$

where $s' = s - s_0$,

$$\boldsymbol{A}_2 = -(\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{C}\,, \tag{5.178}$$

and

$$\hat{\boldsymbol{R}} = (\boldsymbol{G}+s\boldsymbol{C})^{-1}\boldsymbol{R}. \tag{5.179}$$

Now, we more precisely consider the term

$$\begin{aligned} (\boldsymbol{G}+s\boldsymbol{C})^{-1} &= (\boldsymbol{I}-\boldsymbol{\Gamma S}+s_0\boldsymbol{I}+s_0\boldsymbol{\Gamma S})^{-1} = \\ &= \left((1+s_0)\left(\boldsymbol{I}-\frac{1-s_0}{1+s_0}\boldsymbol{\Gamma S}\right)\right)^{-1} = \\ &= \frac{1}{1+s_0}\left(\boldsymbol{I}-\frac{1-s_0}{1+s_0}\boldsymbol{\Gamma S}\right)^{-1}. \end{aligned} \tag{5.180}$$

From (5.180) we see that the utilization of the $\boldsymbol{A}_2$-matrix in the s-domain as well as the utilization of $\hat{\boldsymbol{A}}$ (5.94) in the z-domain for the PVL/PVA and SI approaches require the inversion of the $\boldsymbol{\Gamma S}$ operator. This inversion is prohibitive (in possible at all) by dealing with implicit matrices, because the elements of the matrix are not accessible for the inversion operation.

5.5.2 S-domain versus Z-domain in TLM-MOR

In this subsection we give a brief comparison between the TLM reduced-order models in the z-domain and the s-domain. In general, the transformation of the TLM-MOR from the z-domain into the s-domain yields disadvantages as a result. Some important properties or TLM-MOR are provided below by their comparison in the z- and s-domains.

- The generation of reduced-order models in the LAPLACE domain by direct application of KRYLOV subspace methods already requires the inverse of the implicit general TLM-matrix, that does not occur in the z-domain. This follows by a comparison of the impulse transfer functions (5.44) and (5.176) with operating matrices $\boldsymbol{\Gamma S}$ and (5.174) respectively. If we can figure out how to operate with the inverse of the general TLM operator without forming itself explicitly, then we should be able to proceed with the application of the LANCZOS and ARNOLDI processes for effecting MOR in the s-domain in context of TLM. Thus, from this point od view, the z-domainbased on the TLM-MOR is still favorable.

- It is well-known that the LANCZOS and ARNOLDI algorithms tend to first extract the approximating eigenvalues at the boundary of the convex hull of the eigenvalue spectrum. That means that the spurious intermesh and higher-order modes will be extracted first, if we deal with TLM-MOR in the s-domain (at $s \to \infty$), whereas the utilization of the LANCZOS and ARNOLDI processes for TLM-ROM in the z-domain yields the permanent distribution of extracted eigenfrequencies on the unit circle. That is, we have faster convergence in the z-domain than in the LAPLACE domain due to the specific convergence properties of KRYLOV subspace methods by solving eigenvalue problems.

- The PVL and SI-approaches require the inversion of the $\boldsymbol{\Gamma S}$ -operator in both domains. Operations with $\hat{\boldsymbol{A}}$ (5.94) in the z-domain and with $\boldsymbol{A}_2$ (5.180) in the s-domain exhibit similar convergence properties and have similar computational efforts.

- The length of the state-space vector $|\tilde{\boldsymbol{a}}\rangle$ in (5.42) and (5.172) in both domains is kept the same.

5.6 Post-Processing

In this section we discuss following issues:

- how to accelerate the generation of a reduced-order model,

- how to select the optimal order of a reduced model,

- how to recover the signal in the time-domain by an extracted reduced-order model.

5.6.1 Two-Step Reduction Approach

The two step reduction approach combines a LANCZOS based projection in the first step to pre-reduce the system (5.42) having the transfer function (5.44) with a PADÉ or SI-methods in the second step to get an optimal approximation. When there is a large number of n, the computation of the eigenvalues by Eigenvalue Decomposition (EVD) is very expensive. Instead of that, the second reduction of a system can be applied so as to extract only eigenvalues corresponding to a needed frequency band. The two step reduction algorithm allows us to decrease the computational effort in TLM-MOR and save computational time. In addition, the second step yields an optimal order (or rather a minimal order) of a reduced model.

Let $(\boldsymbol{\Gamma S})_n \in \mathcal{H}^{n\times n}$ be a reduced TLM-operator, which is obtained via similarity transformation by the LANCZOS algorithm (5.72). Inducing the second biorthogonal basis

$$(\breve{\boldsymbol{W}})^T_{m\times n}\breve{\boldsymbol{V}}_{n\times m} = \breve{\boldsymbol{D}}_{m\times m} \tag{5.181}$$

with $m \ll n \ll N$ we can present the the reduced TLM-operator $(\boldsymbol{\Gamma S})_n$ as follows:

$$(\breve{\boldsymbol{W}})^T(\boldsymbol{\Gamma S})_n\breve{\boldsymbol{V}} = \breve{\boldsymbol{D}}\breve{\boldsymbol{T}}\,. \tag{5.182}$$

Using the eigenvalue decomposition of $\breve{\boldsymbol{T}}$

$$\breve{\boldsymbol{T}} = \breve{\boldsymbol{U}}\breve{\boldsymbol{\Lambda}}\breve{\boldsymbol{U}}^{-1}\,, \tag{5.183}$$

and substituting (5.182) and (5.183) into (5.78) we obtain *oblique-oblique projection*

$$\tilde{\boldsymbol{H}}_{n_{21}} = \boldsymbol{Q}\boldsymbol{V}\breve{\boldsymbol{V}}\breve{\boldsymbol{U}}(z\boldsymbol{I} - \breve{\boldsymbol{\Lambda}})^{-1}\breve{\boldsymbol{U}}^{-1}\breve{\boldsymbol{D}}^{-1}\breve{\boldsymbol{W}}^T\boldsymbol{D}^{-1}\boldsymbol{W}^T\boldsymbol{R}. \tag{5.184}$$

Using the orthonormal basis

$$(\breve{\boldsymbol{V}})^T_{m\times n}\breve{\boldsymbol{V}}_{n\times m} = \breve{\boldsymbol{I}}_{m\times m} \tag{5.185}$$

we can write eq. (5.184) as *oblique-orthogonal projection*

$$\tilde{\boldsymbol{H}}_{n_{22}} = \boldsymbol{Q}\boldsymbol{V}\breve{\boldsymbol{V}}\breve{\boldsymbol{U}}(z\boldsymbol{I} - \breve{\boldsymbol{\Lambda}})^{-1}\breve{\boldsymbol{U}}^{-1}\breve{\boldsymbol{V}}^T\boldsymbol{D}^{-1}\boldsymbol{W}^T\boldsymbol{R}. \tag{5.186}$$

The sketch of TLM-MOR based on the two step reduction approach is given below.

Algorithm 5.6.1. The TLM-ROM based on the two step reduction approach

Step 1. The S-symmetric version of the LANCZOS algorithm (Algorithm 5.3.3, steps 1-19) is applied directly to the general TLM-operator $\boldsymbol{\Gamma S}$ in order to gain the reduced TLM-operator $(\boldsymbol{\Gamma S})_n$. The number n is determined during the computation by eigenvalue convergence criteria, which are based on the observation of eigenvalues, corresponding to a needed frequency range. In order to avoid the large computational overheads required by eigenvalue computation at the first step, we can check the eigenvalues in predefined iteration intervals, or we can use some empirical formulas, e.g. the number of iterations required in the first step is of the order $\mathcal{O}(N^{\frac{1}{3}})$. The pre-reduced system is obtained via oblique projection in (5.72). We have to store the tridiagonal matrix $\boldsymbol{T}_n$, diagonal matrix $\boldsymbol{D}$ and the biorthogonal $V - W$ basis.

Step 2. After the first step the model order of $\boldsymbol{\Gamma S}$ is reduced (from N to n), and the second step is applied in order to gain an optimal model size $m \ll n$. For this purpose the PVL-, PVA-, SI-via-LANCZOS-, SI-via-ARNOLDI-approaches can be utilized, depending on the basis to be generated. The *oblique-oblique* or *oblique-orthogonal* projections can be performed by (5.184) and (5.186) respectively. For example, we form the shift-inverse pre-reduced trigiagonal matrix

$$\hat{\breve{\boldsymbol{T}}} = (\breve{\boldsymbol{T}} - z_0 \boldsymbol{I})^{-1} \tag{5.187}$$

and carry out the Algorithm 5.3.5, yielding eigenvalues of $m \times m$-matrix $(\breve{\boldsymbol{T}})$, which also approximate the eigenvalue spectrum of the pre-reduced matrix $\boldsymbol{T}_n$. Thus, the impulse transfer function of the system (5.70) is obtained by (5.186) instead of (5.76). The typical iteration number for the second step is less than 50. The stop criteria are defined very similarly to the ones in the first step.

As we can see, the main disadvantage of the two step reduction approach is to gain the inverse of $(\breve{\boldsymbol{T}} - z_0 \boldsymbol{I})^{-1}$. From this point of view, the use of the SI-via-ARNOLDI-approach is preferential. Even though SI-via-LANCZOS-process also requires

$$\left((\breve{\boldsymbol{T}} - z_0 \boldsymbol{I})^T\right)^{-1} = \left((\breve{\boldsymbol{T}} - z_0 \boldsymbol{I})^{-1}\right)^T ,$$

i.e. one additional computational "chip" transposition operation, the utilization of the ARNODI process for the shift-inverse and/or PADÉ-based approaches seems to be more stable and more accurate .

5.6.2 Model Order Selection

The model order selection represents a trade-off between minimum complexity and maximum accuracy of a reduced-order model, and, thus, allows its optimal order to be gained.

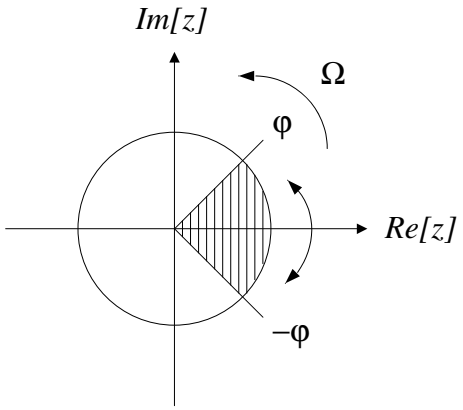


Figure 5.3: Sector of the ROC.

If the KRYLOV subspace methods are applied directly, from the previous sections we have seen, that the order of the reduced model is primarily given by the size of the reduced TLM-operator $(\boldsymbol{\Gamma S})_n$, having the order of n. Such a

reduced-model consists of a complete frequency spectrum of the TLM-system, including the spurious components as well. First of all, we select the eigenvalues lying in a sector of the unit circle, which corresponds to the demanded frequency range $[f_1, f_2]$, and we also discard all spurious eigenvalues as eigenfrequencies outside the needed frequency range. For this purpose we find a relation between an angle φ in the complex plane and a corresponding operating frequency f (Fig. 5.3).

Generally, we can write

$$|\lambda| e^{\jmath\varphi} = e^{\alpha} \cdot e^{\jmath\Omega} . \tag{5.188}$$

From (5.188) we have the attenuation

$$\alpha = \ln(|\lambda|) , \tag{5.189}$$

and the operation frequency

$$f = \frac{\varphi}{2\pi\Delta t} = \frac{\varphi}{\pi} f_{max} \tag{5.190}$$

of a mode. That is the magnitude of an extracted complex eigenvalue λ_i yields the attenuation of the mode, and the angle of λ_i corresponds to the operating frequency f_i. The relevant angles φ_1 and φ_2 for the limit frequencies f_1 and f_2 can be determined by (5.190). Then, we keep the eigenvalues of $(\boldsymbol{\Gamma S})_n$ (and/or poles of the transfer function), which are placed in a sector bounded by angles φ_1 and φ_2 (Fig. 5.4), and in this sector we discard the eigenvalues for which $|\lambda| > 1$ in order to preserve the passivity and stability. The ROC corresponding to the frequency band $[f_1, f_2]$ is shown in Fig. 5.4.

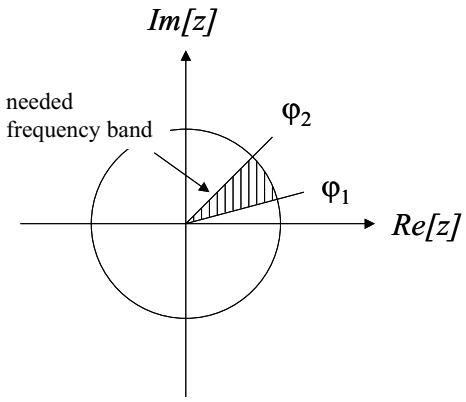


Figure 5.4: Sector of the ROC corresponding to a needed frequency band.

We note due to the symmetry of the poles for the linear case, the order of the model has to be doubled, or in other words, the sector $[\varphi_1, \varphi_2]$ must be mirrored into the lower semicircle with regard to the real z-axis.

An intuitive approach for decreasing the model order would be to observe the resonance peaks (the zeros of the transfer function) and select those which exert dominant influence on the transfer function. This selection can be carried out by the *Signal-to-Noise Ratio* (SNR) technique, which is a powerful tool in digital signal processing that allows the extraction of the important information

about a system from a corresponding reduced system matrix. The reduced matrix consists not only of the signal, but also of noise. The SNR method allows the separation of the desired signal from the coexisting noise. This can be achieved by a selection of important eigenvalues corresponding to a signal in a needed frequency band and, hence, the order of an model extracted by KRYLOV subspace methods can also be reduced.

For example, let we consider the transfer function given by (5.63). We denote the residues of the transfer function from p_1 to port p_2 via c. In order to separate the signal from noise, we sort n extracted eigenvalues so that the consequent resonance peaks are in increasing order. That is

$$|c_1| \gg |c_2| \gg |c_3| \gg \ldots \gg |c_q| \gg \ldots \gg |c_n| \,. \tag{5.191}$$

And we introduce a SNR given by

$$SNR(q) = 10 \log_{10} \frac{\sum_{i=1}^{q} |c_i|^2}{\sum_{j=q+1}^{n} |c_j|^2} \,. \tag{5.192}$$

The model order of the system or signal associated with data matrix T_n is given from the dimension of the underlying signal space, or equivalently from the approximated rank of T_n. The order of the model is selected as q if

$$SNR(q) \geq \mathcal{D}_T \,, \tag{5.193}$$

where $\mathcal{D}_T$ is a threshold, the reasonable value of which can be determined by satisfying the following:

$$|\mathcal{H}_n^{i,j} - \mathcal{H}_q^{i,j}| \ll \epsilon \,. \tag{5.194}$$

In the case of multi-port systems the situation is more complicated. For example, we consider a MIMO-system with p_1 input and p_2 output ports. The thresholds $\mathcal{D}_T^{i,j}$ for each transfer function from port i to port j are determined by satisfying (5.194). The value $\mathcal{D}_T$ is selected as the least value from the set of $\mathcal{D}_T^{i,j}$, i.e.

$$\mathcal{D}_T = \min \{\mathcal{D}_T^{i,j}\} \,, \quad i = 1, 2, \ldots, p_1, \; j = 1, 2, \ldots, p_2 \,. \tag{5.195}$$

Some of the responses $\mathcal{H}_q^{i,j}$ may contain one or more poles, which already exist in other responses. The final set of poles must be unique for all of the $p_1 \times p_2$ responses. These multi-poles should therefore be disregarded, e.g. by a simple sorting. Moreover, some poles can be clustered, and techniques called pole clustering should be applied [120]. This technique consists of averaging poles which numerically differ from a given tolerance and that in theory refer to the same natural frequency. The pole clustering allows to adaptively modify the network's natural frequencies by seeking the most likely values among those found in the entire set of simulated responses. Pole clustering, however, is a very critical operation which strongly effects the model accuracy, and is not considered in the present doctoral work. It could be more propitious to save the complete set of poles for each single input-output transfer function in a MIMO-system, than to perform the pole clustering for the multi-port transfer function of this MIMO system.

In conclusion, let us summarize the procedure of model order selection (MOS).

Algorithm 5.6.2. The Model Order Selection procedure.

(1) Select the poles lying in a sector in the complex z-plane in accord with a frequency range in which the impulse transfer function must be approximated.
(2) Selection of true poles, which have converged and should not change significantly between MOR-algorithm iterations.
(3) Select the stable poles, which are placed inside and on the unit circle.
(4) Select the order of the model using the SNR technique. The most significant natural frequencies have been taken into account. Poles which do not significantly affect the system response (very low value) may be discarded.
(5) Validation of the model, for example by means of the Mean Square Error (MSE) procedure.

5.6.3 Time Response of the TLM Reduced-Order Model

Using the inverse z-transform of (5.63)

$$\tilde{\mathcal{H}}_n^{(i,j)} \xrightarrow{z^{-1}} h^{(i,j)}[k] \tag{5.196}$$

we obtain the impulse response of the TLM-system from the port i to the port j in time-domain as follows:

$$h^{(i,j)}[k] = \sum_{\hat{n}=1}^{n} c_{\hat{n}}^{(i,j)} \left(\lambda_{\hat{n}}\right)^{k-1} . \tag{5.197}$$

Using eq. (5.188) we obtain from (5.197) the expression

$$h^{(i,j)}[k] = \sum_{\hat{n}=1}^{n} c_{\hat{n}}^{(i,j)} \left|\lambda_{\hat{n}}\right|^{k-1} e^{j\Omega(k-1)} . \tag{5.198}$$

Because of the linearity and passivity of the TLM-system the poles of the transfer function (5.63) with $p_1 = p_2$ are complex conjugated (for more details see [3]). Taking this property into account from (5.198) we derive that

$$h^{(i,j)}[k] = \sum_{\hat{n}=1}^{n/2} 2\, c_{\hat{n}}^{(i,j)} \left|\lambda_{\hat{n}}\right|^{k-1} \cos\Omega(k-1) . \tag{5.199}$$

Eq. (5.199) allows us to recover the time-domain response of the TLM-system from the residue/pole distribution of the transfer function in the z-domain.

Chapter 6

Numerical Applications

This chapter contains numerical studies from the application of MOR to the TLM analysis of various lossless and lossy electromagnetic structures. The computational efficiency of MOR-techniques based on the ARNOLDI and LANCZOS processes is discussed here. The convergence properties of certain proposed algorithms and the accuracy of generated reduced-order models are considered in detail. The results are presented from the application of TLM-MOR for the extraction of the matrix transfer function representations of the ROM, such as generalized impedance or scattering-parameter matrices. In all cases, MATLAB 6.5 realizations of the MOR-algorithms are used for computations. The reference computational station is a PENTIUM IV 3 GHz machine with 1 GB RAM.

6.1 Cubic cavity

In order to test the convergence of the ARNOLDI and LANCZOS algorithms we consider a rectangular air-filled cubic 1 cm–resonator with ideally conducting walls. The excitation node and the observation point are placed at the opposite edges of the resonator according to Fig. 6.1.

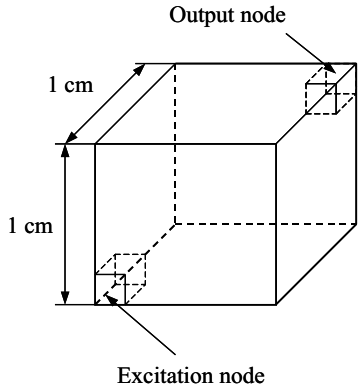


Figure 6.1: Simulated cubic cavity.

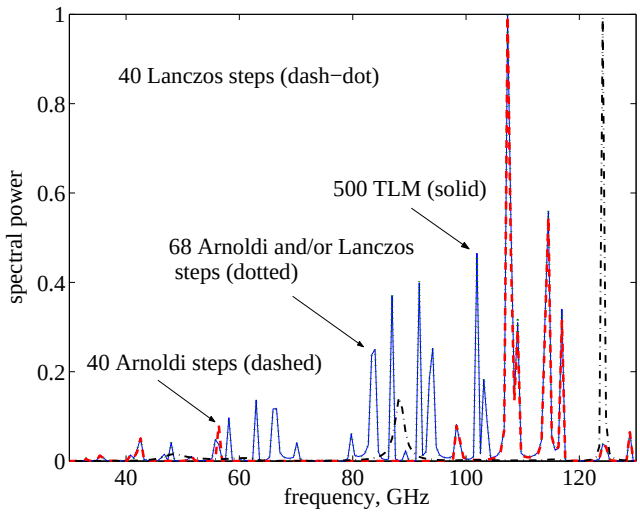


Figure 6.2: Spectral power of the cubic cavity (Fig. 6.1) computed by the ATLM and LTLM algorithms.

The structure is investigated for the 5-cell discretization in each direction. The corresponding state vector of wave amplitudes has a dimension of $12 \times 5 \times 5 \times 5 = 1500$; thus, the dimension of the original discretized lossless model is 1500.

The resonances of the original discretized system are determined by the FOURIER transformation of the TLM time-domain response. The cavity is excited by the DIRAC pulse, and the TLM time evolution scheme is performed for 500 steps.

Then, we applied MOR-algorithms in order to extract the eigenfrequencies of the cavity and to generate a reduced-order model, which approximates the transfer function from the excitation to the observation point.

We considered all possible modes of the cavity, including nonphysical spurious modes as well. For the 5-cell discretization all eigenvalues corresponding to the resonances are converged at 68 ARNOLDI and/or LANCZOS algorithm steps, and they are placed on the unit circle in the complex z-plane due to no losses in the cavity (see Fig. 6.2(f)). Thus, the order of the generated reduced-order model is 68, which corresponds to the reduction of the TLM operator by a factor of $1500/68 \approx 22$.

The spectral power of the cavity obtained by the TLM time evolution scheme, the ARNOLDI (ATLM) and non-symmetric LANCZOS algorithm (LTLM) is shown in Fig. 6.2. This figure shows the better convergence of the ARNOLDI algorithm in comparison to the LANCSZOS procedure at 40 steps. This phenomena is also demonstrated in Fig. 6.6, where the convergence of eigenvalues computed by MOR-algorithms is presented. Converged eigenvalues can almost certainly be recognized by their position close to the unit circle in the complex plane (see Fig. 6.6(f)). The S-symmetric LANCZOS (S-LTLM) algorithm has a similar convergence behavior as its general nonsymmetric variant.

In general, the ARNOLDI process exhibited better convergence than LANCZOS based types for this simple cavity. This well-known phenomenon can be

explained through the fast loss of the bi-orthogonality in the non-symmetric and S-symmetric procedures, whereas the ARNOLDI process generates the orthonormal basis nearly without losses of orthogonality. In order to study losses of orthogonality (and/or biorthogonality) we computed scalar values δ_{ij}, given by

$$\delta_{ij}^{Lanczos} = 20 \log_{10} |\langle \boldsymbol{w}_i | \boldsymbol{v}_j \rangle| \tag{6.1}$$

$$\delta_{ij}^{Arnoldi} = 20 \log_{10} |\langle \boldsymbol{v}_i | \boldsymbol{v}_j \rangle|, \quad i,j = 1,2,\dots 100, \tag{6.2}$$

and depicted them in Fig. 6.3 and Fig. 6.4. These figures show how the LANCZOS algorithms lose the biorthogonality property; already after 15 steps the biorthogonality of both the non-symmetric and the S-symmetric versions is deteriorated, that is clearly to see in Fig. 6.4(b)-(c). By comparing Fig. 6.3(a) and Fig. 6.3(b) we can conclude that the ARNODI procedure generates an "ideal" orthonormal basis.

In order to improve the deteriorated basis biorthogonality in the LANCZOS algorithms the two-sided modified GRAM-SCHMIDT (TSMGS) process process has been proposed in the recommended literature (see APPENDIX B). However, investigations showed that the rebiorthogonalization by TSMGS not only not recovered the biorthogonal property, but also additionally deteriorated it. A comparison between the results with and without the rebiorthogonalization at the various step numbers is presented in Fig. 6.4(b) and Fig. 6.5. For example, the performance of rebiorthogonalization by TSMGS at each 3rd step destroys the biorthogonal basis completely, and the TSMGS process at each 7th step significantly corrupts the biorthogonality, that can be seen by comparing Fig. 6.4(b) and Fig. 6.5(c).

TABLE 6.1 and TABLE 6.2 contain the norms of residual vectors. The small norms of the ARNOLDI residuals (see TABLE 6.1) and left and right LANCZOS residuals (see TABLE 6.2) at 68 steps indicate the convergence of the extracted eigenvalues. We note that the residual norms in the ARNOLDI process are less than left and right residual norms in the non-symmetric LANCZOS algorithms by a factor of $\mathcal{O}(10^2)$, whenever the convergence is achieved.

A vector of the condition numbers for the eigenvalues of the reduced upper HESSENBERG $\boldsymbol{H}_n$ and tridiagonal $\boldsymbol{T}_n$ matrices is observed as well. These condition numbers s_i are the reciprocals of the cosines of the angles between the left and right eigenvectors, and they are unique only if the extracted eigenvalues are simple. Roughly speaking, if perturbations of the order ϵ are caused in the matrix $(\boldsymbol{\Gamma S})_n$ reduced by KRYLOV subspace methods, then an eigenvalue λ_i can be perturbed by an amount $\epsilon \cdot s_i$. Thus, if s_i is big, then small changes in $(\boldsymbol{\Gamma S})_n$ can induce large changes in the eigenvalue λ_i. In this case λ_i is regarded as ill-conditioned. Comparing TABLE 6.1 and TABLE 6.3, we conclude that eigenvalues of the LANCZOS trigiagonal matrix $\boldsymbol{T}_n$ are more sensitive to its changes than eigenvalues of the ARNOLDI matrix are to the changes in $\boldsymbol{H}_n$. (The eigenvalues of the reduced matrices are obtained by means of the eigenvalue decomposition with balancing).

Large condition numbers imply that $(\boldsymbol{\Gamma S})_n$ is near a matrix with multiple eigenvalues. The proof and more details are given in [48].

Indeed, at 68 steps the unique eigenvalues $0.9045 \pm \jmath 0.4265$ are presented in the eigenvalue spectrum extracted by both the LANCZOS and the ARNOLDI algorithms. Performing the algorithms for 75 steps, we obtain multiples of

Table 6.1: Convergence and condition numbers of the ARNOLDI eigenvalues.

	$\|\|r_{res}\|\|_2$			condition number		
steps	min	mean	max	min	mean	max
10	0.059	0.337	0.411	1.00	1.14	1.19
20	0.008	0.246	0.359	1.00	1.12	1.19
30	0.001	0.231	0.420	1.00	1.35	2.48
40	$3.4 \cdot 10^{-5}$	0.131	0.249	1.00	1.07	1.20
50	$3.6 \cdot 10^{-7}$	0.095	0.311	1.00	1.13	2.00
60	$8.2 \cdot 10^{-10}$	0.031	0.088	1.00	1.00	1.01
68	$\mathbf{1.4 \cdot 10^{-17}}$	$\mathbf{8.6 \cdot 10^{-8}}$	$\mathbf{5.8 \cdot 10^{-7}}$	**1.00**	**1.00**	**1.00**
70	$8.3 \cdot 10^{-18}$	$4.6 \cdot 10^{-8}$	$2.9 \cdot 10^{-6}$	1.00	11.86	316.44
75	$2.2 \cdot 10^{-18}$	0.001	0.023	1.00	152.98	$5.2 \cdot 10^{3}$

these eigenvalues: in the LANCZOS eigenvalue spectrum we observe $0.9042 \pm \jmath 0.4266$, and the ARNOLDI eigenvalue spectrum additionally contains the values $0.9024 \pm \jmath 0.4276$. In addition, doubles for 1 and -1 are presented in the extracted spectra. This phenomenon is very well indicated by a large condition number of the order $\mathcal{O}(10^3)$, which is computed after 75 ARNOLDI steps are performed (see TABLE 6.1).

The non-symmetric LANCZOS algorithm exhibits relatively large condition numbers already after 15-20 steps, which indicates the loss of the basis biorthogonality, and the algorithm tends to yield spurious multiples of the real existing eigenvalues. The S-symmetric LANCZOS method generates a tridiagonal matrix similar to one formed by the nonsymmetric version, i.e. computed eigenvalues have the same sensitivity.

Table 6.2: Convergence of the LANCZOS eigenvalues.

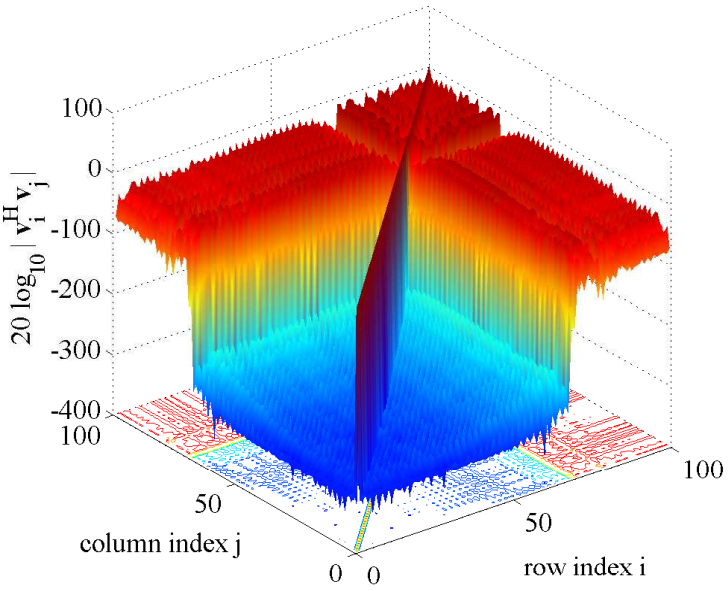
	$\|\|r_{res}\|\|_2$			$\|\|l_{res}\|\|_2$		
steps	min	mean	max	min	mean	max
10	0.047	0.420	0.818	0.011	0.202	0.900
20	0.003	0.010	0.019	$1.7 \cdot 10^{-4}$	0.004	0.015
30	$1.7 \cdot 10^{-4}$	0.095	0.832	$6.8 \cdot 10^{-7}$	0.123	1.133
40	$1.1 \cdot 10^{-5}$	0.025	0.230	0.001	0.045	0.112
50	$3.2 \cdot 10^{-7}$	0.030	0.460	$3.6 \cdot 10^{-5}$	0.028	0.535
60	$2.1 \cdot 10^{-10}$	0.001	0.007	0.002	0.011	0.038
68	$\mathbf{1.3 \cdot 10^{-15}}$	$\mathbf{1.2 \cdot 10^{-6}}$	$\mathbf{1.1 \cdot 10^{-5}}$	$\mathbf{1.1 \cdot 10^{-6}}$	$\mathbf{1.2 \cdot 10^{-5}}$	$\mathbf{4.4 \cdot 10^{-5}}$
70	$3.3 \cdot 10^{-16}$	0.006	0.370	$4.6 \cdot 10^{-21}$	0.006	0.356
75	$4.6 \cdot 10^{-15}$	0.009	0.654	$4.7 \cdot 10^{-20}$	0.010	0.657

Table 6.3: Condition numbers of the LANCZOS eigenvalues.

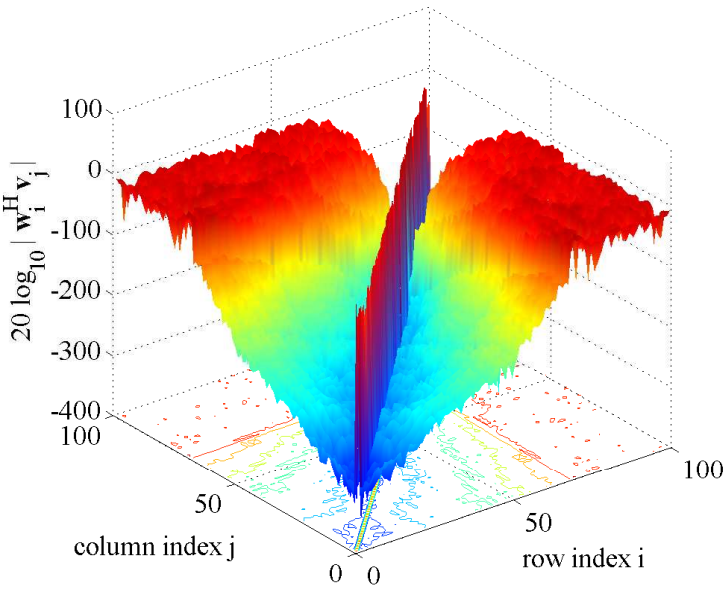
	condition number		
steps	min	mean	max
10	3.14	5.89	9.33
20	11.92	59.18	107.81
30	2.40	56.78	114.47
40	2.52	55.17	151.86
50	2.52	67.86	152.79
60	10.05	65.26	142.20
68	**10.05**	**63.33**	**117.09**
70	1.00	61.55	117.09
75	1.00	57.56	117.09

From thorough investigations we can conclude:

- the LANCZOS methods exhibit slower convergence than the ARNOLDI algorithm due to their losses of basis biorthogonality;
- the utilization of the very expensive rebiorthogonalization does not improve the biorthogonality property, and the biorthogonality is even destroyed in some cases;
- the ARNOLDI algorithm generates an orthonormal basis nearly without the loss of orthogonality;
- eigenvalues of the reduced LANCZOS trigiagonal matrix are more sensitive to its changes than eigenvalues of the reduced ARNOLDI to its respective changes;
- an approximation error of the LANCZOS algorithms is larger than that of the ARNOLDI process due to the fast losses of biorthogonality and the high sensitivity of computed eigenvalues to small changes in the reduced tridiagonal matrix. The biorthogonality loss contributes errors to elements of the reduced matrix, and these errors can additionally make a significant impact on the accuracy of the eigenvalue computation.

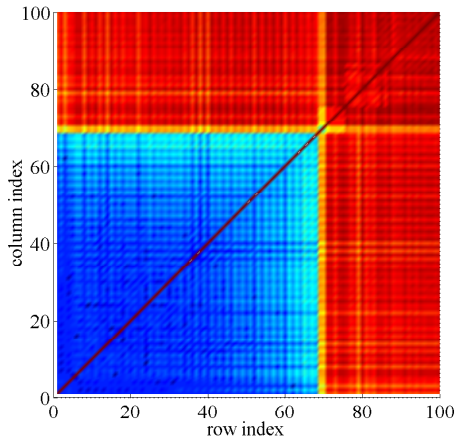


(a) Orthogonality of the Arnoldy basis, i.e. $\boldsymbol{V}^H \boldsymbol{V} = \boldsymbol{I}$

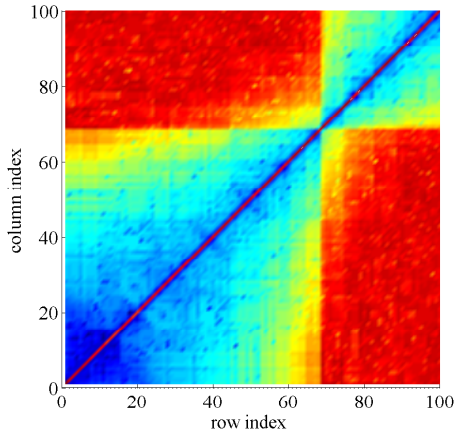


(b) Biorthogonality of the LANCZOS basis, i.e. $\boldsymbol{W}^H \boldsymbol{V} = \boldsymbol{D}$

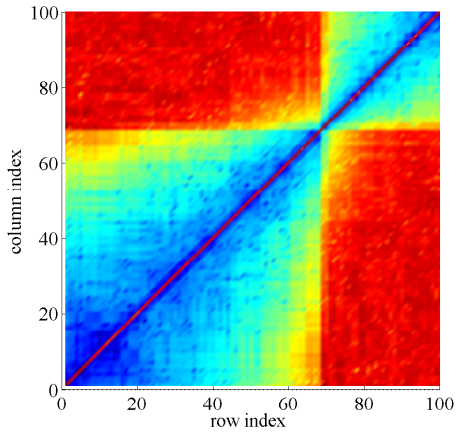
Figure 6.3: Orthogonality and biorthogonality of bases generated by the ARNOLDI and LANCZOS algorithms, respectively, for the cubic cavity, 3-D view.



(a) ATLM orthogonal basis

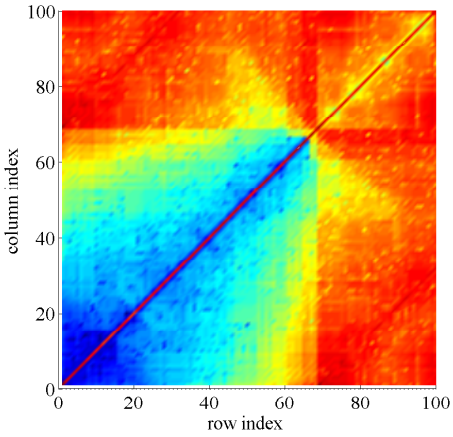


(b) LTLM biorthogonal basis

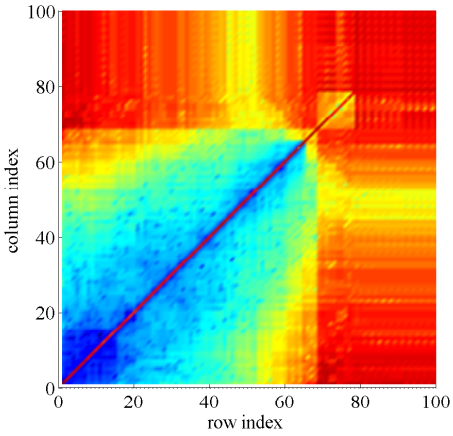


(c) LTLM-S biorthogonal basis

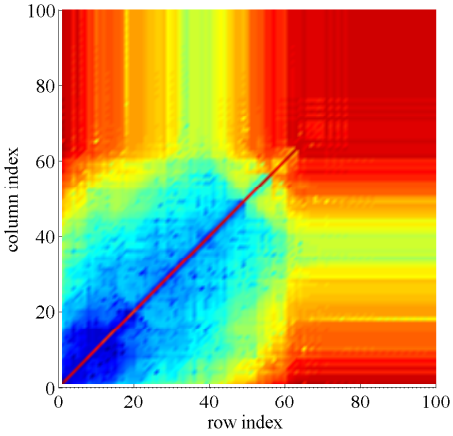
Figure 6.4: Orthogonality and biorthogonality of bases generated by the ARNOLDI and LANCZOS algorithms, respectively, for the cubic cavity, 2-D view.



(a) Rebiorthogonalization at each 33rd step

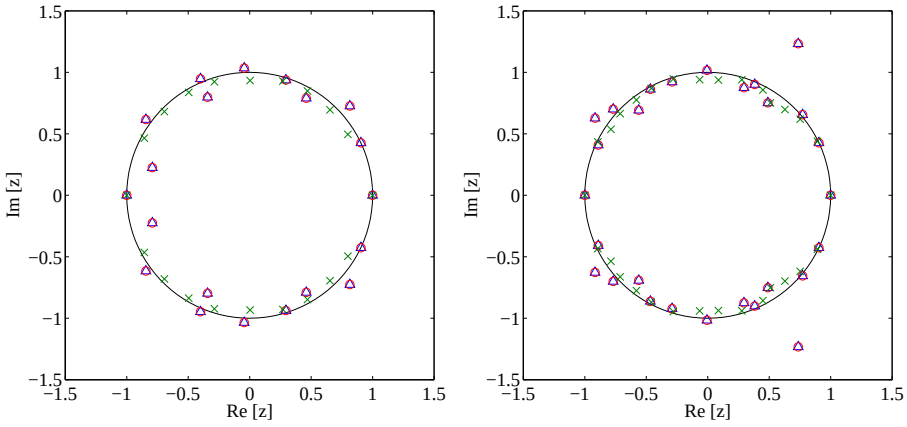


(b) Rebiorthogonalization at each 13th step



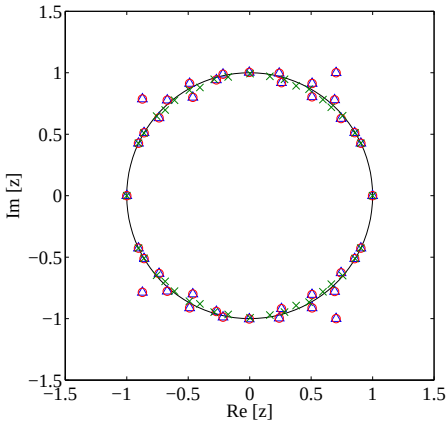
(c) Rebiorthogonalization at each 7th step

Figure 6.5: Influence of rebiothogonalization on the biorthogonal basis generated by the LANCZOS algorithm.

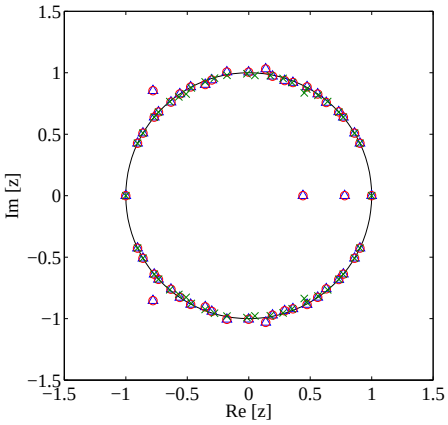


(a) after 20 steps

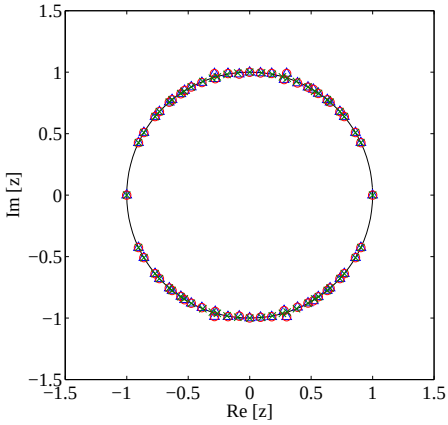
(b) after 30 steps



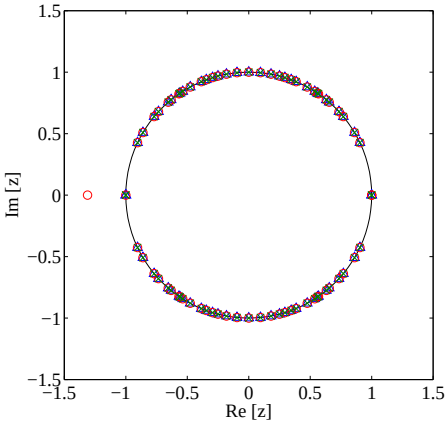
(c) after 40 steps



(d) after 50 steps



(e) after 60 steps



(f) after 70 steps

Figure 6.6: Convergence of eigenvalues extracted via the ARNOLDI (crosses), nonsymmetric (circles) and S-symmetric (triangles) LANCZOS algorithms.

6.2 Ideal Waveguide Resonator

In order to test the accuracy and convergence of the proposed MOR and evaluate its computational efficiency we applied it to the eigenmode analysis of an ideal resonator based on the WR-1800 waveguide. More specifically, taking advantage of the symmetry of the structure, only one eighth of the resonator was analyzed, as depicted in (Fig. 6.7). The front and left walls of the eighth are ideal magnetic, other walls are ideal electric, when the H_{10n}-mode is to be considered. The modeled domain has been discretized using a $20 \times 10 \times 20$ TLM grid of sizes along the $x-, y-$ and $z-$ directions of $\Delta x = \Delta y = 11.43$ mm, $\Delta z = 9.94$ mm.

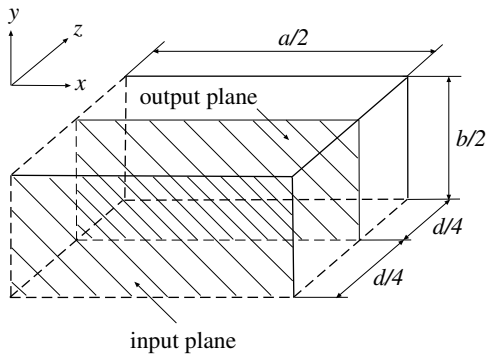


Figure 6.7: Simulated eighth of the ideal resonator.

The Arnoldi, nonsymmetric and S-symmetric Lanczos algorithms

This structure is simulated by the ARNOLDI, non-symmetric and S-symmetric LANCZOS algorithms, and results are compared with the resonator spectrum obtained by the FOURIER transform of the TLM time response. The E_y-field component is observed.

In order to test the accuracy of the implemented MOR-techniques, the resonance behavior of the resonator is computed as follows:

- a reduced-order model approximating the transfer function for the E_y-component from the input plane to the output plane with an optimal order of 14 is generated;
- time response of the resonator is recovered in the output plane ;
- the resonator spectrum is obtained by the FOURIER transform of the time response in the output plane.

Fig. 6.8 shows the comparison between TLM and the S-symmetric LANCZOS process. The spectrum of the resonator obtained by the ARNOLDI procedure is presented in Fig. 6.8. We remark that with the small number of KRYLOV subspace method steps it is impossible to look ahead, which resonances will be obtained first. Fig. 6.8(a) and Fig. 6.8(b) depicted after 150 S-symmetric LANCZOS and ARNOLDI steps have been performed, are good examples of this point.

The calculated spectra are in excellent agreement with resonance frequencies calculated in an analytical way.

As the second start vector for the non-symmetric LANCZOS algorithm the vector $|\tilde{\boldsymbol{w}}_1\rangle = \boldsymbol{S}|\tilde{\boldsymbol{v}}_1\rangle$ has been chosen, which is similar to the S-symmetric version. We only note that if the second start vector is randomly generated, more algorithm steps have to be performed in order to achieve the convergence.

From the whole eigenlaue spectrum (Fig. 6.9) only 14 eigenvalues in the complex plane are necessary in order to reconstruct the spectrum of the resonator in the frequency band $0 - 5.5$ GHz.

The check of norms of residual vectors is used as a stopping criterion. The process is stopped when the norms of residual vectors for approximating eigenvalues have achieved 10^{-4}.

The BiConjugate Gradient Based S-symmetric Lanczos Algorithm

In order to solve the equation

$$|\boldsymbol{v}_n\rangle = (\boldsymbol{A} - z_0\boldsymbol{I})|\tilde{\boldsymbol{v}}_{n+1}\rangle \tag{6.3}$$

with respect to $|\tilde{\boldsymbol{v}}_{n+1}\rangle$ for $|\boldsymbol{v}_n\rangle = |\boldsymbol{v}_1\rangle_{start}$ by an iterative projection, BCG-LTLM-S requires 489 steps at the error $\epsilon = 10^{-4}$. However, the convergence for eigenvalues approximating the resonator in the frequency band $0 - 5.5$ GHz is already achieved at 420 steps with the norms of residual vectors about 10^{-4}. Thus, we do not need to solve eq. (6.3) by further projections, and we can extract the eigenvalue spectrum from the first iterative solution. That allows us to save computational time and memory consumption essentially, because the biorthogonal basis generated during the solution of (6.3) for $|\tilde{\boldsymbol{v}}_2\rangle$ must not be stored for further oblique projection (to obtain solutions for $|\tilde{\boldsymbol{v}}_3\rangle$, $|\tilde{\boldsymbol{v}}_4\rangle$, ..., $|\tilde{\boldsymbol{v}}_n\rangle$), which is actually omitted. In this case the BCG-LTLM-S variant is similar to LTLM-S (with another start vector), and can be used for an alternative and/or additional check of the convergence of LTLM-S by checking the convergence of $|\tilde{\boldsymbol{v}}_2\rangle$ at each iterative step. On the other hand, BCG-LTLM-S requires more computations and memory for one algorithm step than LTLM-S. Additionally, in this case the convergence of BCG-LTLM-S is not dependent on the location of the expansion point z_0. A comparison of resonator spectra obtained by the TLM time evolution scheme and BCG-LTLM-S is shown in Fig. 6.10.

The Shift-Inverse Approach Using the S-symmetric Lanczos Algorithm

In order to obtain seven approximating eigenvalues the Shift-Inverse approach based on the S-symmetric LANCZOS process (SI-LTLM-S) is applied. The algorithm extracts eigenvalues closed to an expansion point z_0 first. For this purpose we have to solve the equation eq. (6.3) with $|z_0| > 1$ for each algorithm step n by a simple iterative way.

The spectrum of the resonator and approximating poles obtained via 44 SI-LTLM-S steps for $|z_0| = 1.1$ and 166 steps for $|z_0| = 2.0$ are shown in Fig. 6.13 and Fig. 6.12 respectively. To solve eq. (6.3) when $\epsilon = 10^{-6}$ and $|z_0| = 1.1$, 155 iterations are needed; and when $|z_0| = 2.0$ only 23 iterations are necessary. If the expansion point is chosen nearby the unit circle, SI-LTLM-S becomes very expensive. And on the other hand, the expansions around a remote point

exhibit the higher approximation error (see TABLE 6.4). The norms of the left and right residual vectors are large, in some cases up to 10^{-1}. On the other hand, if we change the stopping criteria to the smallest residuals, in which the order of norms is 10^{-4} or less, then additional SI-LTLM-S steps can lead to its crash due to the fast biorthogonality loss. Numerous simulations have shown that SI-LTLM-S is very critical to the number of algorithm steps and suffers from instability.

The S-symmetric Lanczos and the Padé-via-Lanczos Algorithms

The S-symmetric LANCZOS process generates a reduced matrix having the tridiagonal form. In order to get the eigenvalues of this matrix, the EigenValue Decomposition (EVD) can be utilized, but that can be very computationally expensive if the pre-reduced tridiagonal matrix has a large order. Admittedly, we often need a few poles approximating our electromagnetic system in a specific frequency band. For this reason we can only extract eigenvalues which are in a sector of the unit circle corresponding to the frequency band $0 - 5.5$ GHz. The PADÉ-via-LANCZOS (PVL) approximation and the Shift-Inverse-via-S-symmetric-LANCZOS (SI-LTLM-S) approach are very suitable for this task. Fig. 6.14 shows the eigenvalue spectrum gained via 30 PVL steps, and compared with the eigenvalues obtained by EVD. Even though the expansion point can be located inside or on the unit circle in comparison to the SI-LTLM-S, the best results are achieved if $|1 - z_0| \approx 0.1$.

Table 6.4: Comparison of simulation approaches for the ideal waveguide resonator (Fig. 6.7).

Method[1]	Steps, N	CPU time, %	Memory[2], %	error, %
TLM	2 500	100	100	-
ATLM	360	22× 100	N× 100	1
LTLM	417	65	600	3.5
LTLM-S	410	50	400	3.5
BCG-LTLM-S	420	140	11 × 100	3.5
SI-LTLM-S$_{\|z_0\|=2.0}$	116 (× 23)	340	600	12.5
SI-LTLM-S$_{\|z_0\|=1.1}$	44 (× 155)	910	600	4.5
LTLM-S + PVL	410 (+ 30)	50	400	3.5

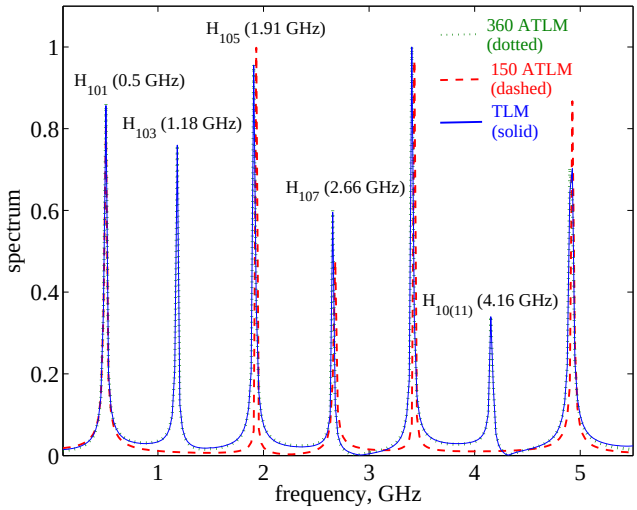
[1] ATLM - ARNOLDI-based TLM-MOR, LTLM - non-symmetric LANCZOS-based TLM-MOR, LTLM-S - S-symmetric LANCZOS-based TLM-MOR, BCG-LTLM-S - BiConjugate Gradient TLM-MOR based on the S-symmetric LANCZOS algorithm, SI-LTLM-S - Shift-Inverse-based TLM-MOR using the S-symmetric LANCZOS algorithm, LTLM-S + PVL - the first - S-symmetric LANCZOS-based TLM-MOR and the second - Pade-via-LANCZOS-based TLM-MOR

[2] Memory needed for one algorithm step

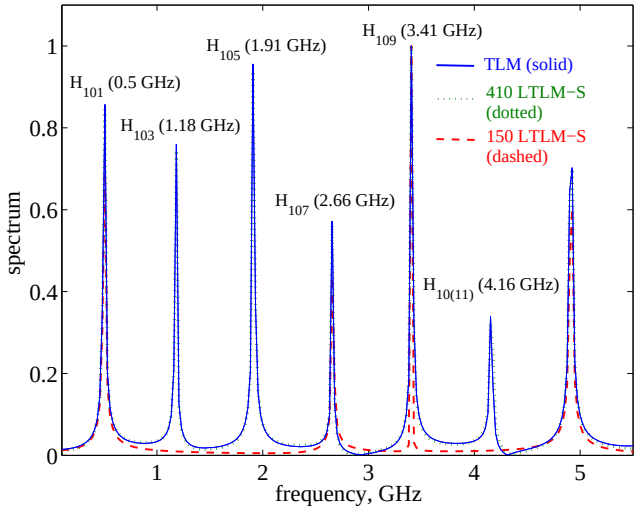
Concluding Remarks

In conclusion we give the comparative studies for the proposed MOR-techniques. A short summarization is presented in TABLE 6.4.

From all proposed MOR-techniques the ARNOLDI algorithm exhibits the smallest error. However, huge computation time and memory requirements have to be taken into account. The LANCZOS procedures converge slower and allocate bigger errors in comparison to ARNOLDI due to a growing loss of biorthogonality with the increasing number of steps. The S-symmetric version of the LANCZOS algorithm is assumed to spare computational resources, but results in a slightly bigger error, also due to the loss of biorthogonality. For this simple structure, the S-symmetric LANCZOS algorithm is 15 % faster than the classical non-symmetric LANCZOS and 44 times faster than the ARNOLDI algorithm. In addition, the S-symmetry requires less memory versus the general non-symmetric approach, because we do not need to store two previous left vectors. The utilization of very expensive rebiorthogonalization in the LANCZOS methods does not significantly improve the approximation quality. The usage of PVL and SI-LTLM-S for post-processing allows the extraction of a few eigenvalues approximating the resonator in a needed frequency band and presents an alternative to the use of standard EVD. The biconjugate-based LTLM-S algorithm provides alternative and/or additional stopping criteria; however, its resource requirements make it less attractive. The absolute outsider is the sift-inverse approach with a simple iterative scheme due to its approximation error, instability and high computational time.



(a) Spectrum of the resonator obtained by the time-domain TLM and ATLM



(b) Spectrum of the resonator obtained by the time-domain TLM and LTLM-S

Figure 6.8: Spectrum of the resonator (Fig. 6.7). The analytically calculated resonance frequencies (in parentheses) for $H_{10\mathrm{m}}$-modes are compared with simulation results obtained by the time-domain TLM, ATLM and LTLM-S.

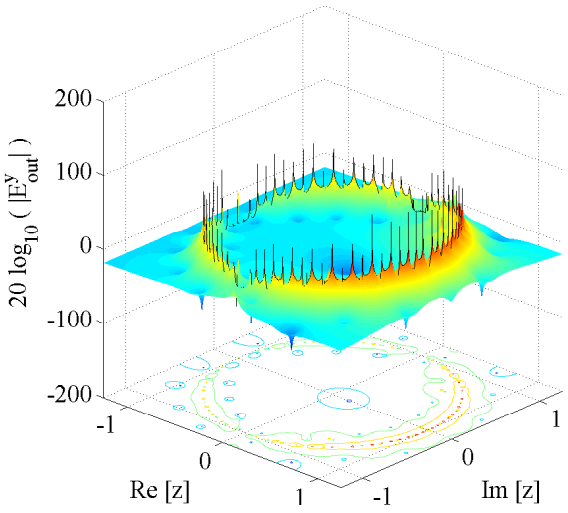


Figure 6.9: Eigenvalue spectrum of the resonator (Fig. 6.7) in 3-D-view extracted via 410 S-symetric LANCZOS algorithm steps.

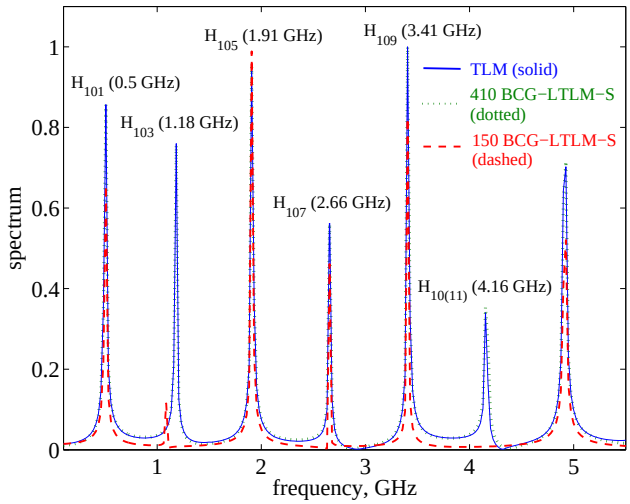
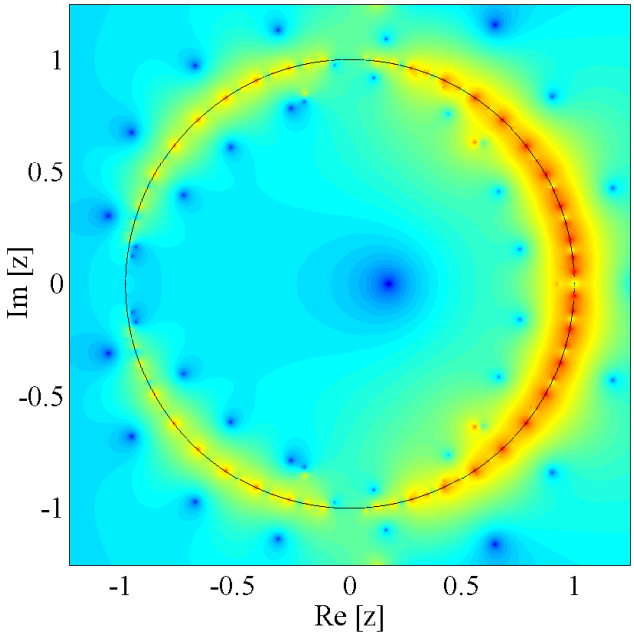
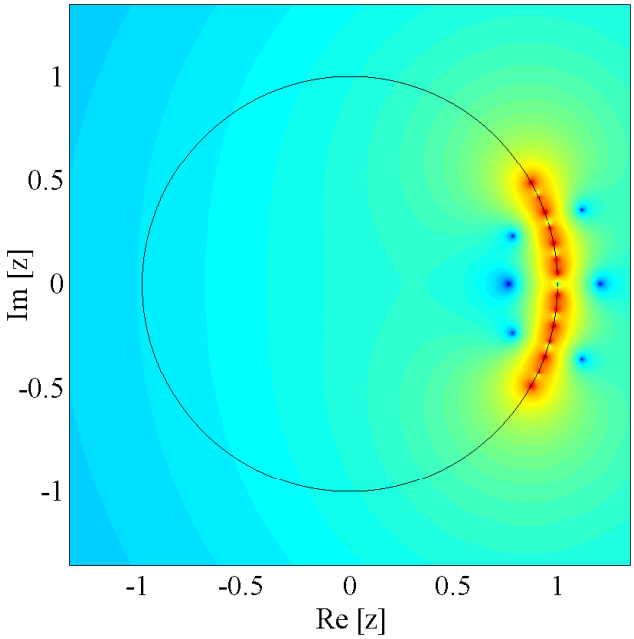


Figure 6.10: Spectrum of the resonator (Fig. 6.7) obtained by time-domain TLM and BCG-LTLM-S.

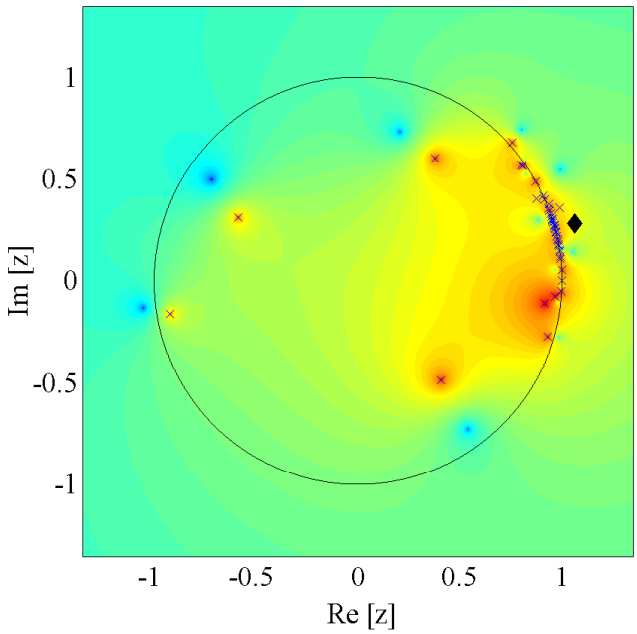


(a) Eigenvalue spectrum (410 values) without refining

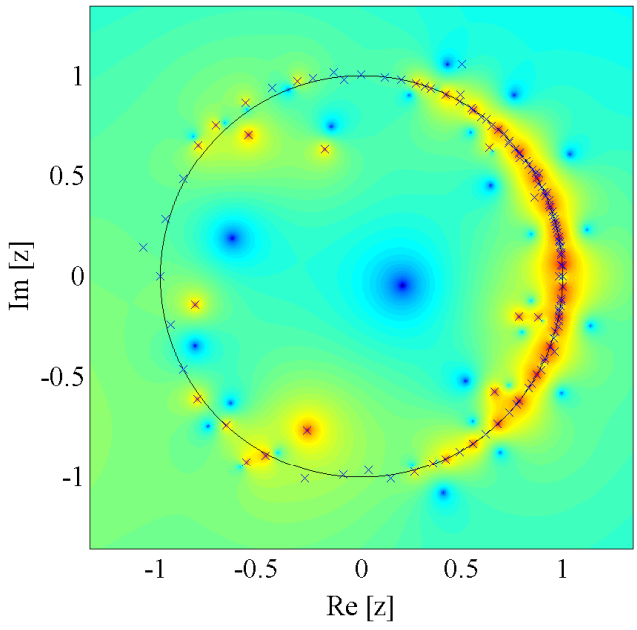


(b) Eigenvalue spectrum (14 values) after refining

Figure 6.11: Eigenvalue spectrum of the resonator (Fig. 6.7) in 2-D-view extracted via 410 S-symetric LANCZOS algorithm steps with and without refining.



(a) 44 poles extracted at $|z_0| = 1.1$



(b) 116 poles extracted at $|z_0| = 2.0$

Figure 6.12: Eigenvalue spectrum of the resonator (Fig. 6.7) in 2-D-view extracted via SI-LTLM-S algorithm steps without refining. Crosses denote poles, whereas the rhombus denotes the expansion point z_0.

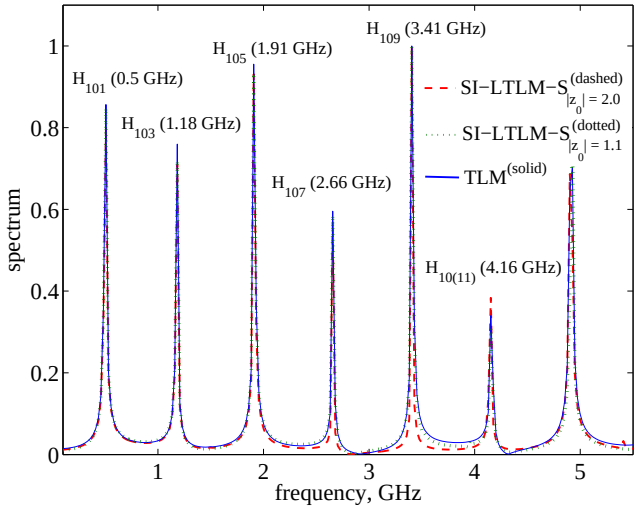


Figure 6.13: Spectrum of the resonator (Fig. 6.7) obtained by the time-domain TLM and SI-LTLM-S for $|z_0| = 1.1$ and for $|z_0| = 2.0$.

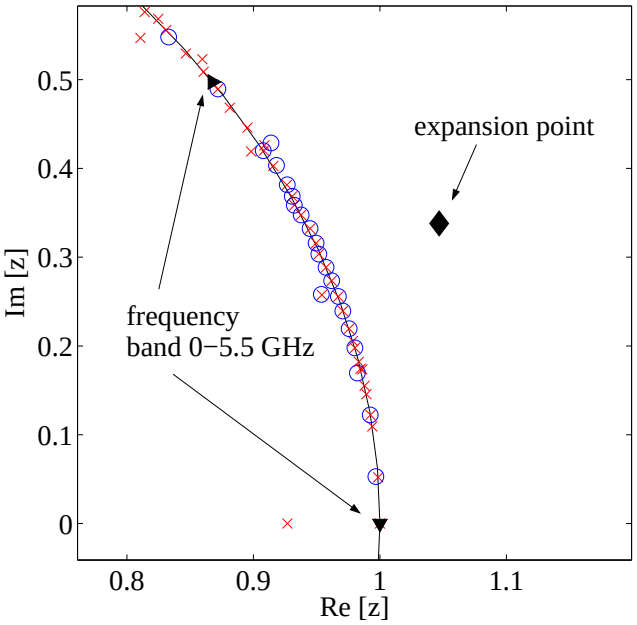


Figure 6.14: Pole distribution of the resonator (Fig. 6.7) in the complex z-plane obtained via EVD (crosses) and PVL (circles) from 410×410 reduced matrix generated by LTLM-S.

6.3 Microstrip line Z-bend

A two-port sharp microstrip line Z-bend is investigated. The dimensions of the Z-bend are shown in Fig. 6.15. The Z-bend is simulated by means of the Method of Moments (MoM), the TLM time evolution scheme, the ARNOLDI, and the nonsymmetric and S-symmetric LANCZOS MOR-algorithms.

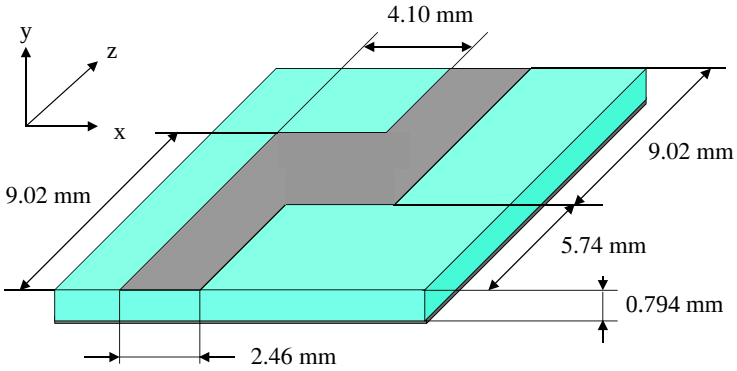
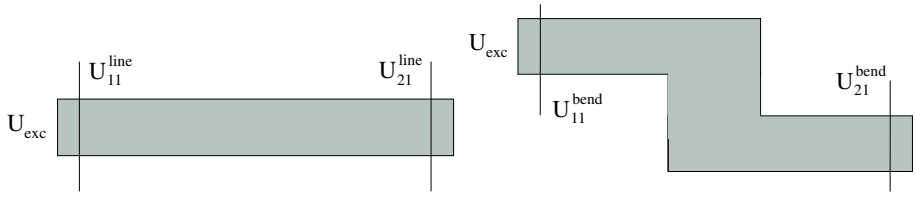


Figure 6.15: Dimensions of the Z-bend.

The used space steps are $\Delta x = 0.82$ mm, $\Delta y = 0.397$ mm, $\Delta z = 0.82$ mm, and the number of cells along x, y and z are 14, 5, and 18, respectively. The microstrip line is of the width $3\Delta x$. The substrate with the thickness $h = 0.794$ having the relative permittivity $\varepsilon_r = 2.2$ and $\tan\delta = 0.001$ is $2\Delta y$. The copper sheet in the TLM is modeled as a very thin layer with the surface impedance $R_s = 0.025\ \Omega$. In the MoM simulations the metal layer with the conductivity $\sigma_{cu} = 5.8 \cdot 10^7$ S/m is 35 µm thick.

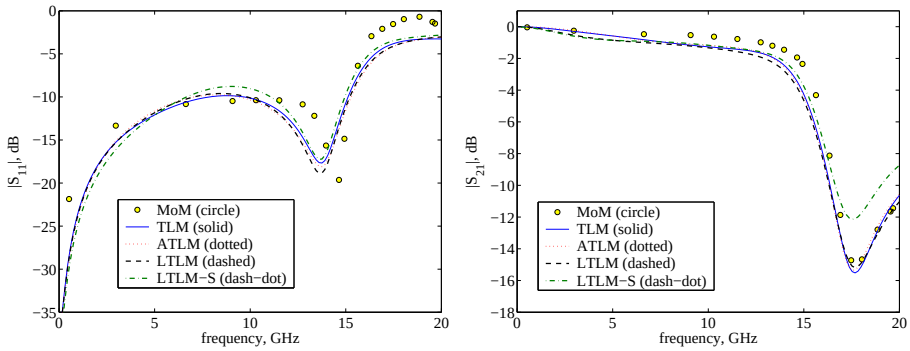
First of all, the structure is simulated by means of the time-domain TLM scheme. The Z-bend is excited by a GAUSS pulse voltage in the plane $z = 1\Delta z$ (see Fig. 6.15). The S-parameters are calculated using a microstrip "through line" of a length of 14.76 mm (see Fig. 6.16(a)). Reference planes for reflected and for transmitted waves are placed at $z = 2\Delta z$ and $z = 17\Delta z$, respectively. The Z-bend is embedded in a box with ABC as reflection-free walls.



(a) Reference microstrip line layer (b) Z-bend layer

Figure 6.16: Signal layers of the reference line and the Z-bend (Fig. 6.15).

Next, the MOR-techniques are applied to extract the S-parameters of the Z-bend. The S-parameters obtained from simulations are presented in Fig. 6.17.



(a) Magnitudes of S_{11}-parameters (b) Magnitudes of S_{21}-parameters

Figure 6.17: S-parameters for a Z-bend (Fig. 6.15) obtained by the MoM, the TLM and MOR-techniques.

The MOR-techniques give very good results compared with the TLM time evolution scheme. For this example the goal of MOR is to approximate the voltage transfer functions $\tilde{\mathcal{H}}^U_{21}$ and $\tilde{\mathcal{H}}^U_{11}$ of the Z-bend from port 2 to port 1 (transmission) and from port 1 to port 1 back (reflection), respectively and to generate a reduced-order model with an optimal order.

According to Fig. 6.16, the S-parameters of the Z-bend can be obtained as following

$$S_{11}^{TLM}(f) = \frac{U_{11}^{bend}(f) - U_{11}^{line}(f)}{U_{11}^{line}(f)}, \quad S_{11}^{MOR}(f) = \frac{\tilde{\mathcal{H}}^U_{11}(f)U_{exc}(f) - U_{11}^{line}(f)}{U_{11}^{line}(f)},$$
$$S_{21}^{TLM}(f) = \frac{U_{21}^{bend}(f)}{U_{11}^{line}(f)}, \quad S_{21}^{MOR}(f) = \frac{\tilde{\mathcal{H}}^U_{21}(f)U_{exc}(f)}{U_{11}^{line}(f)}. \tag{6.4}$$

A brief comparison between simulation approaches with respect to computational efforts are coalesced in TABLE 6.5. Since the investigated structure does not indicate high resonance behavior, all MOR-techniques do not expedite the simulation processes in comparison to the time-domain TLM. The number of algorithm steps is chosen to minimize the approximation error.

Table 6.5: Comparison of simulation approaches for the microstripline Z-bend (Fig. 6.15).

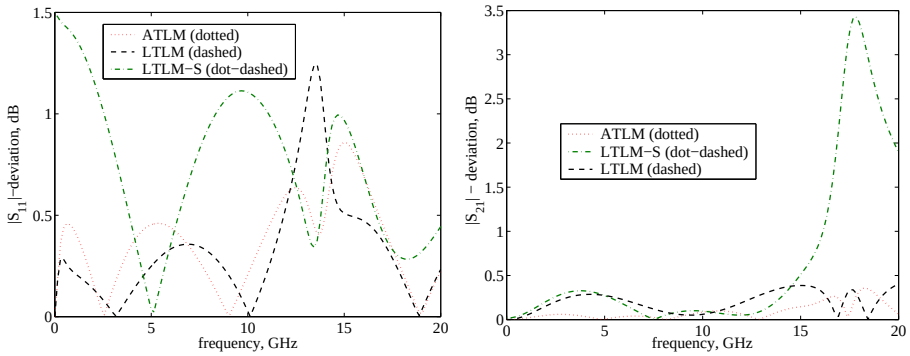
Method	TLM	ATLM	LTLM	LTLM-S
Steps for $\tilde{\mathcal{H}}^U_{11}$, N	1 000	240	290	270
Steps for $\tilde{\mathcal{H}}^U_{21}$, N	1 000	230	250	250
CPU time[1], %	100	26 × 100	135	105

[1] The elapsed time is referred to the largest number of algorithm steps.

Though LTLM-S has the computational time effort similar to TLM, this algorithm gives a bigger error for S-parameters in comparison to the ARNOLDI and to the nonsymmetric LANCZOS procedures. This behavior can be explained that the S-symmetric algorithm loses biorthogonality faster than the nonsymmetric LANCZOS algorithm, because the left vectors in LTLM-S are constructed

without accounting their previous ones. The left start vectors in the LANCZOS algorithms are chosen as one time scattered excitation vectors. The loss of biorthogonality in LTLM-S causes an error accumulation and, as following, an instability in the results. After 300 steps the algorithm crashes, which does not occur in ATLM and LTLM. The ATLM algorithm is proven to be very stable.

Fig. 6.18 indicates deviations of the S-parameters obtained by the MOR-techniques from S-parameters gained via the FOURIER transforms of the TLM time responses. The ATLM algorithm features a slightly lesser error than LTLM. However, ATLM is an "absolute outsider" if we compare the computational time.



(a) Deviations of S_{11}-magnitudes (b) Deviations of S_{21}-magnitudes

Figure 6.18: Deviation of S-parameters obtained by MOR-techniques for a Z-bend (Fig. 6.15) obtained by MOR-techniques.

Table 6.6: Dependance of the Z-bend (Fig. 6.15) model$^{1\to 2}$ order and S_{21}-deviation on SNR.

SNR, dB	Model order	Deviation, dB
∞	64	0.44
100	62	0.44
90	62	0.44
80	60	0.44
70	56	0.44
60	50	0.46
50	46	0.47
48 (SNR_{opt})	**44**	**0.40**
40	40	0.89
30	24	2.40

In order to approximate $\tilde{\mathcal{H}}_{21}^{U}(f)$- and $\tilde{\mathcal{H}}_{11}^{U}(f)$-transfer functions, 64 and 70 poles/residues are required, respectively. The poles are obtained by means of the eigenvalue decomposition, followed by poles selection in a sector of the unit circle

Table 6.7: Dependance of the Z-bend (Fig. 6.15) model$^{1\to1}$ order and S_{11}-deviation on SNR.

SNR, dB	Model order	Deviation, dB
∞	70	1.25
100	68	1.25
90	64	1.25
80	60	1.22
70	56	1.29
60	46	1.24
58 (SNR_{opt})	**44**	**1.15**
55	40	1.27
50	36	1.42
40	28	3.5

corresponding to the frequency band $0-20$ GHz. Fig. 6.19 exhibits the pole-residue distribution of the voltage transfer function $\tilde{\mathcal{H}}_{21}^{U}$ in the complex z-plane (The solid line corresponds to the unit circle). The eigenvalues approximating the transfer functions $\tilde{\mathcal{H}}_{21}^{U}$ and $\tilde{\mathcal{H}}_{11}^{U}$ are concentrated in a sector of the circle between $-42°$ and $42°$ with $0.9 < |z| < 1.0$. We note that two generated reduced-oder models are characterized by their relatively high order.

The model's order can be additionally reduced by introducing a so-called threshold. We can calculate the SNR for the residues and compare it with the threshold. The residues, passing over the threshold, form the order of the reduced model, and no trespassed residues/poles are discarded. The non-symmetric version of the LANCZOS algorithm is selected for Model Order Selection (MOS) investigations. Strictly speaking, this is an optimal order of the macro-model obtained by the nonsymmetric LANCZOS process, selected using SNR. The dependance of the model orders and S-parameter deviations on the SNR is presented in TABLES 6.6 and 6.7. The orders of the optimal reduced models are chosen at $\mathrm{SNR} = \mathrm{SNR}_{opt}$ if the deviations of the S-parameters are minimal. In this case the orders of the one-port models are reduced down to 44 and the order of the two-port microstrip line Z-bend is arranged to be 88 at the initial state vector dimension of $N = 22\,680$. Thus, the corresponding reduction factor for the two-port model is 257, and for the one-port model is 515, respectively. Fig. 6.20 shows the influence of SNR values on the accuracy of computed S-parameters, and depicts the S_{21}-parameters and their deviations from the original TLM simulation.

Poles positions in the complex z-plane for the microstrip line Z-bend for $\mathrm{SNR} = \infty$ and $\mathrm{SNR} = \mathrm{SNR}_{opt}$ cases are shown in Fig. 6.21. Comparing these figures we see that several computed eigenvalues overlap; however, most of them are distant from each other with some tolerance and gathered in a cluster or in a pair. Furthermore, some of these poles are very difficult to distinguish when they are in a pair or in a cluster.

The pole clustering allows the reduction of a multi-port macro-model order by searching the most likely values among those found in the port transfer

responses. Pole clustering, however, is a very critical operation which strongly effects the model accuracy. The store of all poles/residues for each port is recommended, even though in this case the order of the generated multi-port macro-model will be larger. Excluding the overlapped poles (see Fig. 6.21(b)), the two-port ROM has an order of 77.

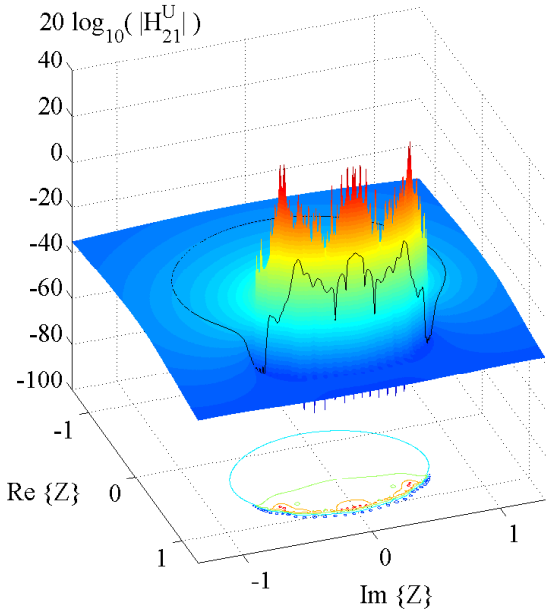
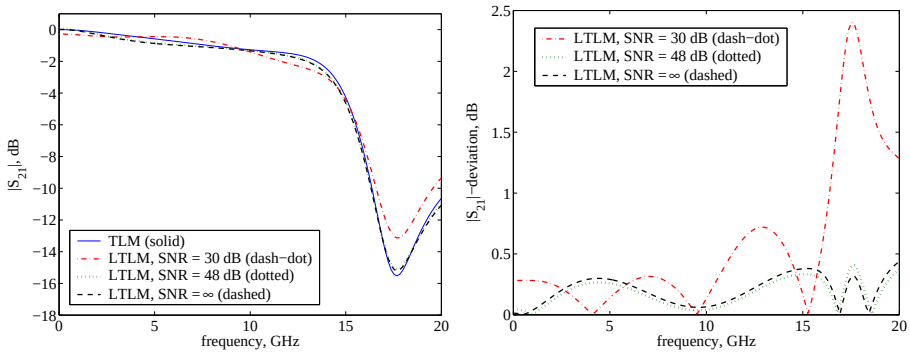


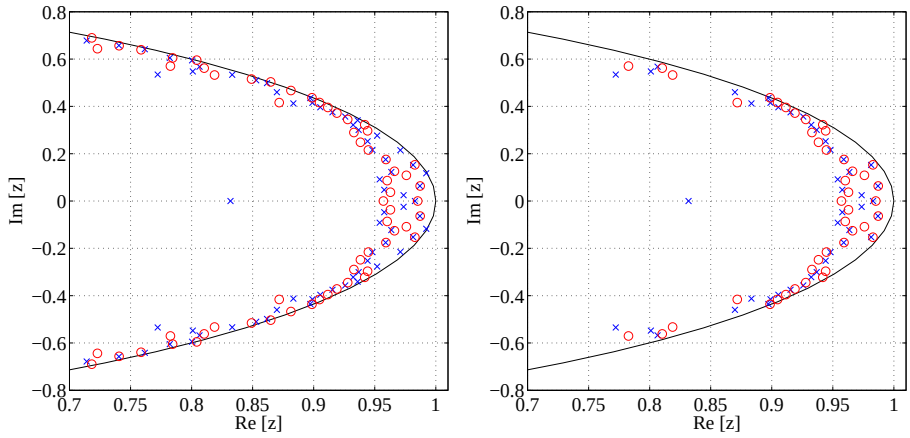
Figure 6.19: Pole-residue distribution of voltage transfer function from port 1 to port 2 in the complex z-plane with SNR $= \infty$ for the microstrip line Z-bend (Fig. 6.15).

From our investigations we can conclude that the nonsymmetric LANCZOS algorithm is optimal for the generation of the two-port macro-model of the Z-bend in view of errors and elapsed time. The order of the extracted reduced macro-models can be additionally decreased by introducing a SNR threshold. In this particular case the order of the two-port model is secondary reduced from 134 down to 88, where the approximation errors are minimized. Even though the pole overlapping and clustering allow us to additionally decrease the order of the model, they are very critical operations and can affect the model accuracy and stability (for more detail see [120]).



(a) S_{21}-magnitudes (b) Deviations of S_{21}-magnitudes

Figure 6.20: Dependance of S_{21}-parameters and their deviations on SNR thresholds for the microstrip line Z-bend (Fig. 6.15).



(a) Poles with SNR $= \infty$ (b) Poles with SNR $=$ SNR$_{opt}$

Figure 6.21: Poles position in the complex z-plane for a microstrip line Z-bend (Fig. 6.15); circles denote poles responding for $\tilde{\mathcal{H}}_{21}^{U}$ approximation; crosses denote poles responding for $\tilde{\mathcal{H}}_{11}^{U}$ approximation; solid line corresponds to the unit circle.

6.4 Coplanar Waveguide Impedance Transition

In this section a coplanar waveguide impedance transition is analyzed. Step changes in the width of the center strip conductor of a Coplanar Waveguide (CPW) give changes of the wave impedance. For the structure depicted in Fig. 6.22, the middle CPW-section of a length of 4 mm has a lower wave impedance than the first and the third CPW-sections of a length of 2 mm.

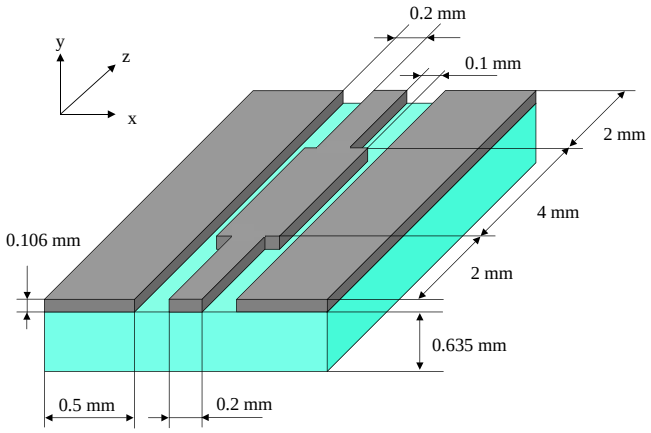


Figure 6.22: Dimensions of the CPW impedance transition.

Due to the symmetry of the structure in respect to the yz-plane only half of the impedance transition is simulated. Since the even mode has to be considered we can put an ideal magnetic wall in the middle of the strip line. Further, we will consider the right half of the CPW wave impedance transition with dimensions shown in Fig. 6.22. The half of the structure is embedded in a box with ABC (reflection-free walls), except for the left wall, which is perfectly magnetic.

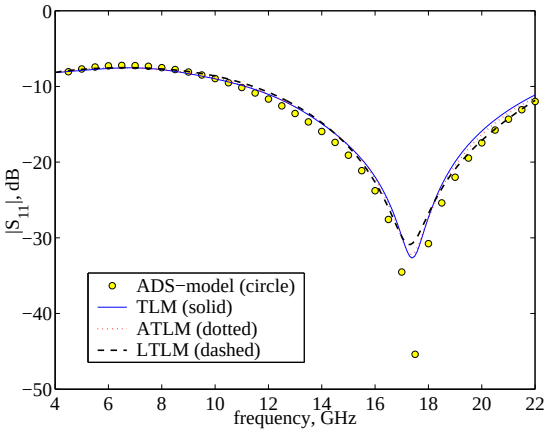
The transition is simulated with the space steps $\Delta x = 0.1$ mm, $\Delta y = 0.106$ mm, $\Delta z = 0.4$ mm; the number of cells along x, y and z are 8, 15 and 22, respectively. The substrate with the thickness $h = 0.635$, having the relative permittivity $\varepsilon_r = 12.9$ and $\tan\delta = 10^{-4}$ is modeled with $6\Delta y$. The copper conductors are $1\Delta y$, i.e. one node thick. The external and middle signal strips have widths of $1\Delta x$ and $2\Delta x$, respectively. The ground plane is $5\Delta x$ in width.

First, the impedance transition is simulated by the time-domain TLM method with an excitation as the GAUSS pulse voltage in the plane $z = 1\Delta z$ (see Fig. 6.22). The S-parameters are calculated using a "through" CPW of a length of 8.8 mm. Reference planes for reflected waves are placed at $z = 3\Delta z$.

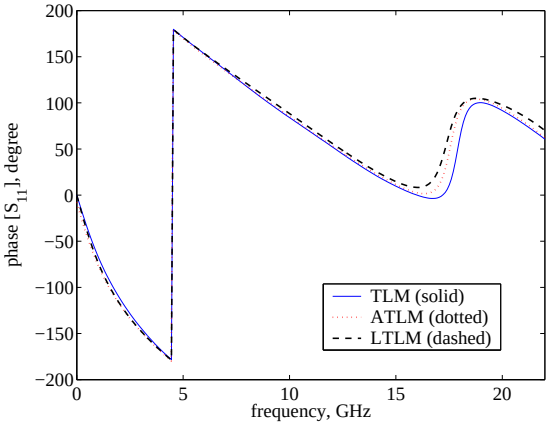
Then, in order to obtain the S-parameters using MOR-techniques we approximated the transfer function $\tilde{\mathcal{H}}_{11}^{U}$ for the reflected wave (from port 1 to port 1 back), i.e.

$$S_{11}^{TLM}(f) = \frac{U_{11}^{trans}(f) - U_{11}^{line}(f)}{U_{11}^{line}(f)}, \quad S_{11}^{MOR}(f) = \frac{\tilde{\mathcal{H}}_{11}^{U}(f)U_{exc}(f) - U_{11}^{line}(f)}{U_{11}^{line}(f)}. \tag{6.5}$$

Altogether, the impedance transition is simulated by the time-domain TLM scheme, ARNOLDI, non-symmetric LANCZOS, and S-symmetric LANCZOS algorithms. Fig. 6.23 shows the S-parameters obtained from simulations compared with results provided by the ADVANCED DESIGN SYSTEM (ADS)-model. All simulations are in a very good agreement.



(a) Magnitudes of S_{11}-parameters



(b) Magnitudes of S_{21}-parameters

Figure 6.23: S-parameters for a CPW impedance transition (Fig. 6.22) obtained by the ADS-model, TLM and MOR-techniques.

The ARNOLDI algorithm yields a very good approximation of the initial system discretized by TLM. The non-symmetric LANCZOS algorithm features some error, especially for higher frequencies. The obtained results are in agreement with the modeling by means of ADS. The second start vector $\tilde{\boldsymbol{w}}_1$ in LTLM is taken to be filled with ones. A choice of the left start vector as the scattered right start vector, i.e. $\tilde{\boldsymbol{w}}_1 = \boldsymbol{S}\tilde{\boldsymbol{v}}_1$ gives rise to an unacceptable error. The right start vectors $\tilde{\boldsymbol{v}}_1$ are taken as the excitation vector.

Convergence cannot be achieved using the S-symmetric LANCZOS procedure, because the algorithm is obviously destroyed due to the radical loss of the biorthogonality.

The whole discrete system is observable as a bad-convergent system, which can be explained by the excitation operation. Only two nodes in the CPW gap are excited with E_x field components corresponding to the excitation voltage U_{exc}.

A comparison between the time-domain TLM, ARNOLDI and LANCZOS algorithms is outlined in TABLE 6.8. The implemented MOR-techniques do not expedite TLM simulations; moreover, they are slower in comparison to them.

Table 6.8: Comparison of simulation approaches for the impedance transition (Fig. 6.22).

Method	Steps[1], N	CPU time, %	Order of the model
TLM	2 500	100	47 520
ATLM	1 000	very large	140
LTLM	1 120	370	156

[1] The time needed for the generation of a completed reduced model including the post-processing.

Limiting the number of approximating residues/poles by means of SNR, we may secondary reduce the model orders down to 112 at $\mathrm{SNR}_{opt} = 55$ dB for both cases in TABLE 6.8.

From the above described results and investigations, we can firstly conclude that the nonsymmetric LANCZOS algorithm gives an acceptable error at an acceptable elapsed time. Secondly, the ARNOLDI approach approximates the CPW impedance transition discrete system very precisely; however, its computation time is a big drawback and makes the algorithm unattractive. The generated macro-model is characterized by its large order, even after the optimal residues/poles selection is made. The achieved reduction factor for the TLM-operator is only 424.

6.5 Rectangular Microstrip Patch Antenna

Next, a rectangular microstrip patch antenna has been investigated. The antenna design and the measurement are borrowed from [108]. The actual dimensions of the microstrip patch antenna are shown in Fig. 6.24.

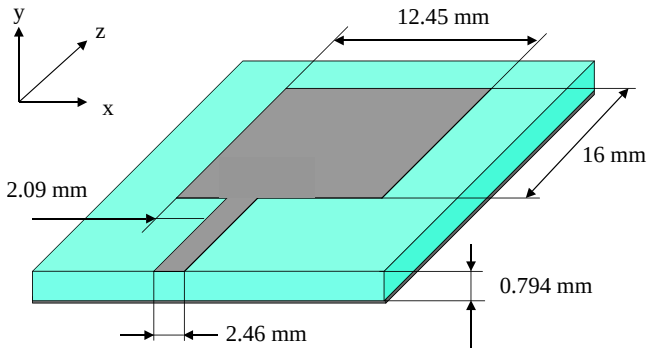


Figure 6.24: Schematic view of the rectangular microstrip patch antenna.

The space steps used are $\Delta x = 0.389$ mm, $\Delta y = 0.400$ mm, $\Delta z = 0.265$ mm, and the number of cells along x, y and z are, respectively, 60, 16 and 100. The rectangular path is $32\Delta x \times 40\Delta z$. The feed microstrip line of the width $6\Delta x$ is $50\Delta z$ long. The substrate thickness is modeled with $3\Delta y$, the additional 13 nodes in y-direction are chosen to model the free space. The copper layer is modeled as a very thin sheet with the surface impedance $R_s = 0.025\ \Omega$.

The patch antenna is excited by a GAUSS pulse voltage in the plane $z = 2$ (Fig. 6.24). Using the time-domain TLM scheme the simulation is performed for 6000 time steps due to a resonant behavior of the antenna; after that, 1380 S-symmetric Lanczos steps are done. The reference plane for reflected waves is defined by $10\Delta z$ from the antenna path. The S-parameters of the antenna are computed through obtained impedances as

$$S_{11} = \frac{Z_{in} - Z_{ref}}{Z_{in} + Z_{ref}}\,, \tag{6.6}$$

where the input impedance is calculated as

$$Z_{in}^{TLM}(f) = \frac{U_{11}^{ant}(f)}{I_{11}^{ant}(f)}\,, \quad Z_{in}^{MOR}(f) = \frac{\tilde{\mathcal{H}}_{11}^{UU}(f)}{\tilde{\mathcal{H}}_{11}^{IU}(f)}\,, \tag{6.7}$$

here $U_{11}^{ant}(f)$ and $I_{11}^{ant}(f)$ are the reflected voltage and current obtained by the FOURIER transform of the TLM time-domain responses; $\tilde{\mathcal{H}}_{11}^{UU}$ and $\tilde{\mathcal{H}}_{11}^{IU}$ are transfer functions of $U_{exc}(f)$ to the voltage and to the current in the reference plane, respectively, which are approximated by LTLM-S. The reference impedance Z_{ref} is taken 50 Ω.

The comparison between measurement and numerical solutions is presented in Fig. 6.25. The S-symmetric LANCZOS algorithm yields a good approximation, especially for the operating resonance at 7.5 GHz. In this case the LTLM-S

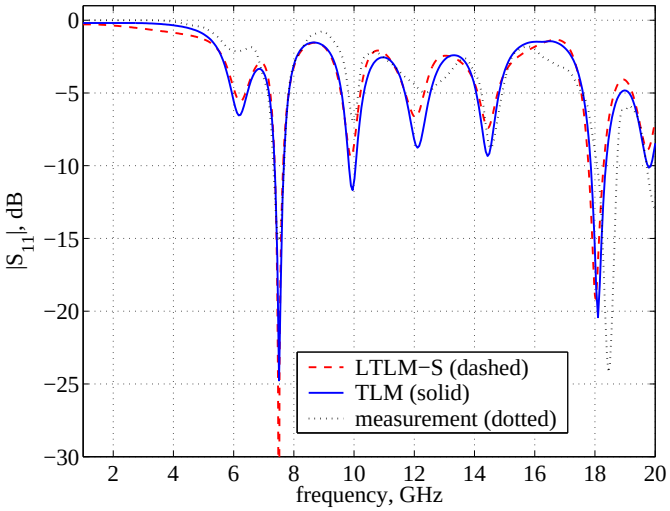


Figure 6.25: Return loss of the rectangular microstrip patch antenna (Fig. 6.24).

algorithm is about 50 % faster (inclusive post-processing routines) than using the FOURIER transform of the TLM time responses.

From the extracted 1380 eigenvalues, only 314 are needed for an approximation of the patch antenna in the frequency band $0 - 20$ GHz. The approximating eigenvalues are located in a sector of circle in the complex plain between $-46°$ and $46°$ with $0.85 < |z| < 1.0$ (see Fig. 6.26). Thus, the order of the generated macro-model is 314 for an initial state vector dimension of $N = 1\,728\,000$, which corresponds to the reduction factor of more than 5 000.

Using a SNR threshold of 40 dB, we can additionally reduce the model order down to 248; in this case the reduction factor is about 7 000. The positions of the eigenvalues passed over the threshold are shown in Fig. 6.27. Comparing Fig. 6.26 and Fig. 6.27 we can see that the passed eigenvalues are slightly rarefied, that we can especially observe at sector boundaries.

Next, we compare models extracted by MOR-techniques and System Identification (SI) approaches. Simulations of the patch antenna (Fig. 6.24) by PRONY Model based Signal Identification (PMSI) is given in [27]. In order to extract the system poles (and to approximate the antenna impedance) from the time-domain transient responses obtained by the time-domain TLM scheme, two main approaches have been considered: the polynomial Last Square (LS) approximations [79] and the pencil matrix (PM) method [57]. The order of the model obtained by PMSI is significantly lower than the order of the reduced-order model generated by S-LTLM. Pole positions of the models obtained by SI are shown in Fig. 6.28. In [27] it has been shown that the PM model of order 30 yields a very good approximation of the S-parameters in the frequency band $0 - 20$ GHz.

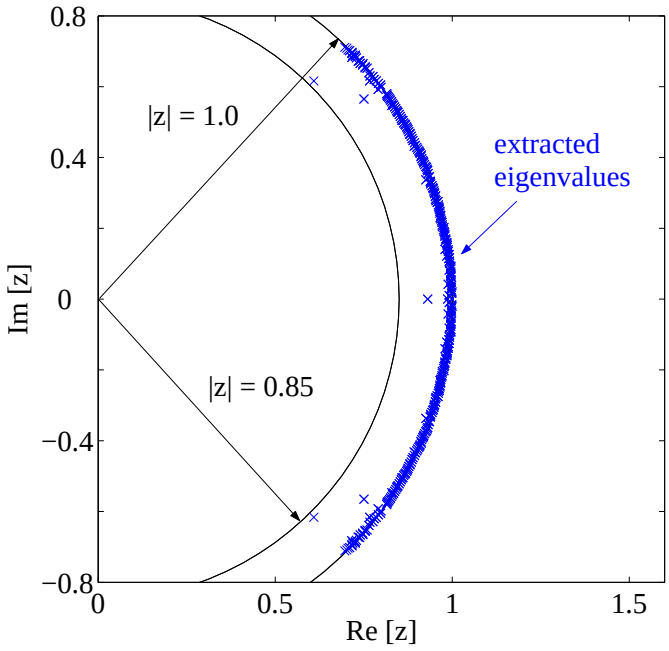


Figure 6.26: Extracted eigenvalues approximating the rectangular microstrip patch antenna (Fig. 6.24) in the frequency band $0 - 20$ GHz.

A short comparison between LTLM-S and PMSI for the analysis of the patch antenna (Fig. 6.24) is provided in TABLE 6.9.

Table 6.9: Comparison of simulation approaches for the microstrip patch antenna (Fig. 6.24).

Method	Model order n	CPU time reduction, %
LTLM-S	248	50
PMSI	30	53

The system identification techniques as well as the model order reduction algorithms expedite the time-domain based TLM analysis by a factor of about two and allow to generate a compact equivalent model of the TLM-system. The order of the model obtained by PMSI is significantly lower than the order of the reduced-order model generated by S-LTLM.

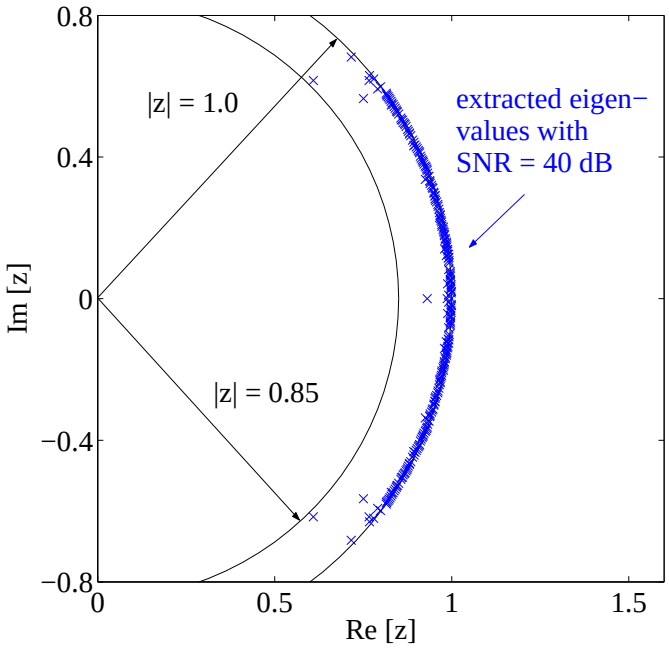


Figure 6.27: Extracted eigenvalues approximating the rectangular microstrip patch antenna (Fig. 6.24) in the frequency band $0 - 20$ GHz with SNR $= 40$ dB.

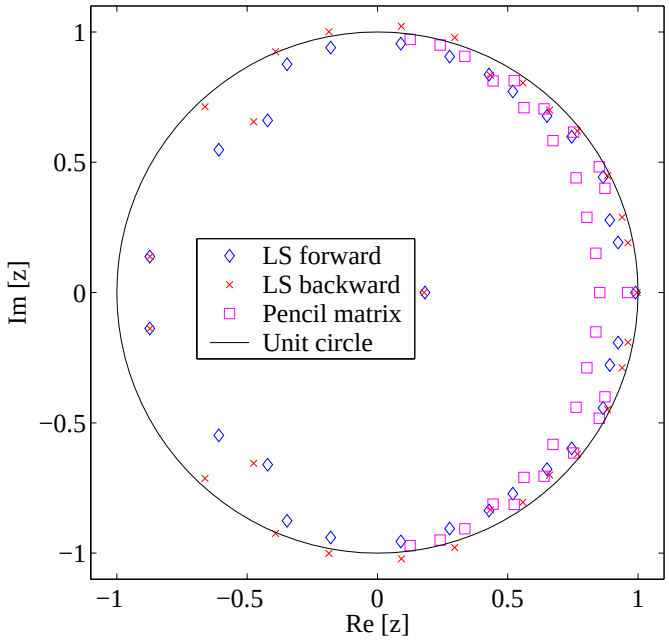


Figure 6.28: Pole positions in the z-plane for the three models obtained by PMSI for the input impedance of the patch antenna (Fig. 6.24) in the frequency band $0 - 20$ GHz. LS forw. means forward Least Square algorithm; LS back. denotes backward Least Square algorithm; Pencil Matrix is Pencil Matrix Method.

6.6 Dielectric Resonator Waveguide Filter

Fig. 6.29 shows a cutoff-coupled rectangular dielectric resonator waveguide bandpass filter. The filter presents a rectangular metal walled waveguide filled with a dielectric material with $\varepsilon_r = 10$ at its ports and in its center. Between three dielectric blocs are two air-filled sections. The lengths of the propagating dielectric blocks l_d=26.1 mm appoint the resonance frequency at about 3.102 GHz for the TE_{10} mode and the lengths of the evanescent air sections l_a=26.1 mm affect on the coupling between the propagating dielectric sections, and consequently, the loaded quality factor of the filter.

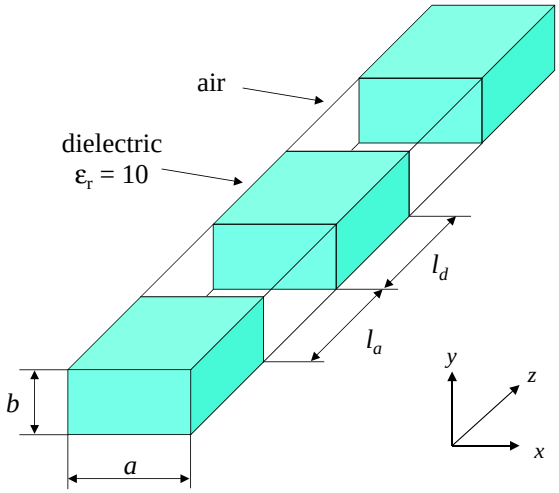


Figure 6.29: Schematic view of a cutoff-coupled rectangular dielectric resonator waveguide filter.

The length of the dielectric blocs can be obtained from the theoretical equation given in [13] as

$$d = \frac{\pi - \arctan\left(\dfrac{2\alpha\beta}{\alpha^2 - \beta^2}\right)}{\beta} = 26.105 \text{ mm}\,, \tag{6.8}$$

where α is the attenuation constant in the evanescent air sections

$$\alpha = -\jmath\sqrt{\beta_0^2 - \beta_\perp^2} = 105.248 \text{ m}^{-1}\,, \tag{6.9}$$

and β is the propagation constant in the propagating dielectric sections

$$\beta = \sqrt{\varepsilon_r\beta_0^2 - \beta_\perp^2} = 164.045 \text{ m}^{-1}\,. \tag{6.10}$$

Here $\beta_\perp = \dfrac{\pi}{a} = 123.685 \text{ m}^{-1}$ and $\beta_0 = \dfrac{2\pi f_0}{c_0} = 64.968 \text{ m}^{-1}$ at the resonance frequency $f_0 = 3.102$ GHz.

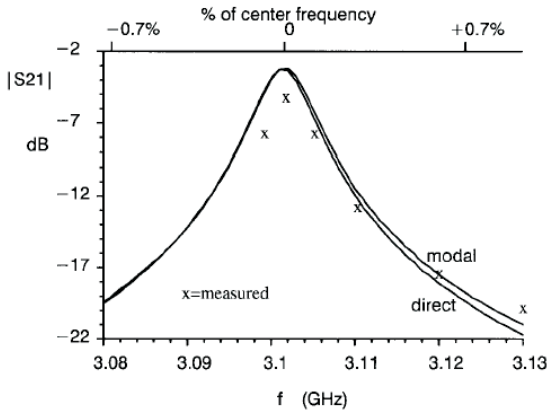


Figure 6.30: Magnitudes of the S_{21}-parameters for the cutoff-coupled rectangular dielectric resonator waveguide filter (Fig. 6.29) with l_d=26.4 mm and l_a=26.1 mm. Curves are computed by the finite-element direct and modal methods, while the crosses are measured data [13].

Simulations by means of finite-element based methods and their comparison with measurements are presented in Fig. 6.30. Here, the model includes the dielectric loss with $\tan\delta = 7 \cdot 10^{-4}$.

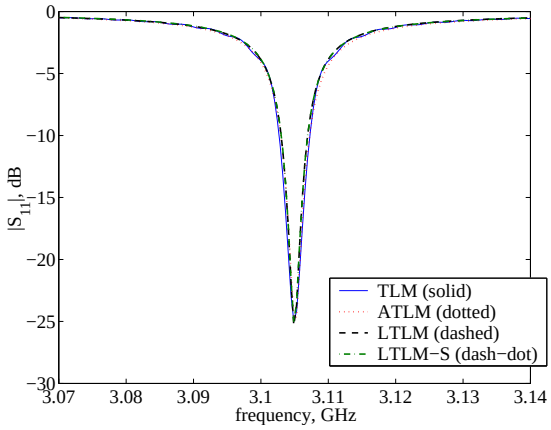
The filter is analyzed by the time-domain TLM, and the ARNOLDI, nonsymmetric and S-symmetric LANCZOS MOR-algorithms.

Due to the symmetry of the 25.4×12.7 mm waveguide only the right half of the dielectric resonator waveguide filter (12.7×12.7 mm) is analyzed. While the H_{10} wave has to be considered, we can put in the filter middle the PMC along the propagation axes; the right, top and bottom walls are PEC; the front and back walls are terminated by ABC. Each half of the section is discretized by $5 \times 5 \times 6$ nodes in x, y and z directions, respectively, with the space steps $\Delta x = 2.54$ mm, $\Delta y = 2.54$ mm, $\Delta z = 4.35$ mm. The waveguide filter is excited by the E_y-field profile corresponding to the H_{10} mode. The waveguide wall loss and dielectric loss are not included for consideration.

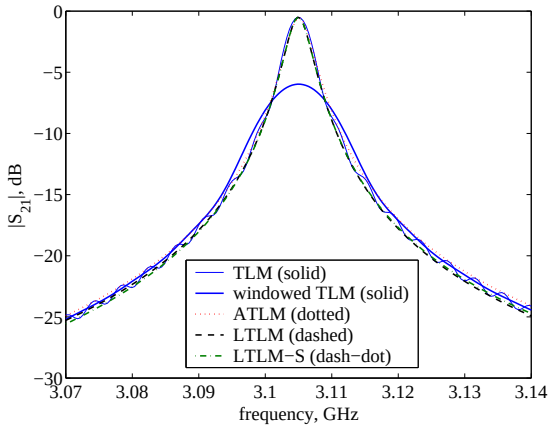
Fig. 6.31 shows the reflection and transmission coefficients $|S_{11}|$ and $|S_{21}|$ near to the resonant frequency $f_0 = 3.102$ GHz. The computed curves indicate the resonance at the frequency of about 3.105 GHz, which corresponds to the frequency error of 0.1 % from the theoretical resonance. The agreement between the time-domain TLM and MOR-algorithms is excellent, in addition TLM-based simulations coincide with the measurement and FEM results in Fig. 6.30.

The filter exhibits a high resonant behavior; therefore a large number of time steps is required. Even after 100 000 TLM time evolution scheme steps the S-parameter curves show small ripples (see Fig. 6.31). Application of the windowing to the time response eliminates the ripples but yields some error. For example, the $|S_{21}|$-magnitude in Fig. 6.31(b); after the HANNING windowing of the 50 000 time steps has been done, the TLM response is about 5 dB lower and slightly wider than the $|S_{21}|$-magnitude obtained via the FOURIER transform of the 100 000 time steps response.

The S-parameters are calculated using a reference 135 mm long waveguide filled with dielectric. The reference planes for the reflected and transmitted



(a) Magnitudes of S_{11}-parameters



(b) Magnitudes of S_{21}-parameters

Figure 6.31: S-parameters for the cutoff-coupled rectangular dielectric resonator waveguide filter (Fig. 6.29) calculated by the time-domain TLM and MOR-techniques.

waves are placed at $z = 2$ and $z = 29$. The transfer functions for the transmission (from port 1 to port 2) and reflection (from port 1 back to port 1) of E_y-field components are approximated by MOR-algorithms. The S-parameters can be obtained from simulations as

$$S_{11}^{TLM}(f) = \frac{U_{11}^{filter}(f) - U_{11}^{ref}(f)}{U_{11}^{ref}(f)}, \quad S_{11}^{MOR}(f) = \frac{\tilde{\mathcal{H}}_{11}^{U}(f)U_{exc}(f) - U_{11}^{ref}(f)}{U_{11}^{ref}(f)},$$
$$S_{21}^{TLM}(f) = \frac{U_{21}^{filter}(f)}{U_{11}^{ref}(f)}, \quad S_{21}^{MOR}(f) = \frac{\tilde{\mathcal{H}}_{21}^{U}(f)U_{exc}(f)}{U_{11}^{ref}(f)}, \tag{6.11}$$

where $U_{11}^{filter}(f)$ and $U_{21}^{filter}(f)$ are the FOURIER transforms of the TLM time-domain voltage responses at $z = 2$ and $z = 29$; $U_{11}^{ref}(f)$ is the FOURIER transform of the TLM time-domain voltage response for the reference waveguide at $z = 2$; $\tilde{\mathcal{H}}_{11}^{U}(f)$ and $\tilde{\mathcal{H}}_{21}^{U}(f)$ are reflection and transmission voltage functions ap-

proximated by MOR-algorithms; $U_{exc}(f)$ is the excitation GAUSS pulse signal.

In order to gain the S-parameters 600 ARNOLDI and LANCZOS steps are performed. All MOR-algorithms demonstrate similar convergence. The norms of the right and left residual vectors after 600 MOR-algorithm steps have an order of 10^{-6}. The comparison of the elapsed time is presented in Table 6.10, where the computational time for the S-symmetric LANCZOS algorithm is taken for 100 %.

Table 6.10: Comparison of simulation approaches for the cutoff-coupled dielectric filter (Fig. 6.29).

Method	TLM	ATLM	LTLM	LTLM-S
Time, %	85 × 100	43 × 100	137	100

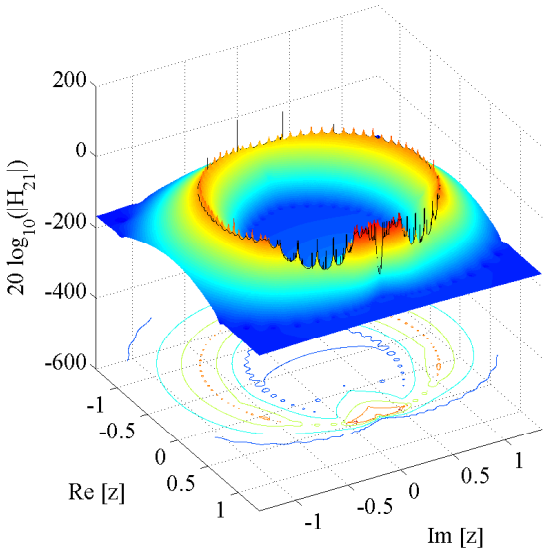
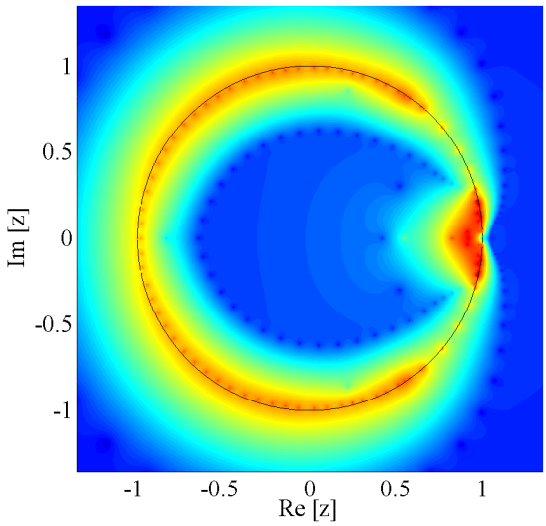


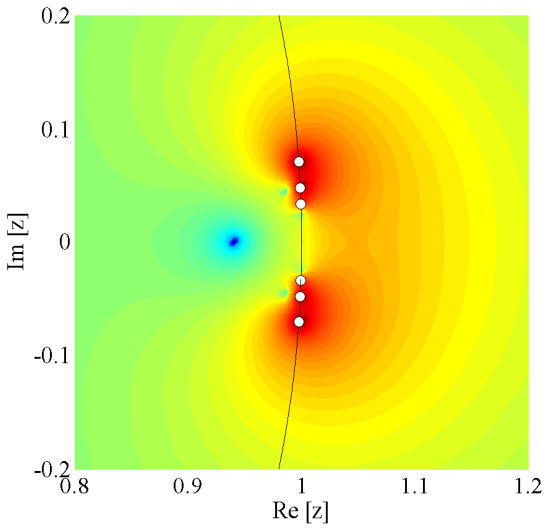
Figure 6.32: Pole-residue distribution of the voltage transfer function from port 1 to port 2 for the cutoff-coupled rectangular dielectric resonator waveguide filter (Fig. 6.29) in the complex z-plane, 3D view.

Fig. 6.32 and Fig. 6.33(a) show the pole-residue distribution of the voltage transfer function $\tilde{\mathcal{H}}_{21}^{U}$ in the complex z-plane after 600 ARNOLDI steps have been performed. After refining only 6 eigenvalues in Fig. 6.33(b) are kept so as to approximate the whole system in the frequency band $3-4$ GHz. The approximating eigenvalues are placed on the complex unit circle in the sector according to the frequency band $0-5$ GHz. Therefore, the order of the two-port macro-model is 6, which comes up to a reduction factor of 2 250.

Hence, we can conclude that both LANCZOS-based methods are very good and suitable for the generation of the macro-model of this cutoff-coupled rectangular dielectric resonator waveguide bandpass filter, whereas LTLM-S is about 37 % faster than the general LANCZOS procedure. The extracted two-port macro-model is characterized by its low order, which corresponds to a significant reduction of the TLM-operator.



(a) Pole-residue distribution after 600 Arnoldi steps



(b) Pole-residue distribution after refining. White circles correspond to the six extracted approximating eigenvalues

Figure 6.33: Pole-residue distribution of the voltage transfer function from port 1 to port 2 for the cutoff-coupled rectangular dielectric resonator waveguide filter (Fig. 6.29) in the complex z-plane, 2D view.

6.7 H-plane Waveguide Filter

The next example is a direct-coupled cavity waveguide filter shown in Fig. 6.34. Inductive irises in the H-plane spaced along the waveguide form four resonators, which are approximately $\lambda_0/2$ in length. The resonators represent a 4th order CHEBYSHEV filter with a 0.5 dB ripple in the passband with a center frequency at about 11 GHz. Design equations and other details for direct-coupled cavity waveguide filters can be found in literature [29].

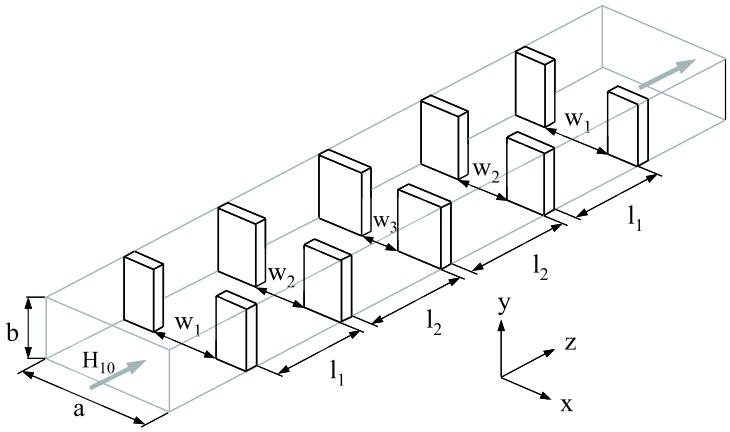


Figure 6.34: Schematic view of a direct-coupled four-resonator H-plane waveguide filter.

All irises are 1 mm thick. The main rectangular waveguide is the WR75 guide (a = 19.050 mm, b = 9.525 mm). The resonator lengths are l_1 =15.712 mm and l_2 =17.671 mm and the gaps are w_1 =10.372 mm, w_2 =6.878 mm, w_3 = 6.19 mm.

The filter is discretized by $53 \times 10 \times 110$ nodes in x, y and z directions, respectively, with the space steps Δx = 0.359 mm, Δy = 0.9525 mm, Δz = 1.043 mm. The waveguide filter is excited on its one side by the E_y-field profile corresponding to the TE_{10} mode with the amplitude E_{y_0}, or to be precise,

$$E_y = |E_{y_0}| \sin\left(\frac{\pi x}{x_{max}}\right). \tag{6.12}$$

The waveguide walls and irises are modeled by PECs. In order to study MOR-techniques the whole structure is simulated, even though due to the symmetry of the filter it is enough to consider only one half.

The S-parameters are calculated using a reference 11.47 cm long WR-75 waveguide. We have been interested in the transfer function of the filter. The reference planes for the reflected and transmitted waves are placed at $z = 5$ and $z = 105$. The transfer function for the E_y-field component (and corresponding voltage) is approximated by MOR-algorithms. The S-parameters of the direct-

coupled cavity waveguide filter can be obtained from simulations as

$$S_{21}^{TLM}(f) = \frac{U_{21}^{filter}(f)}{U_{11}^{ref}(f)}, \quad S_{21}^{MOR}(f) = \frac{\tilde{\mathcal{H}}_{21}^{U}(f)U_{exc}(f)}{U_{11}^{ref}(f)}, \tag{6.13}$$

where $U_{21}^{filter}(f)$ is the FOURIER transform of the TLM time-domain voltage response at $z = 105$; $U_{11}^{ref}(f)$ is the FOURIER transform of the TLM time-domain voltage response for the reference waveguide at $z = 5$; $\tilde{\mathcal{H}}_{21}^{U}(f)$ is the voltage transfer function approximated by MOR-algorithms; $U_{exc}(f)$ is the excitation GAUSS pulse voltage.

In order to obtain the S-parameters of the filter, 50 000 TLM time steps are performed due to a high resonance behavior in the filter. For the S-parameters approximation in the frequency band $10 - 12$ GHz the scattering symmetric LANCZOS algorithm is carried out for 2 080 steps until the convergence for the right and left residual vectors is achieved. The norms of the residual vectors for 10 approximating eigenvalues have an order of 10^{-6}. The extracted S_{21}-parameters presented in Fig. 6.37 are in excellent agreement. The obtained passband ripples are about 0.5 dB.

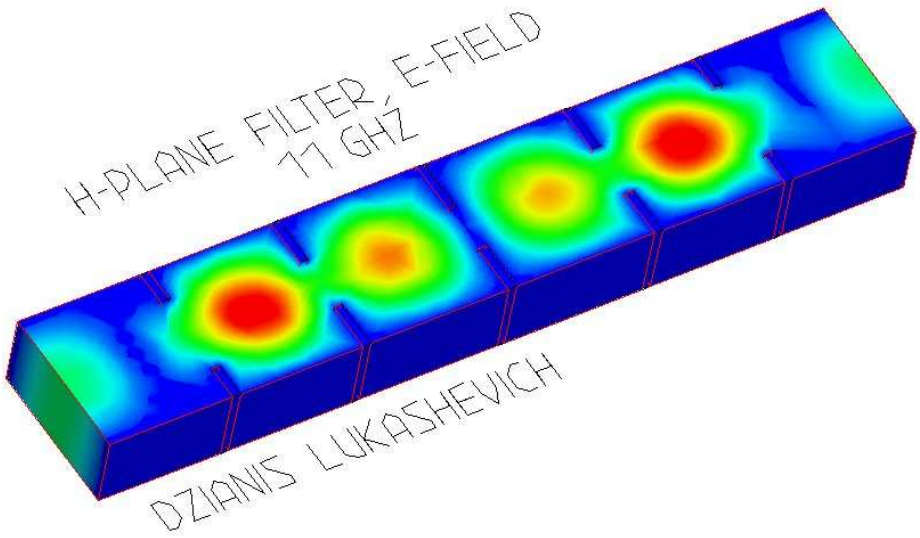


Figure 6.35: Electric field distribution of the direct-coupled four-resonator H-plane waveguide filter (Fig. 6.34) at 11 GHz.

After 2 080 LTLM-S steps have been performed, we have the 2080×2080 reduced tridiagonal matrix containing a couple of eigenvalues approximating the filter transfer function in the frequency band $10 - 12$ GHz. Since EVD is time expensive and has an order of $\mathcal{O}(n^3)$, it makes sense to extract only a few approximating eigenvalues from the pre-reduced tridiaginal matrix, instead to perform pole estimation and selection operations for 2 080 eigenvalues after EVD has been done. For this reason the SIST-via-Lanczos approach has been utilized for the second projection. Because the elapsed time of SIST is similar to $\mathcal{O}(n)$, since only the tridiagonal matrix must be inverted, it allows us to

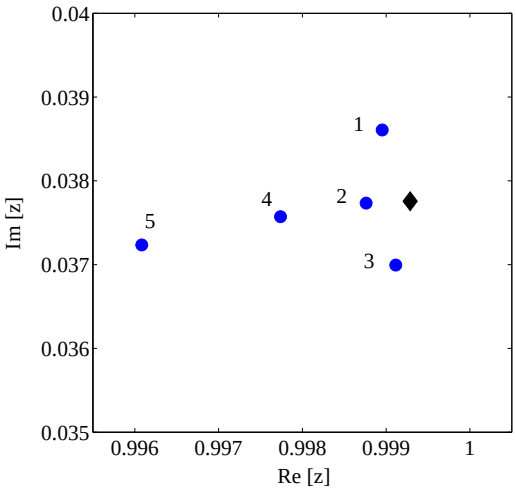


Figure 6.36: Extracted eigenvalues approximating the direct-coupled four-resonator H-plane waveguide filter (Fig. 6.34) in the frequency band $10 - 12$ GHz. Circles denote eigenvalues, whereas the rhombus denotes the expansion point z_{shift}.

conserve computational time additionally. A comparison of simulation methods with reference to the computational time is given in Table 6.11.

Table 6.11: Comparison of simulation approaches for the direct-coupled H-plane filter (Fig. 6.34).

Method	TLM	LTLM-S	LTLM-S + PVL
Time[1], %	6.7 × 100	**100**	83
Time[1], %	8 × 100	120	**100**

[1] The time needed for the generation of a complete reduced-order model including the post-processing.

In this case the treatment of the eigenvalue problem by means of the SIST-approach was approximately 13 times faster than the usage of the standard MATLAB EVD routine, that allows to save about 20% of total computational time. Only 9 LANCZOS steps in the SIST-approach were necessary to extract 5 eigenvalues (in the upper half of the unit circle), which were closed to an expansion point. The extracted 5 approximating eigenvalues closed to the expansion point are shown in Fig. 6.36. The expansion point z_{shift} is chosen on the unit circle in the complex plane in accord with the center resonance frequency $f_0 = 11$ GHz

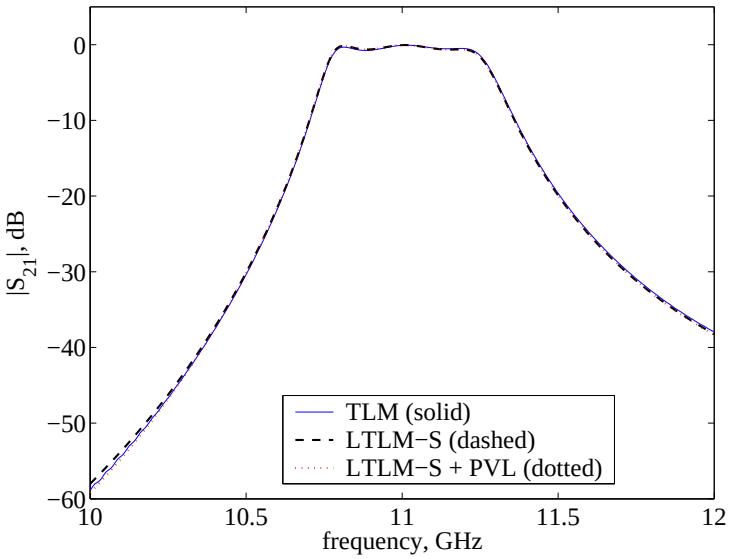
$$z_{shift} = \exp(\jmath 2\pi f_0 \Delta t) = 0.9993 + \jmath 0.0378\,.$$

The three eigenvalues (1, 2 and 3) closest to z_{shift} determine the ripple of the filter (see Fig. 6.37 (b)), whereas the other two (4 and 5) have influence on the approximation accuracy.

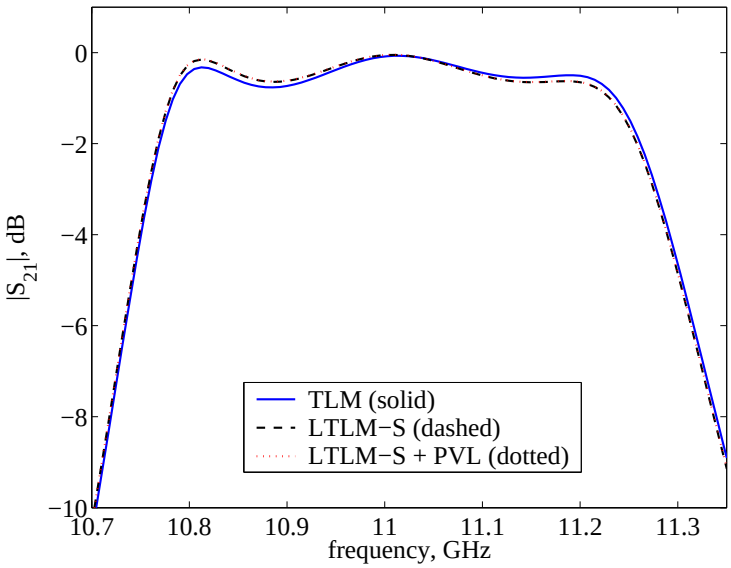
Thus, the order of the generated ROM is 10 and the corresponding reduction factor of the TLM-system is 104 940.

In order to validate our results we re-simulate the filter by means of ANSOFT HFSS using the Fast Frequency Sweep (FFS) based on the LANCZOS process applied to the Finite Element Method [10]. The central frequency for FFS has been chosen as well as 11 GHz. The total mesh contains 13 012 tetrahedra versus 58 300 18^{th}-port cells in TLM. The obtained S-parameters are presented in Fig. 6.38. Even though the filter has four resonators the zoomed regions of computed S-parameters (see Fig. 6.37(b) and Fig. 6.38(b)) exhibit three ripples instead expected four ones [29]. Fig. 6.35 shows the distribution of the electric field in all four resonators of the filter. Obviously, two central poles describing the second and the third resonator are merged in the pole 2 in Fig. 6.36 or two right poles are merged in the pole 3 in Fig. 6.36. This is also possible that the lost pole will appear with increasing number of time steps or algorithm iterations.

In conclusion, we can notice that the scattering symmetric LANCZOS algorithm is assumed to be robust and very suitable for the high order discretized high-Q resonance structures. The results obtained through the implemented MOR-techniques are in very good agreement with simulations performed by the time-domain TLM scheme and frequency-domain FEM, which also utilizes the LANCZOS algorithm for FFS. As an alternative to the conventional eigenvalue decomposition procedure for the treatment of pre-reduced tridiagonal matrices, generated by MOR-techniques, the Padé approximations and/or shift-invert spectral transformations can be involved to extract a few poles, which allows us to save a significant amount of post-processing computational time, especially for loss-free structures with high-order discretizations.

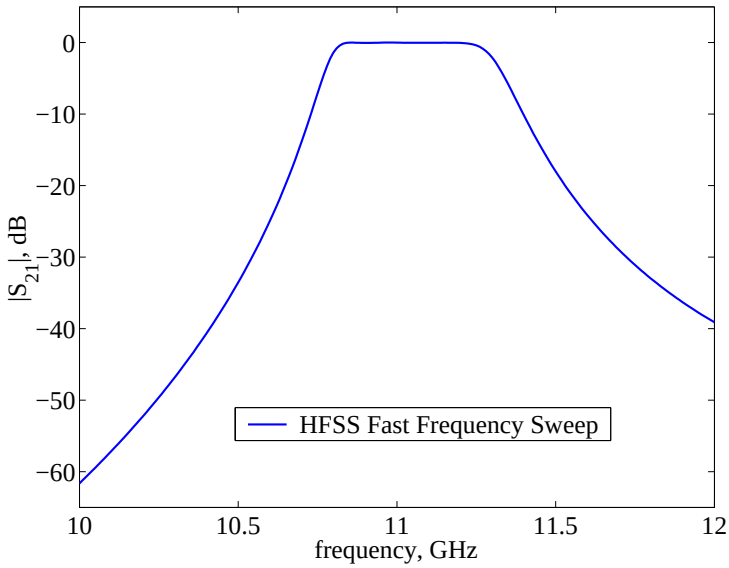


(a) Magnitudes of S_{21}-parameters

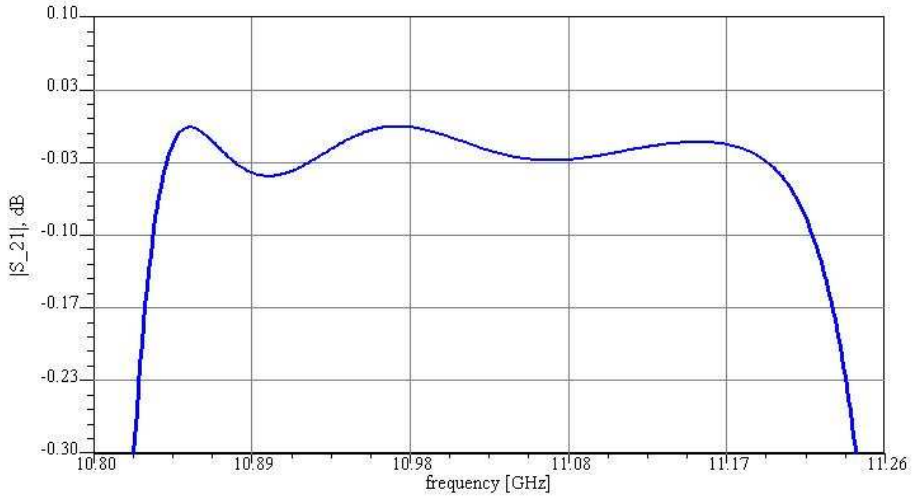


(b) Zoomed magnitudes of S_{21}-parameters

Figure 6.37: S-parameters for the direct-coupled four-resonator H-plane waveguide filter (Fig. 6.34) computed by the time-domain TLM and MOR-techniques.



(a) Magnitude of S_{21}-parameter



(b) Zoomed magnitude of S_{21}-parameter

Figure 6.38: S-parameters for the direct-coupled four-resonator H-plane waveguide filter (Fig. 6.34) computed by means of HFSS Fast Frequency Sweep.

Chapter 7

Conclusions

The discrete electromagnetic model resulting from the application of the Transmission Line Matrix (TLM) method is cast in a state-space matrix form involving wave amplitudes as state variables. For structures of practical interest the dimension of this matrix can be very large, up to several millions. The use of Model Order Reduction (MOR) techniques in the context of TLM-based electromagnetic modeling yields considerable reduction in the computational effort necessary for the solution of the discrete models. Furthermore, they facilitate the direct construction of compact discrete models of the electromagnetic structure under investigation. More specifically,

- MOR yields an acceleration factor of 10-30 in the computation of the response of loss-free, high-Q filters and resonators, while 30-50 % reduction in simulation time have been observed in obtaining the electroamgnetic response of microstrip-line and coplanar waveguiding structures and a patch antenna;

- in terms of model order reduction, for loss-free structures the order reduction by a factor of several thousands is possible, compared to the order of the original TLM model; for structures exhibiting loss the order reduction by a factor of several hundreds can be achieved.

The emphasis of our application of MOR-techniques to TLM has been on the use of KRYLOV subspace methods based on the LANCZOS and ARNOLDI processes. The properties of these methods, both in terms of computational efficiency, convergence and solution accuracy, have been examined through their application to the analysis of several electromagnetic structures containing up to 1 728 000 unknown space-state variables. One-port and two-port structures have been investigated and their appropriate macro-models have been generated. The generated reduced-order models exhibit very good accuracy, as validated through comparisons with results obtained by other means. Optimal model order selection procedures have also been proposed. More specifically, the order of the extracted reduced macro-models can be decreased further by introducing a Signal-to-Noise-Ratio (SNR) threshold.

The implicit time evolution TLM-scheme has been utilized for the purposes of MOR. For this purpose, some modifications of the TLM scattering algorithm introduced originally by JOHNS and field mappings needed for MOR implementations have been performed. The influence of various types of boundary

conditions encountered in practice on the MOR convergence has been discussed. Also, a brief comparison has been provided between FDTD, FIT and TLM with regards to the computational efficiency of MOR-methods for expediting their numerical integration.

Reduced-order models of the TLM system in both the z-domain and the LAPLACE domain have been derived. An important attribute of the model order reduction processes described in this thesis is the use of the congruence transformations for the construction of the state-space matrices of the reduced-order system. Under the assumption of stability and passivity of the original discrete system, the use of such congruence transformations guarantees the stability and passivity of the reduced-order system. Conservation of stability and passivity is an essential requirement for the reduced-order system, since the stability and passivity of the generated macro-model alone is not sufficient to guarantee the stability and stability of the overall system that results from the integration of the macro-model with other multi-port circuits. It has been shown that the obtained TLM approximations and associated reduced-order macro-models are always stable and passive.

It has been shown that the non-symmetric properties of the TLM-operator require the use of the general ARNOLDI and LANCZOS algorithms. Even though the implementation of the LANCZOS algorithm is more efficient and algorithmically simpler than ARNOLDI, one LANCZOS step is equivalent to two time updates in the standard TLM scheme, due to the imperative to deal with both the TLM matrix and its transpose for the construction of two biorthogonal KRYLOV subspaces. One of our contributions has been a method that takes advantage of specific symmetric properties of the TLM-operator to develop a modified LANCZOS algorithm that does not require use of the transpose of the TLM matrix. This new algorithm is called the scattering-symmetric or S-symmetric LANCZOS process, and it is shown to be up to 35 % faster than the conventional non-symmetric LANCZOS algorithm. Furthermore, the use of the S-symmetry results in reduction in computer memory resources compared to the general non-symmetric approach.

As already mentioned above, the LANCZOS and ARNOLDI procedures provide numerically stable and computationally efficient approaches to the construction of the TLM projection matrices. From comparative studies of KRYLOV based processes the following conclusions are reached:

- the nonsymmetric, and the S-symmetric LANCZOS algorithms are assumed to be sparing concerning the computational resources; huge computation time and memory requirement of ARNOLDI-based algorithms is a big drawback and makes them unattractive if not prohibitive;

- the LANCZOS methods exhibit slower convergence than ARNOLDI algorithm due to loss of the biorthogonality property in the numerical construction of the bases of the KRYLOV subspaces;

- the utilization of very expensive re-biorthogonalization in the LANCZOS processes does not necessarily improve biorthogonality, and, in some cases it may even destroy the biorthogonality of the generated basis vectors;

- the ARNOLDI algorithm generates an orthonormal basis nearly without loss of orthogonality;

- eigenvalues of the reduced LANCZOS trigiagonal matrix are more sensitive to its changes than eigenvalues of the reduced ARNOLDI upper HESSENBERG matrix to appropriate changes, that can additionally make a significant impact on the accuracy of eigenvalues computation by Eigenvalue Decomposition (EVD);

- the S-symmetric version of the LANCZOS algorithm is assumed to be spare concerning the computational resources, but gives slightly bigger error due to the biorthogonality loss;

- the biconjugate based S-symmetric LANCZOS method provides alternative and/or additional stopping criteria, however its resource requirements make it less attractive;

- the iterative Shift-Invert Spectral Transformation (SIST) approach (applied for order reduction of the original implicit state-space system in the first step) exhibits large approximation error, instability and high computational time.

As an alternative to the conventional EVD procedures for treatment of reduced tridiagonal and upper HESSENBERG matrices generated by MOR-techniques the PADE-via-LANCZOS, PADE-via-ARNOLDI approximation or SIST technique can be involved to extract a few poles approximating the original state-space system. This two-step reduction allows to reduce post-processing computational time significantly, especially, if pre-reduced matrices has a large order.

In addition, the novel and effective hybrid Transmission Line Matrix-Mode Matching (TLM-MM) approach is proposed for a full-wave analysis of passive transmission structures in MMICs. This method is a combination of the frequency-domain Mode-Matching (MM) method for the modeling of homogeneous waveguide structures with the time-domain TLM method for simulation of discontinuities with arbitrary shapes. It has been shown that using the hybrid TLM-MM approach the time and memory requirements can been reduced several times in comparison to pure TLM simulations, while moreover improving accuracy. A comparison between both methods on the basis of numerical simulations is provided. The proposed approach can be easily involved in the macro-modeling process of MMICs, and the generation of a reduced-order macro-model can be simplified and additionally accelerated. The described time/frequency hybrid technique based on MM method can be also applied to FDTD and FIT methods in a straightforward fashion.

With regards to future research, the present work can be extended to address the following issues:

- implementation of the developed MOR-techniques for sub-gridding in TLM, with specific emphasis placed on the extraction of macro-cells and nested macro-cells for expedient and computationally efficient representation of very fine features in an otherwise large-feature-size structure;

- construction of network oriented models via FOSTER canonical forms;

- application to domain-decomposition analysis of large-scale structures, through a) the partitioning of the computational domain into sub-domains,

for each of which a reduced-order model is first generated; b) the subsequent generation of a reduced-order model for the entire structure, through the interconnection of the generated sub-domain macro-models.

Appendix A

Vector Spaces

A.1 Basic Definitions

Definition A.1.1. A set T a with a mapping $d : T \times T \to \mathbb{R}$ and properties $(\forall\ s,t,u \in T)$

(a) $d(s,t) \geq 0$,

(b) $d(s,t) = d(t,s)$,

(c) $d(s,u) \leq d(s,t) + d(t,u)$ (*triangle inequality*),

(d) $d(s,t) = 0 \ \Leftrightarrow s = t$,

is called the *metric space* with the *metric* d.
A sequence $(t_n) \subset T$ converges to $t \in T$ if

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ n > N\ d(t_n, t) \leq \varepsilon$$

A sequence $(t_n) \subset T$ is called *Cauchy sequence* if

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ n,m > N\ d(t_n, t_m) \leq \varepsilon$$

Definition A.1.2. A metric space T is said to be the *complete* if each Cauchy sequence $t_n \subset T$ is convergent.

Definition A.1.3. Let be $\boldsymbol{X}$ a vector space of $\mathbb{K}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. A map $p : \boldsymbol{X} \to \mathbb{R}$ is called the *norm* if

(a) $p(\lambda \boldsymbol{x}) = |\lambda| p(\boldsymbol{x})\ \forall\ \lambda \in \mathbb{K}, \boldsymbol{x} \in \boldsymbol{X}$,

(b) $p(\boldsymbol{x} + \boldsymbol{y}) \leq p(\boldsymbol{x}) + p(\boldsymbol{y})\ \forall\ \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$,

(c) $p(\boldsymbol{x}) = 0 \Rightarrow p(\boldsymbol{x}) = 0$.

The *normed space* is defined by a pair $(\boldsymbol{X}, p)$. On a normed space $(\boldsymbol{X}, \| \ . \ \|)$ a metric (A.1.1) $d(\boldsymbol{x}, \boldsymbol{y}) = \| \ \boldsymbol{x} - \boldsymbol{y} \ \|$ is defined, then $\forall\ \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{X}$

(a) $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$,

(b) $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x})$,

(c) $d(\boldsymbol{x},\boldsymbol{z}) \leq d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$,

(d) $d(\boldsymbol{x},\boldsymbol{y}) = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{z}$.

The sequence of vectors $(\boldsymbol{x}_n)$ is the Cauchy sequence in a normed space $(\boldsymbol{X}, \| \, . \, \|)$ if

$$\forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N}\ \forall\, n, m > N \ \ \| \boldsymbol{x}_n - \boldsymbol{x}_m \| \leq \varepsilon$$

Definition A.1.4. A complete normed space is called the *Banach space.*

Definition A.1.5. A Banach space $\boldsymbol{X}$ with a Hermitian inner product $\langle ., . \rangle : \boldsymbol{X} \times \boldsymbol{X} \to \mathbb{K}$ that satisfies the following properties

(a) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if only if $|\boldsymbol{x}\rangle = 0$,

(b) $\langle \alpha \boldsymbol{x}, \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle$,

(c) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle^*$,

(d) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$,

$\forall\ |\boldsymbol{x}\rangle, |\boldsymbol{y}\rangle, |\boldsymbol{z}\rangle \in \boldsymbol{X}$ and $\forall\ \alpha \in \mathbb{K}$ is called the *Hilbert space* $\mathcal{H}$.

The vectors $\langle \boldsymbol{x}|, \langle \boldsymbol{y}|$ and $\langle \boldsymbol{z}|$ are vectors of a dual to $\boldsymbol{X}$ space to be the set of all linear functionals $\boldsymbol{X}^H = \{f : \ \boldsymbol{X} \to \mathbb{K}\}$. From the definition A.1.5 follows, that $\langle \boldsymbol{x}|^H = |\boldsymbol{x}\rangle$ or $\langle \boldsymbol{x}| = |\boldsymbol{x}\rangle^H$.
The *norm* in HILBERT space is defined by $\| \boldsymbol{x} \|_2 = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$.

Appendix B

The Gram-Schmidt algorithm

B.1 Basic Gram-Schmidt Process

Let

$$\boldsymbol{Z} = \{|\boldsymbol{z}_1\rangle, |\boldsymbol{z}_2\rangle, |\boldsymbol{z}_3\rangle, \ldots, |\boldsymbol{z}_n\rangle\} \tag{B.1}$$

be a set of linearly independent vectors in $\mathcal{C}^N$. The GRAM-SCHMIDT orthogonalization procedure generates a corresponding orthogonal set of vectors

$$\boldsymbol{Q} = \{|\boldsymbol{q}_1\rangle, |\boldsymbol{q}_2\rangle, |\boldsymbol{q}_3\rangle, \ldots, |\boldsymbol{q}_n\rangle\} \in \mathcal{C}^N \tag{B.2}$$

such that

$$\text{span } \{\boldsymbol{Q}\} = \text{span } \{\boldsymbol{Z}\}\,, \tag{B.3}$$

and

$$\boldsymbol{Z} = \boldsymbol{Q}\boldsymbol{R}\,, \tag{B.4}$$

where $\boldsymbol{R} = r_{(ij)}$ is a unit upper triangular matrix. The vectors $|\boldsymbol{q}_k\rangle$ and the matrix $\boldsymbol{R}$ are defined recursively by the following equations. Define $|\boldsymbol{q}_1\rangle = |\boldsymbol{z}_1\rangle$. Then for each $k = 2, \ldots, n$ define

$$|\boldsymbol{q}_k\rangle = |\boldsymbol{z}_k\rangle - \sum_{i=1}^{k-1} r_{ki}|\boldsymbol{q}_i\rangle\,, \tag{B.5}$$

where

$$\boldsymbol{r}_{ki} = \langle \boldsymbol{z}_k|\boldsymbol{q}_i\rangle/d_i\,, \quad d_i = \langle \boldsymbol{q}_i|\boldsymbol{q}_i\rangle\,. \tag{B.6}$$

Observe that

$$\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{D} = \text{diag } \{d_1, d_2, \ldots, d_n\}\,. \tag{B.7}$$

The error propagation which occurs when eq. (B.10) are implemented on a computer is such that they are seldom used in practice. Typically the modified form of the GRAM-SCHMIDT orthogonalization is used.

B.2 Modified Gram-Schmidt Orthogonalization

The following modification of the basic GRAM-SCHMIDT procedure yields an orthogonalization procedure with better numerical stability than the GRAM-SCHMIDT procedure as defined in (B.10). The modified process utilizes following

recursions. The vectors $|\boldsymbol{q}_k\rangle$ and the matrix $\boldsymbol{R}$ are defined recursively by the following equations.

$$\begin{aligned} &\text{for each } k \text{ define } |\boldsymbol{q}_k\rangle = |\boldsymbol{z}_k\rangle^{(k)} \text{ where} \\ &|\boldsymbol{z}_j\rangle^{(k+1)} = |\boldsymbol{z}_j\rangle^{(k)} - r_{kj}|\boldsymbol{q}_k\rangle, \text{ for } j > k \\ &r_{kj} = \langle \boldsymbol{q}_k|\boldsymbol{z}_j\rangle^{(k)}/d_k \quad \text{and} \quad d_k = \langle \boldsymbol{q}_k|\boldsymbol{q}_k\rangle\,. \end{aligned} \tag{B.8}$$

Column interchanges (pivoting) can be used in the orthogonalization process. In particular at any stage in the process the vector in the remaining subset of vectors which has maximal norm can be normalized. That is at stage k, choose vector $|\boldsymbol{z}_s^{(k)}\rangle$ such that

$$\|\boldsymbol{z}_s^{(k)}\|_2 = \max_{k\le j\le n} \|\boldsymbol{z}_j^{(k)}\|_2\,. \tag{B.9}$$

B.3 Two-Sided Modified Gram-Schmidt Biorthogonalization

The two-sided modified GRAM-SCHMIDT may be used in order to rebiorthogonalize a deteriorated biorhogonal basis

$$\boldsymbol{P}^T\boldsymbol{Q} = \boldsymbol{D} = \text{diag}\,\{d_1, d_2, \ldots, d_n\} \tag{B.10}$$

by following recursions

$$\begin{aligned} &\text{for each } i \ = 1, 2, \ldots, j \text{ compute} \\ &|\boldsymbol{q}_{j+1}\rangle = |\boldsymbol{q}_{j+1}\rangle - \langle \boldsymbol{p}_i|\boldsymbol{q}_{j+1}\rangle|\boldsymbol{q}_i\rangle\,, \\ &|\boldsymbol{p}_{j+1}\rangle = |\boldsymbol{p}_{j+1}\rangle - \langle \boldsymbol{q}_i|\boldsymbol{p}_{j+1}\rangle|\boldsymbol{p}_i\rangle\,. \end{aligned} \tag{B.11}$$

Here

$$\boldsymbol{P} = \{|\boldsymbol{p}_1\rangle, |\boldsymbol{p}_2\rangle, |\boldsymbol{p}_3\rangle, \ldots, |\boldsymbol{p}_n\rangle\} \in \mathcal{H}^N \tag{B.12}$$

is a corresponding biorthogonal set of vectors. Eq. (B.11) rebiorthogonalizes two vectors $|\boldsymbol{q}_{j+1}\rangle$ and $|\boldsymbol{p}_{j+1}\rangle$ at each step j.

List of Figures

List of Tables

List of Acronyms and Symbols

Acronyms

ABC	Absorbing Boundary Conditions
AWE	Asymptotic Waveform Evaluation
ATLM	Arnoldi-based TLM-MOR
BCG	BiConjugate Gradient
CFM	Center Field Mapping
CBM	Cell Boundary Mapping
CPW	Coplanar WaveGuide
EM	Electromagnetic
EVD	Eigenvalues Decomposition
FDTD	Finite Difference Time–Domain
FEM	Finite Element Emthod
FIT	Finite Integration Technique
GHM	Generalized Hessenberg method
GIM	Generalized Impedance Matrix
LTLM	nonsymmetric Lanczos-based TLM-MOR
LTLM-S	S-symmetric Lanczos-based TLM-MOR
BCG-LTLM-S	BiConjugate Gradient and S-symmetric Lanczos based TLM-MOR
MIMO	Multi-Input Multi-Output
MM	Mode Matching
MMIC	Microwave Monolithic Integrated Circuits
MoM	Method of Moments
MOR	Model Order Reduction
ML	Matched Layers
PEC	Perfect Electric Conductor
PMC	Perfect Magnetic Conductor
PML	Perfectly Matched Layers
PMSI	Prony model based Signal Identification
PVA	Pade-Via-Lanczos
PVL	Pade-Via-Lanczos
QMR	Quasi-Minimal Residual
ROM	Reduced-Order Model
SISO	Single-Input Single-Output
SI-LTLM-S	Shift-Inverse TLM-MOR based on the S-symmetric Lanczos process
SCN	Symmetrical Condensed Node
SI	Shift-Inverse, (or System Identification)
SIST	Shift-Invert Spectral Transformation
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
TLM	Transmission Line Matrix
TLM-ME	Transmission Line Matrix–Multipole Expansion
TLM-MM	Transmission Line Matrix–Mode Matching

Symbols

$\boldsymbol{\Gamma}$	connection operator
$\boldsymbol{S}$	scattering operator
$\boldsymbol{\Gamma S}$	general TLM operator
$\boldsymbol{S\Gamma}$	transpose general TLM operator
$(\boldsymbol{\Gamma S})_n$	reduced TLM operator
λ	eigenvalues
$\boldsymbol{\Lambda}$	eigenvalue matrix
$\mathcal{R}, \mathcal{C}$	sets of real and complex numbers
$\mathcal{R}^N, \mathcal{C}^N, \mathcal{H}^N$	real, complex and HILBERT spaces
$\mathcal{K}$	KRYLOV subspaces
$\text{diag}(\alpha, \beta, \cdots)$	diagonal matrix constructed from scalars $\alpha, \beta, \ldots$
$\text{span}(\boldsymbol{a}, \boldsymbol{b}, \ldots)$	spanning space of vectors $\boldsymbol{a}, \boldsymbol{b}, \ldots$
$\mathcal{O}(.)$	big O notation
$\boldsymbol{A}^{N\times n}$	indicate that $\boldsymbol{A}$ is N-by-n
$\boldsymbol{A}^*$	conjugate of $\boldsymbol{A}$
$\boldsymbol{A}^T$	transpose of $\boldsymbol{A}$
$\boldsymbol{A}^H$	HERMITIAN of $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	inverse of $\boldsymbol{A}$
$\boldsymbol{A}^\dagger$	HERMITIAN conjugate of $\boldsymbol{A}$
$\det \boldsymbol{A}$	determinant of $\boldsymbol{A}$
$\langle . \vert$	bra-vector in DIRAC notation
$\vert . \rangle$	ket-vector in DIRAC notation
$\langle . \vert . \rangle$	inner product in DIRAC notation
$\vert . \rangle\langle . \vert$	projection operator in DIRAC notation
$\Vert . \Vert_p$	vector and matrix p norm
$\otimes$	tensor product
$\oplus$	direct summation

Printing in boldface is referred to vectors, operators and matrices.

Symbols $\hat{.}$ and $\tilde{.}$ are often used to denote temporal vectors, matrices and scalars.

Bibliography

[1] M. Aidam. *Wavelet-Galerkin-Methoden zur Berechnung electromagnetischer Felder im Zeitbereich.* PhD thesis, Technische Universität München, 1999.

[2] M. Aidam and P. Russer. Derivation of the transmission line matrix method by finite integration. *AEÜ Int. J. Electron. Commun*, 51(1):35–39, January 1997.

[3] B. D. O. Anderson and S. Vongpanitlerd. *Network Analysis and Synthesis.* Prentice Hall, New Jersey, 1973.

[4] A. C. Antoulas and D. C. Sorensen. Projection methods for balanced model reduction. *Technical Report ECE-CAAM Depts., Rice University*, September 2000. Rice University, Houston, TX.

[5] F. Arndt, V. Catina, and J. Brandt. Flexible hybrid MM/MoM technique for the CAD and optimization of arbitrarily shaped 3-D waveguide components. *Proc. 33th European Microwave Conference, Munich*, pages 343–346, October 2003.

[6] B. Bader and P. Russer. The alternating transmission line matrix ATLM scheme. *1995 Int. Microwave Symposium Digest, Orlando*, pages 19–22, 1995.

[7] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide.* SIAM, Philadelphia, 2000.

[8] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition.* SIAM, Philadelphia, PA, 1994.

[9] J. Bracken, V. Raghavan, and R. Rohrer. Interconnect simulation with AWE. *IEEE Trans. Circuits and Systems*, 39(11):869–878, November 1992.

[10] J. Bracken, D. Sun, and Z. Cendes. S-domain methods for simultaneous time and frequency characterization of electromagnetic devices. *IEEE Trans. Microwave Theory Techn.*, 46(9):1277–1290, September 1998.

[11] J. E. Bracken and Z. J. Cendes. Asymptotic waveform evaluation for s-domain solution of electromagnetic devices. *IEEE Trans. Magnetics*, 34(5):3232 – 3235, September 1998.

[12] J. E. Bracken, D. K. Sun, and Z. J. Cendes. S-domain methods for simultaneous time and frequency characterization of electromagnetic devices. *IEEE Trans. Microwave Theory Techn.*, 46:1277–1290, September 1998.

[13] J. Brauer and G. C. Lizalek. Microwave filter analysis using a new 3-D finite-element modal frequency method. *IEEE Trans. Microwave Theory Techn.*, 45(5):810–818, May 1997.

[14] A. Cangellaris, M. Celik, S. Pasha, and L. Zhao. Electromagnetic model order reduction for system level modeling. *IEEE Trans. Microwave Theory Techn.*, 47:840–850, June 1999.

[15] A. Cangellaris, S. Pasha, S. Prince, and M. Celik. A new discrete transmission line model for passive model order reduction and macromodeling of high-speed interconnections. *IEEE Trans. Advanced Packaging*, 22(2):356 – 364, August 1999.

[16] A. Cangellaris and Z. Zhao. Rapid FDTD simulation without time stepping. *IEEE Microwave And Guided-Wave Letters*, 9(1):4–6, January 1999.

[17] Z. J. Cendes. A comparison of fast frequency sweep and transient model order reduction algorithms. *2004 Int. Microwave Symposium Workshop Digest*, June 2004.

[18] J. Chen and S. M. Kang. An algorithm for automatic model-order reduction of nonlinear MEMS devices. *IEEE International Symposium on Circuits and Systems*, 2:445 – 448, May 2000.

[19] J. Chen, S. M. Kang, J. Zou, C. Liu, and J. E. Schutt-Aine. Reduced-order modeling of weakly nonlinear MEMS devices with Taylor-series expansion and Arnoldi approach. *Journal of Microelectromechanical Systems*, 13(3):441 – 451, June 2004.

[20] Z. Chen, M. Ney, and W. Hoefer. A new finite-difference time-domain formulation and its equivalence with the TLM symmetrical condensed node. *IEEE Trans. Microwave Theory Techn.*, 39(12):2160–2169, December 1991.

[21] Z. Chen, M. Ney, and W. Hoefer. Absorbing and connecting boundary conditions for the TLM method. *IEEE Trans. Microwave Theory Techn.*, 41(11):2016–2024, November 1993.

[22] E. Chiprout and M. S. Nakhla. *Asymptotic Waveform Evaluation and Moment Matching for Interconnect Analysis.* Kluwer Academic Publishers, Norwell, MA, 1994.

[23] C. Christopoulos. *The Transmission-Line Modeling Method: TLM.* IEEE Press, Piscataway, NJ, 1995.

[24] V. Chtchekatourov, F. Coccetti, and P. Russer. Direct Y-parameters estimation of microwave structures using TLM simulation and Prony's method. *Proc. 17th Annual Review of Progress in Applied Computational Electromagnetics, Monterey*, pages 580–586, May 2001.

[25] V. Chtchekatourov, W. Fisch, F. Coccetti, and P. Russer. Full-wave analysis and model-based parameter estimation approaches for S- and Y- matrix computation of microwave distributed circuits. *2001 Int. Microwave Symposium Digest, Phoenix*, pages 1037–1040, June 2001.

[26] L. J. Chu. Physical limitations of omni–directional antennas. *J. Applied Physics*, pages 1163–1175, December 1948.

[27] F. Coccetti. *Application of System Identification (SI) to Full-Wave Time Domain Characterization of Microwave and MillimeterWave Passive Structures.* PhD thesis, Technische Universität München, 2004.

[28] R. E. Collin. *Field Theory of Guided Waves.* IEEE Computer Society Press, second edition, 1990.

[29] R. E. Collin. *Foundations for Microwave Engineering.* Mc-Graw Hill, second edition, 1992.

[30] V. Crino, M. Mongiardo, and C. Tomassoni. Efficient CAD of class of waveguide discontinuities via hybrid Finite-Element/Mode-Matching and modified Transverse-Resonance Analysis. *2003 Int. Microwave Symposium Digest, Philadelphia*, pages 1223–1226, 2003.

[31] J. K. Cullum. Arnoldi versus nonsymmetric Lanczos algorithms for solving nonsymmetric matrix eigenvalue problems. *Technical report TR-93-123, TR-3576*, 1, 1996. University of Maryland, USA.

[32] J. K. Cullum and R. A. Willoughby. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol.1 Theory.* Progress in Scientific Computing. Birkhäser, Boston, Basel, Stuttgart, 1985.

[33] D. Day. An efficient implementation of the nonsymmetric Lanczos algorithm. *SIAM J. Matrix Anal. Appl.*, 18(3):566–589, July 1997.

[34] B. P. Demidovich and I. A. Maron. *Basics of numerical mathematics.* Science, Moscow, 1970.

[35] B. Denecker, F. Olyslager, L. Knockaert, and D. De Zutter. Generation of FDTD subcell equations by means of reduced-order modelling. *IEEE Trans. Antennas and Propagation*, 51(8):1806–1817, August 2003.

[36] A. Dounavis, E. Gad, R. Achar, and M. Nakhla. Passive model-reduction of distributed networks with frequency-dependent parameters. *2000 Int. Microwave Symposium Digest, Boston*, 3:1789–1792, June 2000.

[37] W. Dressel. *Subgridding in the TLM method.* PhD thesis, Technische Universität München, 2005.

[38] J. Dubard and D. Pompei. Simulation of Berenger's perfectly matched layer with a modified TLM node. *IEEE Trans. Antennas and Propagation*, 144(3):205–207, March 1997.

[39] T. Ericsson and A. Ruhe. The spectral transformation Lanczos method for the numerical solution of large sparse generalized symmetric eigenvalue problems. *Math. Comp.*, 35:1251–1268, 1980.

[40] C. Eswarappa and W. Hoefer. Implementation of Berenger absorbing boundary conditions in TLM by interfacing FDTD perfectly matched layers. *Electron. Lett.*, 31(15):1264–1266, 1995.

[41] P. Feldman and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos algortihm. *IEEE Trans. Computer-Aided Design*, 14:639–649, 1995.

[42] L. B. Felsen and N. Marcuvitz. *Radiation and Scattering of Waves.* IEEE Press, Piscataway, NJ, second edition, 1994.

[43] R. W. Freund. Krylov-subspace methods for reduced-order modelling in circuit simulation. *Numerical Analysis Manuscript No. 99-3-17*, November 1999. Bell Labs, Murray Hill, NJ, USA.

[44] R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal. An implementation of the look-ahead Lanczos algorithm for non-hermitian matrices. *Technical report No. 99-09*, April 1991. RIACS, NASA Ames Research Center, MOffet Field, CA.

[45] R. W. Freund and N. M. Nachtigal. Software for simplified Lanczos and QMR algorithms. *Appl. Numer. Math.*, 19:319–341, 1995. AT&Bell Laboratories, Murray Hill, NJ.

[46] R. W. Freund and T. Szeto. A transpose-free quasi-minimal residual squared algorithm for non-Hermitian indefinite linear systems. *Advances in Computer Methods for Partial Differential Equations - VII, IMACS*, pages 258–264, 1992. New Brunswick, CA.

[47] R. W. Freund and H. Zha. Simplifications of the nonsymmetric Lanczos process and a new algorithm for Hermitian indefinite linear systems. *Numerical Analysis Manuscript*, 1995. AT&Bell Laboratories, Murray Hill, NJ.

[48] G. H. Golub and C. Van Loan. *Matrix computation.* The John Hopkins University Press, Baltimore, third edition, 1996.

[49] R. F. Harrington. *Time Harmonic Electromagnetic Fields.* McGraw-Hill, New York, first edition, 1961.

[50] S. Hein. Consistent finite-difference modeling of Maxwell's equations with lossy symmetrical condensed TLM node. *Int. J. Num. Mod.*, 6:207–220, 1993.

[51] S. Hein. Finite-difference time-domain approximation of Maxwell's equations with nonorthogonal condensed TLM mesh. *Int. J. Num. Mod.*, 7:179–188, 1994.

[52] S. Hein. TLM numerical solution of bloch's equations for magnetized gyrotropic media. *Appl. Math. Mod.*, 21:221–229, 1997.

[53] H. Heldener, H. Körner, A. Mitchell, M. Schwerd, and U. Seidel. Comparison of copper damascene and aluminum RIE metallization in BiCMOS technology. *Proc. of the Materials for Advanced Metalization Conference*, March 2000. Stresa, Italy.

[54] J. Herring. *Developments on the Transmission-Line Modelling Method for Electromagnetic Compatibility Studies.* PhD thesis, University of Nottingham, 1993.

[55] J. Herring and C. Christopoulos. Solving electromagnetic problems using multiple grid transmission line modeling method. *IEEE Trans. Antennas and Propagation*, 42(12):1654–1658, December 1994.

[56] W. J. R. Hoefer. The transmission line matrix (TLM) method. In T. Itoh, editor, *Numerical Techniques for Microwave and Millimeter Wave Passive Structures*, pages 496–591. John Wiley & Sons, New York, 1989.

[57] Y. Hua and T.K. Sarkar. Further analysis of three modern techniques for pole retrieval from data sequence. *in Proc. 30^{th} Midwest Symp. Circuits Syst., Syracuse, NY*, August 1987.

[58] B. Isele and P. Russer. Improved skin effect model in TLM. *COMPUMAG*, 1993.

[59] H. Jin and R. Vahldieck. The frequency-domain TLM method. *IEEE Trans. Microwave Theory Techn.*, 42(12):2554–2561, December 1994.

[60] P. B. Johns. A symmetrical condensed node for the TLM method. *IEEE Trans. Microwave Theory Techn.*, 35(4):370–377, April 1987.

[61] G. A. Baker Jr. and P. Graves-Morris. *Padé Approximants.* Cambridge University Press, New York, NY, second edition, 1996.

[62] J. Kessler. *Untersuchung hochtemperatursupraleitender planarer Wellenleiter mittels Partialwellensynthese.* PhD thesis, Technische Universität München, 1992.

[63] J. Kessler, R. Dill, and P. Russer. Field theory investigation of hight-T_c superconducting coplanar waveguide transmission lines and resonators. *IEEE Trans. Microwave Theory Techn.*, 39(9):1566–1574, September 1991.

[64] K. Krohne and R. Vahldieck. A fast filter optimization scheme based on model order reduction. *2004 Int. Microwave Symposium Digest, Fort Worth*, pages 21–24, June 2004.

[65] M. Krumpholz. *Über die Grundlagen der TLM-Methode.* PhD thesis, Technische Univertität München, 1994.

[66] M. Krumpholz, B. Bader, and P. Russer. On the theory of the discrete TLM green's functions in three-dimensional tlm. *IEEE Trans. Microwave Theory Techn.*, 43(7):1472–1483, July 1995.

[67] M. Krumpholz and P. Russer. A field theoretical derivation of TLM. *IEEE Trans. Microwave Theory Techn.*, 42(9):1660–1668, September 1994.

[68] L. Kulas and M. Mrozowski. A simple high-accuracy subgridding scheme. *Proc. 33th European Microwave Conference, Munich*, pages 347–350, October 2003.

[69] L. Kulas and M. Mrozowski. Yee's macrocells in three dimensions. *2004 Int. Microwave Symposium Digest, Fort Worth*, pages 1717–1720, June 2004.

[70] S. Lindenmeier, C. Christopoulos, and P. Russer. Methods for the modeling of thin wire structures with the TLM method. *2000 Int. Microwave Symposium Digest, Boston*, pages 387–390, 2000.

[71] P. Lorenz. Combination of TLM method with multipole expansion of the radiation field. Master's thesis, Technische Universität München, 2003.

[72] P. Lorenz and P. Russer. Hybrid transmission line matrix (TLM) and multipole expansion method for time-domain modeling of radiating structures. *2004 Int. Microwave Symposium Digest, Fort Worth*, pages 1037–1040, June 2004.

[73] J. LoVetri and N. R. S. Simons. A class of symmetrical condensed node TLM methods derived directly from Maxwell's equations. *IEEE Trans. Microwave Theory Techn.*, 41(8):1419–1428, August 1993.

[74] D. Lukashevich, , and P. Russer. Network-oriented models of transmission line structures in MMICs. *Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, pages 178–181, April 2003. Garmisch, Germany.

[75] D. Lukashevich, B. Broido, and P. Russer. Using of transmission lime matrix method and mode matching approach for simulation of MMICs. *2003 Int. Microwave Symposium Digest, Philadelphia*, pages 993–995, June 2003.

[76] D. Lukashevich, L. Vietzorreck, and P. Russer. Numerical investigation of transmission lines and components in damascene technology. *Proc. 32th European Microwave Conference, Milan*, pages 885–889, September 2002.

[77] S. Le Maguer and M. Ney. PML-TLM technique : extension to evanescent waves and to advanced symmetrical condensed nodes. *3rd International Workshop on TLM Modelling*, October 1999. Nice.

[78] T. Mangold. *Feldmodellierung von verteilten Mehrtorschaltungen und systematische Extraction von Ersatzschaltungen aus konzentrierten Elementen.* PhD thesis, Technische Universität München, 2001.

[79] S.L. Marple. *Digital Spectral Analysis.* Prentice Hall, Englewood Cliffs, NJ, 1997.

[80] V. H. C. Melo, C. A. Tenorio de Carvalho Jr., and L.R.A.X. de Menezes. High performance TLM simulation using adjustable time steps. *2004 Int. Microwave Symposium Digest, Fort Worth*, pages 1029–1032, June 2004.

[81] L. Menezes and W. Hoefer. Modeling of general constitutive relationships in SCN TLM. *IEEE Trans. Microwave Theory Techn.*, 44(6):854–861, July 1996.

[82] B. C. Moore. Principal component analysis in linear systems: Controllability, observability and model reduction. *IEEE Trans. Automatic Control*, 26:17–32, 1981.

[83] J. A. Morente, J. A. Porti, and M. Khalladi. Adsorbing boundary conditions for the TLM method. *IEEE Trans. Microwave Theory Techn.*, 40(11):2095–2098, November 1992.

[84] I. Munteanu, T. Wittig, T. Weiland, and D. Ioan. FIT/PVL circuit parameter extraction for general electromagnetic devices. *IEEE Trans. Magnetics*, 39(7):1421–1424, July 2000.

[85] A. Odabasioglu, M. Celik, and L. T. Pileggi. PRIMA: Passive reduced–order interconnect macromodeling algorithm. *IEEE Trans. Computer-Aided Design*, 17:645–653, August 1998.

[86] A. V. Oppenheim and R. W. Schafer. *Digital Signal Processing.* Prentice Hall, New Jersey, 1975.

[87] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications.* Norwell, MA, Kluwer, 1996.

[88] B. N. Parlett. A new look at the Lanczos algorithm for solving symmetric systems of linear equations. *Linear Algebra Appl.*, 29:323–346, 1980.

[89] B. N. Parlett, D. R. Taylor, and Z.A.Liu. A look-ahead Lanczos algorithm for unsymmetric matrices. *Math. Comp.*, 44:105–124, 1985.

[90] N. Pena and M. M. Ney. Absorbing-boundary conditions using perfectly matched-layer (PML) technique for three-dimensional TLM simulations. *IEEE Trans. Microwave Theory Techn.*, 10:1749–1755, October 1997.

[91] L. Pierantoni, G. Gerri, S. Lindenmeier, and P. Russer. Theoretical and numerical aspects of the hybrid MOM-FDTD, TLM-IE and ARB methods for the efficient modelling of EMC problems. *Proc. 29th European Microwave Conference, Munich*, October 1999.

[92] L. Pierantoni, S. Lindenmeier, and P. Russer. A hybrid time domain TLM-integral equation method for solution of radiating problems. *Proc. 14th Annual Review of Progress in Applied Computational Electromagnetics, Monterey*, pages 320–325, 1998.

[93] L. T. Pillage and R. A. Rohrer. Asymptotic waveform evaluation for timing analysis. *IEEE Trans. Computer-Aided Design*, 9:352–366, 1990.

[94] J. N. Rebel. *On the foundation of the transmission line matrix method.* PhD thesis, Technische Universität München, 2000.

[95] J. N. Rebel, M. Aidam, and P. Russer. On the convergence of the classical symmetrical condensed node - TLM scheme. *IEEE Trans. Microwave Theory Techn.*, 49(5):954–963, May 2001.

[96] R. F. Remis and P. M. van den Berg. A modified Lanczos algorithm for the computation of transient electromagnetic wavefields. *IEEE Trans. Microwave Theory Techn.*, 45(12):2139–2149, December 1997.

[97] M. Rewienski and J. White. A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices. *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, 22(2):155 – 170, 2003.

[98] P. Russer. The alternating rotated transmission line matrix ARTLM scheme. *Electromagnetics*, 16(5):537–551, 1996.

[99] P. Russer. *The Transmission Line Matrix Method*, chapter Applied Computational Electromagnetics, pages 243–269. NATO ASI. Springer, Cambridge, Massachusetts, London, England, 2000.

[100] P. Russer and A. Cangellaris. Network–oriented modeling, complexity reduction and system identification techniques for electromagnetic systems. *Proc. 4th Int. Workshop on Computational Electromagnetics in the Time–Domain: TLM/FDTD and Related Techniques*, pages 105–122, 2001.

[101] P. Russer, P. So, and W. Hoefer. Modeling on nonlinear active regions in TLM. *IEEE Microwave And Guided-Wave Letters*, 1(1):10–13, 1991.

[102] Y. Saad. On the Lanczos method for solving symmetric linear systems with several right-hand sides. *Math. of Comput.*, 48(178):651–662, April 1987.

[103] Y. Saad. *Iterative Methods for Sparse Linear Systems*. The Pws Series in Computer Science. Pws Pub Co, 1996.

[104] R. Scaramuzza and A. J. Lowery. Hybrid symmetrical condensed node for the TLM method. *Electron. Lett.*, 28(23):1947–1949, July 1990.

[105] R. Schmidt. *Vollwellenanalyse von verlustbehafteten koplanaren Leitungen und Leitungsdiskontinuitäten*. PhD thesis, Technische Universität München, 1996.

[106] R. Schmidt and P. Russer. Modeling of cascaded coplanar waveguide discontinuities by the Mode-Matching approach. *IEEE Trans. Microwave Theory Techn.*, 43(12):2910–2917, December 1995.

[107] P. Sewell, J. G. Wykes, C. Christopoulos, T. M. Benson, D. W. P. Thomas, and A. Vukovic. TLM for unstructured triangular meshes. *Progress in Electromagnetics Research Symposium*, March 2004. Pisa, Italy.

[108] D.M. Sheen, S.M. Ali, M.D. Abouzahra, and J.A. Kong. Application of three–dimesional finite–difference time–domain method to the analysis of planar microstrip circiuts. *IEEE Trans. Microwave Theory Techn.*, 38(7):849–857, July 1990.

[109] R. N. Simons. *Coplanar Waveguide, Circuits, Components and Systems*. Wiley series in microwave and optical engineering. John Wiley and Sons, 2001.

[110] P. P. M. So and W. J. R. Hoefer. Time domain TLM modeling of metamaterials with negative refractive index. *2004 Int. Microwave Symposium Digest, Fort Worth*, pages 1179–1182, June 2004.

[111] T. Tang and M. Nakhla. Analysis of lossy multiconductor transmission lines using the AWE technique. *IEEE Trans. Microwave Theory Techn.*, 39(12):2107–2116, December 1991.

[112] G. Tardioli and W. Hoefer. Derivation of the SCN-TLM node scattering matrix from the integral formulation of Maxwell's equations. *The 2nd Int. Workshop on TLM - Theory and Applications*, pages 34–46, October 1997. Munich, Germany.

[113] W. Thiel, L. P. B. Katehi, and K. Sabet. A hybrid approach for modeling complex antenna systems on vehicular platforms. *2003 Int. Microwave Symposium Digest, Philadelphia*, pages 2097–2100, 2003.

[114] J. W. Thomas. *Numerical Partial Differential Equations.* Springer, New York, 1995.

[115] V. Trenkic. *The Development and Characterization of Advanced Nodes for the TLM Method.* PhD thesis, University of Nottingham, 1995.

[116] V. Trenkic, C. Christopoulos, and T. Benson. Theory of the symmetrical super-condensed node for the TLM method. *IEEE Trans. Microwave Theory Techn.*, 43(6):1342–1348, July 1995.

[117] V. Trenkic, C. Christopoulos, and T. Benson. Advanced node formulations in TLM - the adaptable symmetrical condensed node (ASCN) for the TLM method. *IEEE Trans. Microwave Theory Techn.*, 44(12):2473–2478, December 1996.

[118] V. Trenkic, C. Christopoulos, and T. Benson. Analytical expressions of the dispersion relation for the TLM condensed nodes. *IEEE Trans. Microwave Theory Techn.*, 44(12):2223–2230, December 1996.

[119] V. Trenkic, C. Christopoulos, and T. Benson. Development of a general symmetrical condensed node for the TLM method. *IEEE Trans. Microwave Theory Techn.*, 44(12):2129–2135, December 1996.

[120] K. L. Virga, M. Elzinga, and J. L. Prince. Improved global rational approximation macromodeling algorithm for networks characterized by frequency-sampled data. *IEEE Trans. Microwave Theory Techn.*, 48(9):1461–1468, 2000.

[121] H. F. Walker. Implementation of the GMRES method using Householder transformations. *SIAM J. Sci. Statist. Comput.*, 9:152–163, 1988.

[122] J. M. Wang, C. C. Chu, Q. Yu, and E. S. Kuh. On projection-based algorithms for model-order reduction of interconnects. *IEEE Trans. Circuits and Systems I*, 49(11):1563 – 1585, 2002. November.

[123] K. Weiss, S. Riedel, S. Schulz, H. Heldener, M. Schwerd, and T. Gessner. Characterization of electroplated and MOCVD cooper for trench fill in damascene architecture. *Proc. of the Materials for Advanced Metalization Conference*, pages 125–132, September 1999. Orlando, USA.

[124] J. H. Wilkinson. *The algebraic eigenvalue problem.* Oxford Science Publication. Clarendon Press, Oxford, 1965, reprinted 1992.

[125] T. Wittig, I. Munteanu, R. Schuhmann, and T. Weiland. Model order reduction and equivalent network generation for a FIT curl-curl formulation. *Proc. 18th Annual Review of Progress in Applied Computational Electromagnetics, Monterey*, pages 265–272, 2002.

[126] T. Wittig, I. Munteanu, R. Schuhmann, and T. Weiland. Two-step Lanczos algorithm for model order reduction. *IEEE Trans. Magnetics*, 38(2):673 – 676, March 2002.

[127] X. Yuan and Z. J. Cendes. A fast method for computing the spectral response of microwave devices over a broad bandwidth. *Proc. of the APS/URSI International Symposium*, page 196, June 1993. Ann Arbor, MI.

[128] B. Zermühl and S. Falk. *Matrizen und Ihre Anwendungen.* Springer Verlag, Berlin, Heidelberg, New York, Tokyo, 5-th edition, 1984.

[129] X. M. Zhang and J. F. Lee. Application of the AWE method with the 3-D-TVFEM to model spectral responses of passive microwave components. *IEEE Trans. Microwave Theory Techn.*, 46:1735–1741, November 1998.