Model Order Reduction Methods for Parameterized Systems in Electromagnetic Field Simulations

Zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation von Dipl.-Ing. Kynthia Katherina Stavrakakis
2012 — Darmstadt — D 17

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Elektrotechnik und Informationstechnik
Institut für Theorie Elektromagnetischer Felder (TEMF)

Model Order Reduction Methods for Parameterized Systems in Electromagnetic Field Simulations

Vom Fachbereich Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt

zur Erlangung
der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

von

Dipl.-Ing. Kynthia Katherina Stavrakakis
geboren am 15.03.1982 in Ostfildern, Stuttgart

Darmstadt 2012

Referent:	Prof. Dr.-Ing. Thomas Weiland
Korreferent:	Prof. Dr. Romanus Dyczij-Edlinger
Tag der Einreichung:	09 Dezember 2011
Tag der mündlichen Prüfung:	27 April 2012

D 17
Darmstädter Dissertation

Dedicated to my family

Contents

Kurzfassung

Das Thema dieser Arbeit fällt in das Gebiet der Systemapproximation. Ein ursprünglich großes und parameterabängiges Modell wird auf ein vereinfachtes kleineres Modell reduziert, welches die Übertragungsfunktion des ursprünglichen Modells approximiert, wobei die Parameterabhängigkeit im reduzierten Modell erhalten bleibt. Das vereinfachte Modell wird dann anstelle des ursprünglichen Modells in einer Vielzahl von Anwendungen verwendet.

Die in dieser Arbeit behandelten Modelle werden aus den Maxwell Gitter Gleichungen gewonnen, welche wiederum aus den kontinuierlichen Maxwell Gleichungen mit Hilfe der Methode der Finiten Integration (FIT) resultieren. Prinzipiell kann in elektromagnetischen Problemen eine Vielzahl von Größen als Parameter dienen. Diese Arbeit beschränkt sich jedoch auf Material- und Geometrie-Parameter, wobei Letzteres besondere Beachtung erhält.

Zu diesem Zweck werden zwei unterschiedliche Ansätze verfolgt. Der Erste basiert auf bereits bestehenden Ansätzen zur parametrischen Ordnungsreduktion, welche jedoch nicht auf Systeme in der direkt aus den Maxwell Gitter Gleichungen resultierenden Form angewendet werden können. Es wird ein Linearisierungsschritt entwickelt, welcher der Ordnungsreduktion vorausgestellt wird, mit dessen Hilfe die FIT- Systeme in die erwünschte Form gebracht werden können.

Ein alternativer Ansatz basiert auf der Verwendung der Systeme in der direkt aus den Maxwell Gitter Gleichungen erhaltenen Form. Die Projektionsmatrix des parametetrischen Systems wird als Komposition lokaler Projektionsmatritzen definiert. Diese Methode erweist sich als flexibler bezüglich der Variation geometrischer Parameter, wobei der Variationsspielraum zunächst auf die Größe der entspechenden Gitterlänge begrenzt ist. Es wurden jedoch Überlegungen angestellt, diese Limitierung aufzulockern.

Die in dieser Arbeit entwickelten Methoden wurden an mehreren Beispielen angewendet sowie untereinander verglichen.

Abstract

The topic of this thesis falls in the field of system approximation. The original large-scale and parameter-dependent model is reduced to a smaller model which approximates the transfer function of the original model while the parameter dependence is retained in the reduced model. The simplified model is then successfully used instead of the original model in a large variety of applications.

The models stem from the Maxwell grid equations, which are obtained from the continuous Maxwell equations with the help of the Finite Integration Theory (FIT). In general, a variety of parameters is possible for electromagnetic problems. In this work though, the focus has been set to material and geometry parameters with particular emphasis on the latter.

To this purpose, two main strategies are followed. The first strategy is based on already existing works on parametric model order reduction. These methods are not applicable to systems in the form naturally obtained by the Maxwell grid equations. Therefore, in order to be able to apply these methods, a linearization step is shown that appropriately adapts the FIT systems to the desired form.

An alternative approach is based on using the system in the form directly obtained from the Maxwell grid equations, and defines the projection matrix of the parametric system as the composition of local projection matrices. This method provides more flexibility in the geometry variation than the approach described above. Naturally, the variation range of geometrical parameters in this method is restricted to the mesh-cell size, though considerations have been performed to loosen this restriction.

The methods developed in this thesis have been applied to several numerical examples.

Chapter 1

Introduction

Unfolding the title of this thesis brings out three main subjects, namely electromagnetic field simulations, parametric systems and model order reduction methods. Numerical simulations are nowadays an indispensable tool in many engineering fields. They are used in the design process, where they allow for a faster optimization procedure, they shorten development cycles and reduce costs. The time consuming prototype construction is often avoided. Furthermore, they admit valuable insight into relations in cases where measurements are not feasible. Numerical simulations have also gained entry in the field of electrodynamics. The discrete equations on which they are based form a model which allows for a system-theoretic approach. In general, systems in practice are large, so that the need arises to replace them by smaller approximating systems. Model order reduction techniques are a common method for this purpose. This introductory chapter is intended to provide a quick insight into each of the three topics mentioned above.

1.1 Motivation

Modeling the real world by simple abstract representations is a main characteristic of human behavior and is being considered as the most important ability of Homo sapiens compared to its less developed predecessor, the Neanderthal. The cave-paintings in stone-age are the first abstract representations of real-world objects. The numbers, for counting and marking on bones, are the first recognizable models and were documented since about 30.000 BC. However, it was the Ancient Near East and Ancient Greek cultures which educed the modeling breakthrough.

With the independent development of astronomy and architecture by the Babylonians, Egyptians and Indians, as well as the development of the mathematical theory by the Greeks, a development was started, that is continued until today. Modeling started becoming mathematical, i. e. the real world was described by mathematical objects with the advantage of a precise analysis by means of mathematical tools. Sciences and mathematics followed thereafter a parallel mutually interacting development. Nowadays, mathematical modeling is widely used in a large variety of other disciplines, such as in engineering, in economics, in biosciences, but also in the social

sciences. The mathematical modeling of electromagnetic phenomena started in the 18th century. A brief historical overview will be given in the next chapter.

1.2 Systems

System (from the greek word $\sigma\acute{\upsilon}\sigma\tau\eta\mu\alpha$ (systima), "whole compounded of several parts or members, system", literary "composition" [33]) is a set of interacting or interdependent components forming an integrated whole. The word system has a long history which can be traced back to Plato, Aristotle and Euclid.

The concept of a "system" was first developed in the natural sciences in the 19th century by the French physicist Nicolas Léonard Sadi Carnot in the field of thermodynamics. In 1850, the German physicist Rudolf Clausius added the concept of the surroundings to the picture. However, it was not before 1945 that the biologist Ludwig von Bertalanffy pioneered with his general systems theory. He generalized the concept of systems to be irrespective of the particular kind or the nature of its elements, and the relations between them. Furthermore, models, principles, and laws that apply to them or their subclasses were introduced. Since then, significant development was achieved especially in the use of mathematics to study systems.

Naturally, the elements of a system are scoped from the elements of the surrounding, that is some entities are inside the system and some are outside. Formally, this can be achieved by viewing a system as an "exclusion law", declaring that some outcomes are possible while others are not [41, p. 1]. If $\mathbb{U}$ denotes the set in which the outcomes of a phenomenon are produced and $\mathfrak{B}$ the subset of $\mathbb{U}$ of all possible outcomes, called the behavior, a system can be formally defined as the pair $(\mathbb{U}, \mathfrak{B})$. The behavior $\mathfrak{B}$ can be viewed as a law excluding the occurrence of certain outcomes. In science, $\mathfrak{B}$ is usually described by equations, the behavioral equations. In this case, it contains those variables satisfying the equations. Different equations can define the same system. Of course, also behavioral inequalities exist.

One important class of systems are dynamical systems. A dynamical system is a system which evolutes in time, that is, the set $\mathbb{U}$ contains objects that are functions of time. The time axis needed to formalize the notion of the dynamical system is $\mathbb{T}$. For continuous-time systems $\mathbb{T} = \mathbb{R}$ or $\mathbb{R}_+$ and for discrete time systems $\mathbb{T} = \mathbb{Z}$ or $\mathbb{Z}_+$, or an interval in $\mathbb{R}$ or $\mathbb{Z}$, respectively. In this work only continuous time systems are considered, therefore it is assumed $\mathbb{T} = \mathbb{R}$. The family of time functions take their values in the space $\mathbb{W}$, which is called the signal space. Thus, the behavior $\mathfrak{B}$ is a subset of $\mathbb{W}^{\mathbb{T}}$ ($\mathbb{W}^{\mathbb{T}}$ is the standard mathematical notation for the collection of all maps from $\mathbb{T}$ to $\mathbb{W}$). With this notation, a dynamical system is formally defined as a triple $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$. The issue of dynamical systems will be revisited in chapter 3.

The notions of systems and mathematical models are often set equal, though it has to be considered, that both notions go along with different backgrounds, even if the practical result, namely the equations governing the system or the mathematical model, are the same.

1.2.1 Systems in Electrodynamics

In this thesis, the systems come from the field of macroscopic electrodynamics which is described by four fundamental equations, the Maxwell equations. They will be introduced in chapter 2, along with three constitutive equations that are required for their solution. Except for a few special cases, the time-dependent Maxwell equations in general form have no analytical solution. For this reason, they are discretized in space by mapping the continuous space to a finite set of discrete elements leading to a system of differential equations and, for transient problems, in time by introducing discrete instants in time yielding a set of difference equations. This work concentrates on spatially discretized systems. Then, it is tried to find approximate solutions to the Maxwell equations by solving the discrete equations. These are the Maxwell grid equations, which will also be introduced in chapter 2.

1.2.2 Systems obtained from the Discrete Electrodynamics Equations

All systems in this work stem from the Maxwell grid equations, which are formed by a set of partial differential equations. Different representations of this system exist, as will be shown in chapter 3. They all describe the same system, but they exhibit different numerical properties, as the system matrices differ. Furthermore, additional requirements, such as for instance the consideration of losses, result in different behavioral equations.

In the following, the various equations describing the systems treated in this work will be shown. The systems will be defined either in the time domain, or in the frequency domain, with the frequency parameter s. Details about the mathematics underlying this transformation, namely the Fourier and Laplace transforms, can be found e. g. in [67].

As will be shown in chapter 3, the system of Maxwell grid equations can be represented in the classical state-space form

$$\boldsymbol{\Sigma}_{\text{state-space}} : \left\{ \begin{array}{rl} s\mathbf{x} & = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{i}, \\ \mathbf{u} & = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{i}, \end{array} \right. \tag{1.2.1}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are the system matrices, the block-vectors $\mathbf{i}, \mathbf{u}$ are the input and output of the system, respectively, and s is the frequency parameter. The interpretation of these matrices will be explained in chapter 3 and 4.

Alternatively, second-order systems can be derived from the Maxwell grid equations:

$$\boldsymbol{\Sigma}_{\text{2nd-order}} : \left\{ \begin{array}{rl} (\mathbf{A}_2 s^2 + \mathbf{A}_1 s + \mathbf{A}_0)\mathbf{x} & = \mathbf{B}\mathbf{i}, \\ \mathbf{u} & = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{i}. \end{array} \right. \tag{1.2.2}$$

Notice, that it is not the identity matrix that corresponds to s and s^2, as it is the case in the classical state-space notation (1.2.1). With simple matrix multiplication the first equation of (1.2.2) can be reformulated to $s^2\mathbf{x} + s\mathbf{x} = \bar{\mathbf{A}}\mathbf{x} + \bar{\mathbf{B}}\mathbf{i}$. It depends on the context, which notation is preferred.

Higher-order systems with powers of $s > 2$ arise in connection with material losses or frequency-dependent material descriptions, as will be shown in chapter 2:

$$\boldsymbol{\Sigma}_{\text{higher-order}} : \begin{cases} \displaystyle\sum_{k=0}^{\mathrm{N_A}} \mathbf{A}_k \; s^k \mathbf{x} & = \displaystyle\sum_{k=0}^{\mathrm{N_B}} \mathbf{B}_k \; s^k \mathbf{i}, \\ \mathbf{u} & = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{i}. \end{cases} \tag{1.2.3}$$

The main topic of this work, however, are parametric systems. Formally, it can be distinguished between the case in which the frequency parameter s is included in the parameter set and the case in which s is viewed independently from the other parameters. In the first case the notation $\mathbf{s} = (s_1, \ldots, s_r)$ is used, where s is one of the parameters, and in the second case the parameters, excluded s, are contained in the vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{r-1})$.

Different system representations arise at this point. The most general form assumes that the matrices $\mathbf{A_s}$, $\mathbf{B_s}$, $\mathbf{C_s}$ and $\mathbf{D_s}$ depend on the parameters, either in a known or an unknown way. The general form of the resulting system is the following:

$$\boldsymbol{\Sigma}_{\mathbf{s}} : \begin{cases} \mathbf{A_s}(\mathbf{s})\mathbf{x} & = \mathbf{B_s}(\mathbf{s})\mathbf{i}, \\ \mathbf{u} & = \mathbf{C_s}(\mathbf{s})\mathbf{x} + \mathbf{D_s}(\mathbf{s})\mathbf{i}. \end{cases} \tag{1.2.4}$$

In this work, a special case with $\mathbf{A_s}, \mathbf{B_s}, \mathbf{C_s}$ and $\mathbf{D_s}$ being (multivariate) polynomials of time-invariant (frequency-independent) matrices is considered. These systems are called multi-parameter polynomial systems. As a special notation is needed for this kind of systems, they will be introduced in chapter 4.

If the alternative description with $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{r-1})$ is used, again with $\mathbf{A}_{\boldsymbol{\xi}}$, $\mathbf{B}_{\boldsymbol{\xi}}$, $\mathbf{C}_{\boldsymbol{\xi}}$ and $\mathbf{D}_{\boldsymbol{\xi}}$ being polynomials, the system is formulated as follows:

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}} : \begin{cases} \displaystyle\sum_{k=1}^{\mathrm{N_A}} \mathbf{A}_k(\boldsymbol{\xi}) s^k \mathbf{x} & = \displaystyle\sum_{k=1}^{\mathrm{N_B}} \mathbf{B}_k(\boldsymbol{\xi}) \; s^k \mathbf{i}, \\ \mathbf{u} & = \displaystyle\sum_{k=1}^{\mathrm{N_C}} \mathbf{C}_k(\boldsymbol{\xi}) \; s^k \mathbf{x} + \mathbf{D}\mathbf{i}. \end{cases} \tag{1.2.5}$$

In this case, the matrices $\mathbf{A}_k, \mathbf{B}_k, \mathbf{C}_k$ are time-invariant. Notice, that to each parameter vector $\boldsymbol{\xi}, \mathbf{s}$ corresponds one system, denoted by $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}, \boldsymbol{\Sigma}_{\mathbf{s}}$.

1.3 Complexity Reduction and System Approximation

Once a mathematical model or system is set-up, the question of solving it arises. The methods used for this purpose are nowadays assigned to the field of numerical analysis. As exact solutions of practical mathematical problems are often impossible to obtain, this field studies the numerical algorithms for approximating the solutions while maintaining reasonable bounds on errors.

The history of mathematical problem-solving shows that, while the mathematical model complexity steadily increased, the large computational requirements in most cases restricted the solvable problems to a very small number of cases. The development of numerical methods made it possible to automate the solution process and gave rise to the first computers, which, at the beginning of the last century were human workers. The methods often depended on hand interpolation in large printed tables. Nevertheless, the problem size was still limited. In 1945 the first Electronic Numerical Integrator and Computer (ENIAC) was introduced which revolutionized the mathematical problem solving. It was the first time that large practical problems were feasible by mathematical modeling. The field of numerical analysis and in particular numerical linear algebra has known an incredible development. The parallel development in mathematical methods and in computer science which followed enabled the solution of continuously increasing problem sizes and complexities, but this procedure is ever-ongoing. Due to either increased system size or an increased desire for detail there is always a demand to solve larger and more complex models. One example of this development is the use of increasingly higher frequencies in electronic circuits, which have a twofold impact on the mathematical model. On one side, the smaller wavelengths require a finer space discretization leading to increasingly larger systems of equations. On the other side, field effects have received relevance in fields that were once completely assigned to circuit simulation.

However, it is not only a matter of larger and faster computers. It is also a matter of efficient approaches, such that the given possibilities are fully exploited. Even today, where inconceivable sizes compared to the time of first computers are solvable, many dense numerical linear algebra techniques are only computationally feasible for a limited number of variables.

1.3.1 Problem Set-up

One of these efficient approaches, to which the topic of this thesis belongs, falls in a field which is broader known as complexity reduction or system approximation. The original complex model is reduced to a simplified model which captures the main characteristics of the original model. The simplified model is then successfully used instead of the original model in a large variety of applications. Apparently, depending on the use of the simplified model, an appropriate choice among a variety of reduction steps has to be done. If the model is intended to be used for a fast frequency sweep, where the model is used many times, the accuracy and the reduction time are of main interest. Obviously, the time to generate the reduced model should be less than the time to solve the original model. If the model is used in an optimization process, where the invariable parts of the structure are represented by a simplified model and only the remaining parts are optimized, then the preservation of stability or passivity of the system is more important than the reduction time.

In the past years, methods of model order reduction (MOR) have been developed to determine an approximate dynamical system $\hat{\mathbf{\Sigma}}$ by appropriately reducing the number of equations describing the initial system. The general formulation of the problem can be as follows: Given $\mathbf{\Sigma}$, e. g. in a form as described in the previous

section, with $\mathbf{x} \in \mathbb{R}^n$, determine a system $\hat{\mathbf{\Sigma}}$ in the same form as $\mathbf{\Sigma}$ with $\hat{\mathbf{x}} \in \mathbb{R}^m$ with $m \ll n$, such that the approximation error between the original and the reduced system is small in a predefined sense which will be investigated analytically in the respective chapter. Important properties of the original system such as stability and passivity should be preserved also for the reduced order model. Furthermore, the applied procedure should be computationally stable and efficient.

The most common MOR methods are based on projections in appropriate subspaces, as explained in the following. Figure 1.1 visualises the MOR principle.

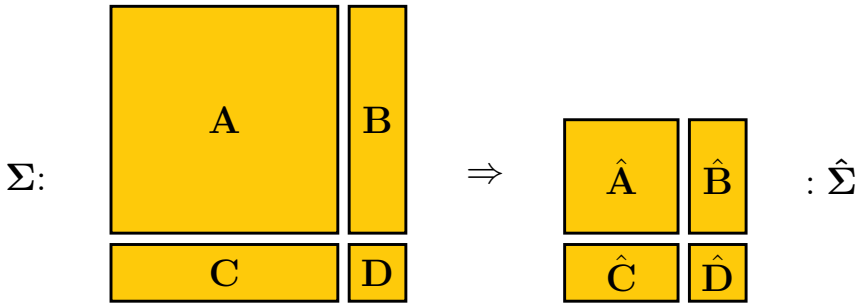


Figure 1.1: *Model order reduction of a dynamical system: a large scale system* $\mathbf{\Sigma}$ *is approximated by a reduced order system* $\hat{\mathbf{\Sigma}}$*. The approximation criteria will be discussed subsequently.*

1.3.2 Approximation by Projection

MOR methods by projection correspond to projection in an appropriate subspace and truncation. Their principle will be shown for the system (1.2.1) in state-space form. Let $\mathbf{x}$ live in $\mathbb{R}^{n\times 1}$ and consider the change of basis $\mathbf{T} \in \mathbb{R}^{n\times n}$ in the state-space $\bar{\mathbf{x}} = \mathbf{T}\mathbf{x}$. The quantities $\mathbf{x}, \mathbf{T}$ and $\mathbf{T}^{-1}$ are partitioned as follows:

$$\bar{\mathbf{x}} = \begin{pmatrix} \hat{\mathbf{x}} \\ \tilde{\mathbf{x}} \end{pmatrix}, \quad \mathbf{T}^{-1} = \begin{pmatrix} \mathbf{V} & \mathbf{T_1} \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} \mathbf{W}^{\mathrm{T}} \\ \mathbf{T}_2^{\mathrm{T}} \end{pmatrix}, \tag{1.3.6}$$

where $\hat{\mathbf{x}} \in \mathbb{R}^m, \tilde{\mathbf{x}} \in \mathbb{R}^{n-m}, \mathbf{V}, \mathbf{W} \in \mathbb{R}^{n\times m}$. Loosely speaking, it is aimed to pack the important information of the system in $\hat{\mathbf{x}}$ and the redundant information in $\tilde{\mathbf{x}}$. Substituting for $\mathbf{x}$ in (1.2.1) and retention of only the first m equations leads to:

$$\begin{cases} \hat{\mathbf{x}}(t) &= \mathbf{W}^{\mathrm{T}}\mathbf{A}(\mathbf{V}\hat{\mathbf{x}} + \mathbf{T}_1\tilde{\mathbf{x}}) + \mathbf{W}^{\mathrm{T}}\mathbf{B}\mathbf{i}, \\ \mathbf{u} &= \mathbf{C}(\mathbf{V}\hat{\mathbf{x}} + \mathbf{T}_1\tilde{\mathbf{x}}) + \mathbf{D}\mathbf{i}. \end{cases} \tag{1.3.7}$$

The approximation occurs by neglecting the term $\mathbf{T}_1\tilde{\mathbf{x}}$, i. e. the redundant part of the system. The resulting approximation is:

$$\hat{\mathbf{\Sigma}} : \begin{cases} \hat{\mathbf{x}}(t) &= \mathbf{W}^{\mathrm{T}}\mathbf{A}\mathbf{V}\hat{\mathbf{x}} + \mathbf{W}^{\mathrm{T}}\mathbf{B}\mathbf{i}, \\ \mathbf{u} &= \mathbf{C}\mathbf{V}\hat{\mathbf{x}} + \mathbf{D}\mathbf{i}. \end{cases} \tag{1.3.8}$$

Thus, the approximating matrices are: $\hat{\mathbf{A}} = \mathbf{W}^{\mathrm{T}}\mathbf{A}\mathbf{V}, \hat{\mathbf{B}} = \mathbf{W}^{\mathrm{T}}\mathbf{B}, \hat{\mathbf{C}} = \mathbf{C}\mathbf{V}$ and $\mathbf{D}$. Analogous results hold for the systems in (1.2.2) to (1.2.5). The choice of $\mathbf{V}$ and $\mathbf{W}$ depends on the system requirements and will be investigated in chapter 5.

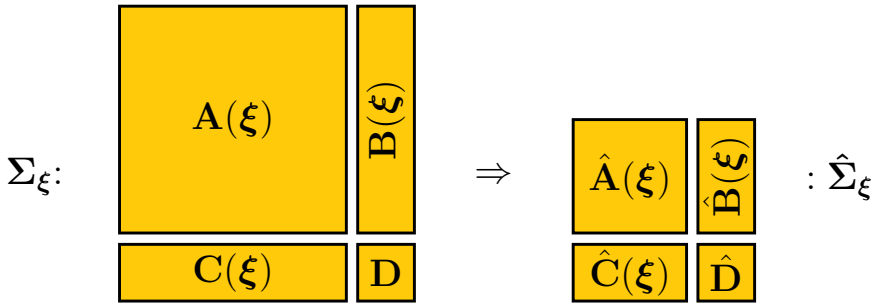


Figure 1.2: *Model order reduction of a parametric system with* $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_{r-1})$ *the vector containing the parameters, excluding the frequency parameter* s.

1.3.3 Problem Formulation for Parametric MOR

The same considerations hold for parametric systems, or sometimes called multivariate systems. As described in the previous section, several system representations exist, based on $\mathbf{s}$ or $\boldsymbol{\xi}$. A parametric MOR (PMOR) method consists in developing a system $\hat{\boldsymbol{\Sigma}}_{\mathbf{s},\boldsymbol{\xi}}$ by appropriately reducing the number of initial differential equations. The additional requirement is to retain the dependence on the parameter-vector $\boldsymbol{\xi}, \mathbf{s}$, respectively, in the reduced system, as visualized in figure 1.2 for $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$.

The reduced parametric dynamical system of (1.2.4) reads:

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{s}} : \begin{cases} \hat{\mathbf{A}}(\mathbf{s})\mathbf{x}(\mathbf{s}) &= \hat{\mathbf{B}}(\mathbf{s})\mathbf{i}(\mathbf{s}), \\ \mathbf{u}(\mathbf{s}) &= \hat{\mathbf{C}}(\mathbf{s})\hat{\mathbf{x}}(\mathbf{s}) + \hat{\mathbf{D}}(\mathbf{s})\mathbf{i}(\mathbf{s}), \end{cases} \tag{1.3.9a}$$

while for (1.2.5) it reads:

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}} : \begin{cases} \displaystyle\sum_{k=1}^{\mathrm{N_A}} \hat{\mathbf{A}}_k(\boldsymbol{\xi}) s^k \hat{\mathbf{x}}(\mathbf{s}) &= \displaystyle\sum_{k=1}^{\mathrm{N_B}} \hat{\mathbf{B}}_k(\boldsymbol{\xi})\ s^k \mathbf{i}(\mathbf{s}), \\ \mathbf{u}(\mathbf{s}) &= \displaystyle\sum_{k=1}^{\mathrm{N_C}} \hat{\mathbf{C}}_k(\boldsymbol{\xi})\ s^k \hat{\mathbf{x}}(s) + \hat{\mathbf{D}}\mathbf{i}(\mathbf{s}). \end{cases} \tag{1.3.9b}$$

Finally, figure 1.3 summarizes pictorially the set-up described in this introduction.

1.4 Dissertation Goals

Order reduction techniques for single-parameter systems have been analyzed thoroughly for many years. An overview of basic single-parameter MOR will be given in chapter 5. In the past years the interest centered upon reduction of systems depending on multiple parameters, whereas in this work the focus has been set to both material and geometry parameters. Different approaches are needed for these two parameter kinds: while material variations do not affect the mesh underlying the space discretization, geometry parameters definitely do. This eventually leads to matrices of different sizes. Most existing methods cannot cope with different-sized matrices. While a large number of works exists on this topic, in particular on

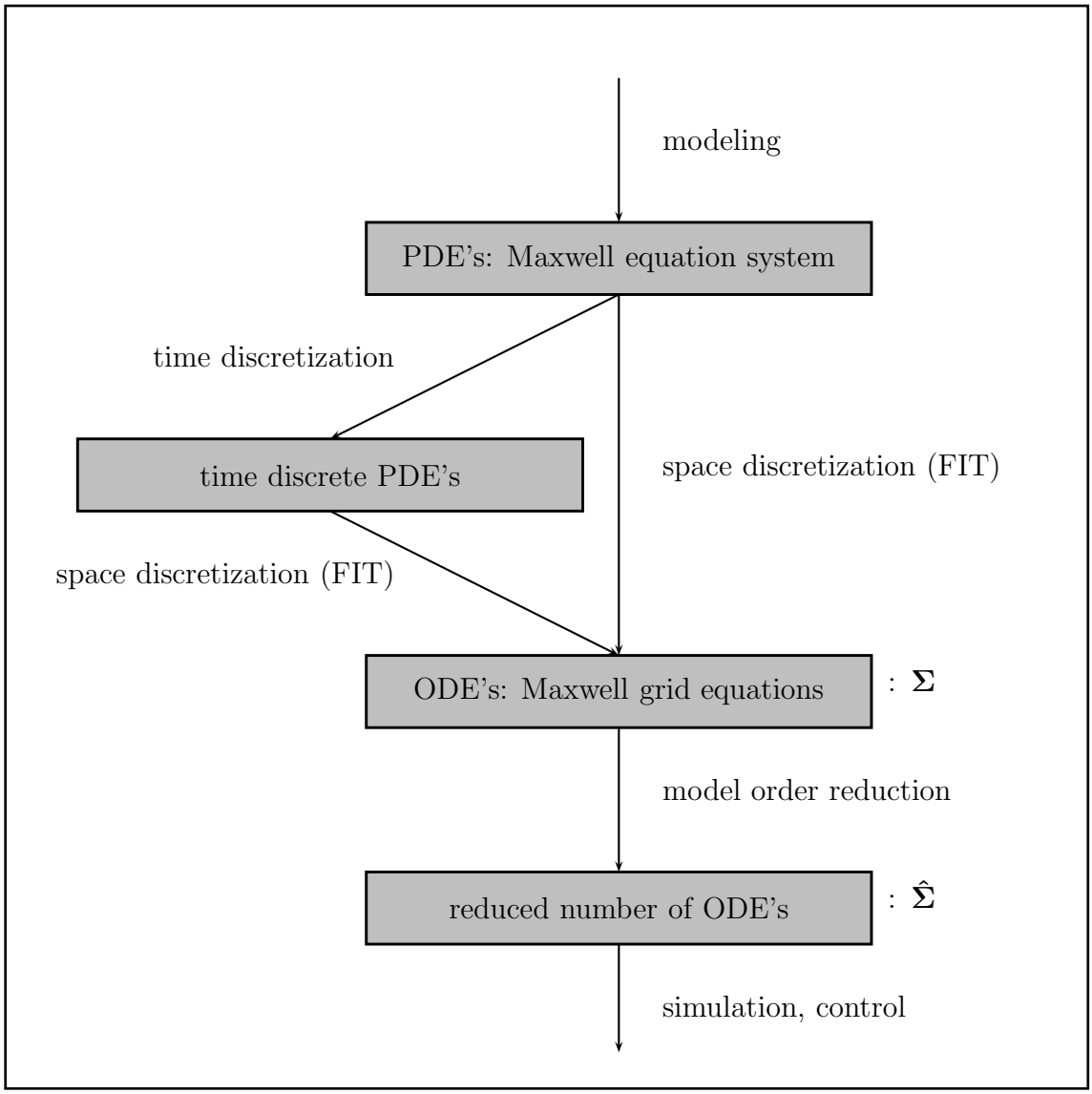


Figure 1.3: *The general set-up: from modeling electrodynamic phenomena to reduced order models*

material parametric systems, for geometrical parameterizations, investigations are still in an initial state.

In view of the highly practical relevance to vary the geometry, e. g. in the development phase of a device, and the lack of existing works taking into account this specialty of geometrical parameter variations, it was an informed decision to emphasize on them. Additionally, while there are several works related to parametric MOR for Finite Element (FE) systems (e. g. [17]), there has been so far no work on PMOR for systems resulting from the Finite Integration Theory (FIT). The cornerstone of this dissertation is, therefore, to develop parametric reduction techniques adapted to the needs of FIT systems, while paying particular attention on geometrical parameter variations.

To this purpose, two main strategies are followed. The systems naturally obtained by the Maxwell grid equations are in the form of equation (1.2.5), as will be shown in

chapter 4. However, already existing works on PMOR are not applicable to systems in this form. One strategy is, therefore, to appropriately adapt the FIT systems obtained directly from the Maxwell grid equations, so that these existing approaches can be applied. These methods are based on higher-order Krylov subspaces and guarantee that the leading moments, i. e. the Laurent series expansion coefficients around one fixed interpolation point $\mathbf{s}$, of the original and the reduced system are precisely the same. Unfortunately, these systems suffer from the difficulty to freely vary the geometry parameters. As will be shown in chapter 6 this has to do with the fact that these methods require a constant mesh topology. However, the computational grid does not necessarily remain the same during geometrical variations. Their usage is therefore limited for geometry parameter variations, while they are suitable for material variations.

An alternative approach is based on using the system in the form directly obtained from the Maxwell grid equations, equation (1.2.5), and defines the projection matrix of the parametric system as the composition of local projection matrices. This method provides more flexibility in the geometry variation, i. e. it does not require a constant mesh topology, it requires only a constant mesh size. Naturally, the variation range of geometrical parameters in this method is restricted to the mesh-cell size. However considerations have been performed to loosen this restriction. For this method, no moment matching properties between original and reduced system can be guaranteed. Nevertheless, the simplicity of this approach combined with the good agreement of original and reduced system is the main advantage in practice.

For each method a detailed description and the required supporting tools are given: techniques for choosing the points in the parameter range where the series coefficients are matched, or for which the singular value decomposition is performed, estimations for the modeling error, the stability issue, treating multiple-input multiple-output systems are some of these.

1.5 Outline

This introduction has given an overview of all topics covered in this thesis, which will be dealt with in detail in the following chapters. In chapter 2, the basic equations describing the continuous electrodynamics as well as their spatial discretization by the FIT will be described. Chapter 3 gives the fundamentals of dynamical systems relevant to the order reduction methods. In chapter 4 it is shown how the different classes of dynamical systems specified in the introduction arise from special applications of the FIT. A brief overview of classical MOR techniques is summarized in chapter 5. Parametric MOR techniques are developed in chapter 6 and tested on numerical examples in chapter 7. Finally, this work concludes with a summary and an outlook in chapter 8.

Chapter 2

Electrodynamics

The Maxwell equations describe all macroscopic phenomena of electromagnetism. In the first part of this chapter they are axiomatically stated along with three constitutive equations that are required for their solution. Except for a few special cases, the time-dependent Maxwell equations in general form have no analytical solution. For this reason they are discretized in space by mapping the continuous space to a finite set of discrete elements and, for transient problems, in time by introducing discrete instants in time. Numerical methods are used in order to find approximate solutions which are determined for every spatial element at each time instant. Numerous discretization techniques exist to discretize integral or differential equations in space. In this work the Finite Integration Technique (FIT) is used. The principles are described in the second part of this chapter.

2.1 Continuous Electrodynamics

While first observations both of electric as well as magnetic phenomena go back to the ancient ages by Thales of Miletus, except of some individual observations during the Middle Ages, it was in the 18th century that first quantitative statements were made by H. Cavendish (1773) and C. A. Coulomb (1785) in the field of electrostatics. In 1826, the relation of electric and magnetic fields was first stated by A. M. Ampère in a law that was later named Ampère's law. In 1831 the induction law was stated by M. Faraday, nowadays also known as Faraday's law. In 1873 James Clerk Maxwell established in his book "A Treatise on Electricity and Magnetism" a closed theory which combined all previous existing works and experimental results and is still nowadays the description basis of all macroscopic electromagnetic phenomena. Maxwell's main contribution, which enabled his closed theory, was to define the displacement current density $\vec{D}$.

2.1.1 The Maxwell Equations

The continuous Maxwell equations combine the magnetic field strength indicated by $\vec{H}$, the magnetic flux denoted by $\vec{B}$, the electric field strength $\vec{E}$, the dielectric flux

density or dielectric displacement $\vec{D}$, the electric current density $\vec{J}$ and the charge density ϱ. They read:

$$\int_{\partial A} \vec{E}(\vec{r},t)\cdot \mathrm{d}\vec{s} = -\int_{A} \frac{\partial}{\partial t}\vec{B}(\vec{r},t)\cdot \mathrm{d}\vec{A} \qquad \forall A \subset \mathbb{R}^3, \tag{2.1.1a}$$

$$\int_{\partial A} \vec{H}(\vec{r},t)\cdot \mathrm{d}\vec{s} = \int_{A} \left(\vec{J}(\vec{r},t) + \frac{\partial}{\partial t}\vec{D}(\vec{r},t)\right)\cdot \mathrm{d}\vec{A} \qquad \forall A \subset \mathbb{R}^3, \tag{2.1.1b}$$

$$\int_{\partial V} \vec{D}(\vec{r},t)\cdot \mathrm{d}\vec{A} = \int_{V} \varrho(\vec{r},t)\,\mathrm{d}V \qquad \forall V \subset \mathbb{R}^3, \tag{2.1.1c}$$

$$\int_{\partial V} \vec{B}(\vec{r},t)\cdot \mathrm{d}\vec{A} = 0 \qquad \forall V \subset \mathbb{R}^3. \tag{2.1.1d}$$

The first two equations are known as Faraday's and Ampère's laws, respectively, and relate the time differential of a field flux over an arbitrary surface area A to the voltage along its boundary ∂A. Equations (2.1.2c) and (2.1.2d) are named Gauss' law and Gauss' law of magnetism and relate the charge in an arbitrary volume V to the fluxes through their closed boundary ∂V. The latter states the absence of magnetic charges.

For stationary media, the Maxwell equations can be expressed equivalently in their differential form using the general Stokes theorem. Application of the Kelvin-Stokes theorem to Ampère's and Faraday's law and the Gauss-Ostrogradsky's theorem to Gauss' laws (2.1.2c) and (2.1.2d) yield the differential form of Maxwell's equations:

$$\nabla \times \vec{E}(\vec{r},t) = -\frac{\partial \vec{B}(\vec{r},t)}{\partial t}, \tag{2.1.2a}$$

$$\nabla \times \vec{H}(\vec{r},t) = \frac{\partial D(\vec{r},t)}{\partial t} + \vec{J}(\vec{r},t), \tag{2.1.2b}$$

$$\nabla \cdot \vec{D}(\vec{r},t) = \varrho(\vec{r},t), \tag{2.1.2c}$$

$$\nabla \cdot \vec{B}(\vec{r},t) = 0. \tag{2.1.2d}$$

The current density $\vec{J}$ in (2.1.1b) and (2.1.2b) can be decomposed into a conductive part $\vec{J}_\kappa$, a convective part $\vec{J}_c$ and the impressed current density $\vec{J}_e$:

$$\vec{J}(\vec{r},t) = \vec{J}_\kappa(\vec{r},t) + \vec{J}_c(\vec{r},t) + \vec{J}_e(\vec{r},t). \tag{2.1.3}$$

2.1.2 Material Equations

The equations described above are universally valid. They are independent of the material properties. The influence of the material is expressed in relations holding between the vectors $\vec{D}$ and $\vec{E}$, $\vec{B}$ and $\vec{H}$ as well as between the external field $\vec{E}_e$, the overall $\vec{E}$, and $\vec{J}$ describing the electric and magnetic properties of a given medium. From a mathematical viewpoint, equations (2.1.1) form a coupled set of

first-order linear partial differential equations. Therefore, each equation cannot be solved individually. The pairwise connection between $\vec{D}$ and $\vec{E}$, $\vec{B}$ and $\vec{H}$ enables their solution. The constitutive equations in the general case are given by

$$\vec{D}(\vec{r},t) = \varepsilon_0 \vec{E}(\vec{r},t) + \vec{P}(\vec{E},\vec{r},t), \tag{2.1.4a}$$

$$\vec{B}(\vec{r},t) = \mu_0 \vec{H}(\vec{r},t) + \vec{M}(\vec{H},\vec{r},t), \tag{2.1.4b}$$

$$\vec{J}(\vec{r},t) = \kappa \vec{E}(\vec{r},t). \tag{2.1.4c}$$

The material constants ε_0 and μ_0 are the permittivity and permeability of the free space while $\vec{P}(\vec{E},\vec{r},t)$ and $\vec{M}(\vec{H},\vec{r},t)$ denote the polarization and magnetization of the medium, respectively. While the first terms in (2.1.4a) and (2.1.4b) describe the linear contributions of the field, $\vec{P}$ and $\vec{M}$ describe the macroscopic behavior of the physical effects inside the materials and are, in general, time-variant, frequency-dependent and non-linear as well as non-isotropic functions of $\vec{E}$ and $\vec{H}$, respectively. Typically, modeling a material behavior of this kind is difficult, but in many cases materials with simplified behavior can be assumed. In case of a linear hysteresis-free material, $\vec{P}$ and $\vec{M}$ can be modeled with the help of the electric and magnetic susceptibility tensors χ_e and χ_m, respectively, using the convolution operator $*$:

$$\vec{P}(\vec{E},\vec{r},t) = \varepsilon_0 \chi_e(\vec{r},t) * \vec{E}(\vec{r},t) + \vec{P}_r(\vec{r}), \tag{2.1.5a}$$

$$\vec{M}(\vec{H},\vec{r},t) = \chi_m(\vec{r},t) * \vec{H}(\vec{r},t) + \vec{M}_r(\vec{r}). \tag{2.1.5b}$$

The terms $\vec{P}_r$ and $\vec{M}_r$ describe the permanent polarization or magnetization of a material and are assumed as zero in the following. The frequency dependence is obtained by considering the equations in frequency domain:

$$\vec{D}(\vec{r},\omega) = \varepsilon_0 \varepsilon_r(\vec{r},\omega) \cdot \vec{E}(\vec{r},\omega), \tag{2.1.6a}$$

$$\vec{B}(\vec{r},\omega) = \mu_0 \mu_r(\vec{r},\omega) \cdot \vec{H}(\vec{r},\omega), \tag{2.1.6b}$$

where ε_r and μ_r are the relative permittivity and permeability of the material, respectively. The frequency dependence is set by atomic and molecular interactions due to an external electric field. Various models exist to describe the resulting macroscopic behavior. Detailed descriptions can be found in [62]. Of importance in this thesis is the complex permittivity with which losses can be described:

$$\varepsilon = \varepsilon + \frac{\kappa}{j\omega}. \tag{2.1.7}$$

For frequency-independent isotropic material ε_r and μ_r are scalar time-invariant values and the constitutive equations read:

$$\vec{D} = \varepsilon_0 \varepsilon_r \, \vec{E}, \tag{2.1.8a}$$

$$\vec{B} = \mu_0 \mu_r \, \vec{H}. \tag{2.1.8b}$$

2.1.3 Boundary Conditions

With the help of Maxwell's integral equations the field behavior on the transition of two different media 1 and 2 with characteristic values $\varepsilon_1, \mu_1, \kappa_1$ and $\varepsilon_2, \mu_2, \kappa_2$ can be determined. Application of Faraday's and Ampère's laws on a surface intersecting the material interface results in the limit case in the following conditions for the fields:

$$\vec{n} \times (\vec{E}_2 - \vec{E}_1) = 0, \tag{2.1.9a}$$

$$\vec{n} \times (\vec{H}_2 - \vec{H}_1) = \vec{J}_A, \tag{2.1.9b}$$

where $\vec{n}$ is the unit normal vector of the interface pointing from medium 1 to medium 2, and $\vec{J}_A$ is the surface current density. Thus, the tangential electric field component is always continuous at material transitions, whereas the tangential magnetic field component is only continuous if there is no surface current on the interface. Analogously, application of Gauss' laws on a closed volume containing both material leads to the following conditions for the fluxes $\vec{D}$ and $\vec{B}$:

$$(\vec{D}_1 - \vec{D}_2)\,\vec{n} = \sigma_A, \tag{2.1.10a}$$

$$(\vec{B}_1 - \vec{B}_2)\,\vec{n} = 0, \tag{2.1.10b}$$

with σ_A being the surface charge density. Thus, while the magnetic flux is always continuous at the interface of two different media, the dielectric flux density $\vec{D}$ is only continuous if there is no surface charge.

2.1.4 Wave Equation

Combining the first two Maxwell equations in differential form, (2.1.2a) and (2.1.2b), and the material equations (2.1.4a) and (2.1.4c), the second-order forms of the Maxwell equations in a homogenous isotropic medium, e.g. vacuum, can be derived:

$$\varepsilon(\vec{r},t)\frac{\partial^2}{\partial t^2}\,\vec{E}(\vec{r},t) + \nabla \times \left(\mu^{-1}(\vec{r},t)\nabla \times \vec{E}(\vec{r},t)\right) = -\frac{\partial}{\partial t}\,\vec{J}(\vec{r},t), \tag{2.1.11a}$$

$$\mu(\vec{r},t)\frac{\partial^2}{\partial t^2}\,\vec{H}(\vec{r},t) + \nabla \times \left(\varepsilon^{-1}(\vec{r},t)\nabla \times \vec{H}(\vec{r},t)\right) = \nabla \times (\varepsilon^{-1}\vec{J}(\vec{r},t)). \tag{2.1.11b}$$

Assuming that the waves propagate far enough from sources, such that $\vec{J} = 0$ and $\varrho = 0$, (2.1.11) lead to:

$$\varepsilon(\vec{r},t)\frac{\partial^2}{\partial t^2}\,\vec{E}(\vec{r},t) + \nabla \times \left(\mu^{-1}(\vec{r},t)\nabla \times \vec{E}(\vec{r},t)\right) = 0, \tag{2.1.12a}$$

$$\mu(\vec{r},t)\frac{\partial^2}{\partial t^2}\,\vec{H}(\vec{r},t) + \nabla \times \left(\varepsilon^{-1}(\vec{r},t)\nabla \times \vec{H}(\vec{r},t)\right) = 0. \tag{2.1.12b}$$

Using the vector calculus relation

$$\nabla \times (\nabla \times \vec{a}) = \nabla(\nabla \cdot \vec{a}) - \triangle\vec{a} \tag{2.1.13}$$

and considering Gauss' laws, the following equations are obtained:

$$\frac{\partial^2}{\partial t^2}\vec{E}(\vec{r},t) - \frac{1}{\mu\varepsilon}\nabla^2\vec{E}(\vec{r},t), \tag{2.1.14a}$$

$$\frac{\partial^2}{\partial t^2}\vec{H}(\vec{r},t) - \frac{1}{\mu\varepsilon}\nabla^2\vec{H}(\vec{r},t). \tag{2.1.14b}$$

Thus, each component of $\vec{E}$ and $\vec{H}$ satisfies the wave equation with a velocity equal to $1/\sqrt{\mu\varepsilon}$.

2.1.5 Waveguides

Of special interest is the field propagation in metallic structures, known as waveguides . It is assumed that all boundary surfaces are perfectly electric conducting and that the profile does not change along the waveguide axis. The Maxwell equations for harmonic time dependence $\mathrm{e}^{-j\omega t}$ and for a guide containing a homogenous, loss-free medium, with permeability μ and permittivity ε, lead to the wave equation for $\vec{E}$ and $\vec{B}$:

$$(\nabla^2 + \mu\varepsilon\omega^2)\left\{ \begin{array}{c} \vec{E} \\ \vec{B} \end{array} \right\} = 0. \tag{2.1.15}$$

Waveguides support in general two field patterns: TE modes (Transverse Electric) with no electric field in the direction of propagation and TM modes (Transverse Magnetic), with no magnetic field in the direction of propagation. Other transverse modes which occur because of boundary conditions imposed on the wave by other waveguide types are the TEM and hybrid modes. TEM modes (Transverse Electro-Magnetic) have neither an electric nor a magnetic field in the direction of propagation. Hybrid modes have nonzero electric and magnetic fields in the direction of propagation. The TEM wave cannot exist in a single guide with perfectly conducting walls. Two or more conductors are required, as e. g. in a coaxial cable or a (ideal) microstrip line.

The superposition of eigenmodes can describe every field in a waveguide. With m designating the mode number, it is:

$$\vec{E}_t = \sum_{m=1}^{M} \vec{E}_{t,m}(a_m e^{-jk_w w} + b_m e^{jk_w w}), \tag{2.1.16a}$$

$$\vec{H}_t = \sum_{m=1}^{M} \vec{H}_{t,m}(a_m e^{-jk_w w} - b_m e^{jk_w w}), \tag{2.1.16b}$$

where a_m denotes the $m-th$ mode traveling in positive coordinate direction and b_m the corresponding part traveling in negative direction.

2.2 Discrete Electrodynamics

The objective of discrete electrodynamics is to find approximate solutions to the Maxwell equations. The first step consists in mapping the domain of interest Ω in the continuous space, $\Omega \subset (\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3)$, to a finite set of discrete elements $\mathcal{G}_i$ which depend on the dimensionality of Ω. In this way, a set of spatial elements is defined which is a topological structure in space and is referred to as computational grid $\mathcal{G}$. Let $\Omega_\mathcal{G}$ be the domain covered by the computational grid. In $\mathbb{R}^3$ the discrete volumes $\mathcal{G}_i$, also called grid cells, are associated with elementary points $P(i,j,k)$, edges $L(i,j,k)$ and surfaces $A(i,j,k)$.

The discrete elements have to be connected such that they completely cover the domain of interest, i. e. $\bigcup \mathcal{G}_i = \mathcal{G}$. However, the elements must not overlap, that is $\bigcap_{n_P} \mathcal{G}_i = \emptyset$. For practical reasons, typically basic geometrical shapes like triangles or rectangles in a two-dimensional domain and tetrahedrons or hexahedrons in three-dimensional domains are used.

Numerical methods are used to solve the resulting discrete equations for every spatial element at each time instance. The solutions comprise approximations to the continuous Maxwell equations. The discretization method of choice in this work is the Finite Integration Technique.

2.2.1 The Finite Integration Technique

The Finite Integration Technique (FIT) was introduced by T. Weiland in 1977 [69, 71]. Initially applied in the frequency domain for the numerical determination of eigenfrequencies in resonant structures, the range of applications was continuously extended. Today, an extremely wide range of electromagnetic components are successfully solved using the FIT. Applications include electro- and magnetostatics, stationary current problems, low and high frequency problems, as well as devices with movement of charged particles.

2.2.2 Spatial Discretization by Computational Grids

For the discretization of the Maxwell equations the FIT makes use of a staggered pair of grids, the primary grid $\mathcal{G}$ and the dual grid $\widetilde{\mathcal{G}}$. This idea goes back to the seminal paper of K.S. Yee in 1966, [79]. The principle idea of the FIT is to split the closed line integrals in the continuous Maxwell equations into integrals along the grid edges and the closed surface integrals into integrals over the grid surfaces. Hereby, Faraday's and Gauss' law of magnetism make use of the primary grid, while Ampère's and Gauss' law of the dual grid. In this way, the resulting discrete equations are exact representations of the continuous Maxwell equations. However, at this point they are still two pairs of uncoupled equations, as no connection has been assumed between the primary and the dual grid and no material conditions have been applied. The inevitable approximations going along with all numerical methods are introduced in the FIT by the discretization of the material relations

(2.1.4), as will be explained more detailed in section 2.2.5. At that point, the also necessary coupling between primary and dual grid is accomplished. However, imposing a specific relation between the two grids gives the opportunity to shape an efficient implementation.

The dual grid is constructed according to the following duality conditions: the grid points $\widetilde{P}$ of the dual grid $\widetilde{\mathcal{G}}$ are co-located with the centers of the primary grid cells. The orientation of a dual (primary) edge coincides with the orientation of the primary (dual) surface. Furthermore, every primary surface is cut by exactly one dual edge and every primary edge is cut exactly by one dual surface. In principal, the respective cutting angles can be arbitrary. If they are 90°, the orthogonality condition is met.

Given an arbitrary primary grid $\mathcal{G}$ the construction of a dual grid $\widetilde{\mathcal{G}}$ fulfilling the first three conditions is generally possible. This does not hold true for the orthogonality condition. Grids fulfilling all conditions are referred to as dual orthogonal. The grid set-up in the Cartesian coordinate system employs a staggered pair of dual orthogonal grids, as will be shown in the following.

Typically, a general, curvilinear, orthogonal coordinate system (u, v, w) is assumed, though in this work problems in the three-dimensional Cartesian space are considered. A suitable pair of dual orthogonal hexahedral grids $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ is used.

The primary grid edges are aligned with the Cartesian coordinates and lead to the following definition:

$$\mathcal{G} = \{(u(i), v(j), w(k)) \subset \mathbb{R}^3 | i \in \{1 \dots I\},\ j \in \{1 \dots J\},\ k \in \{1 \dots K\}\}. \quad (2.2.17)$$

The corresponding edges, surfaces and volumes are given by the following equations:

$$\begin{gathered} L_u(i) = \overline{u(i)u(i+1)}, \quad L_v(j) = \overline{v(j)v(j+1)}, \quad L_w(k) = \overline{w(k)w(k+1)}, \\ A_u(j,k) = L_v(j)L_w(k), \quad A_v(i,k) = L_u(i)L_w(k), \quad A_w(i,j) = L_u(i)L_v(j), \\ V(i,j,k) = A_u(i)A_v(j)A_w(k), \end{gathered} \quad (2.2.18)$$

with $i \in \{1 \dots I-1\},\ j \in \{1 \dots J-1\}, k \in \{1 \dots K-1\}$.

With I, J, K the number of points in each coordinate direction u, v and w, and i, j, k the corresponding indices, the n_P mesh points are numbered as follows:

$$n(i,j,k) = i + (j-1)I + (k-1)IJ. \quad (2.2.19)$$

Figure 2.1 shows the Cartesian grid with the above definitions.

The dual grid is defined by

$$\widetilde{\mathcal{G}} = \{(\tilde{u}(i), \tilde{v}(j), \tilde{w}(k)) \subset \mathbb{R}^3 | i \in \{1 \dots I\},\ j \in \{1 \dots J\},\ k \in \{1 \dots K\}\}, \quad (2.2.20)$$

with

$$\begin{aligned} \tilde{u}(i) &= \tfrac{1}{2}(u(i) + u(i+1)), \\ \tilde{v}(j) &= \tfrac{1}{2}(v(j) + v(j+1)), \\ \tilde{w}(k) &= \tfrac{1}{2}(w(k) + w(k+1)). \end{aligned} \quad (2.2.21)$$

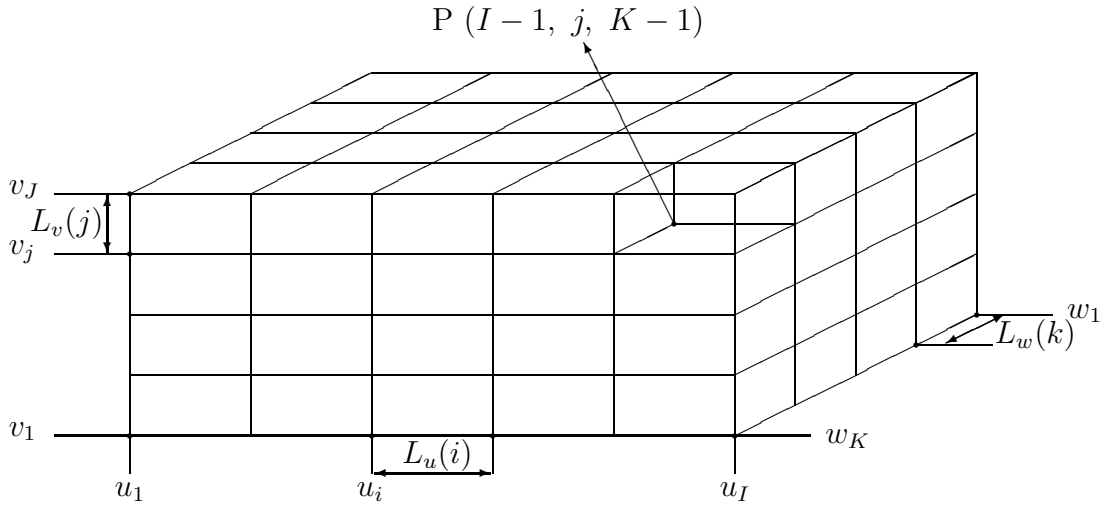


Figure 2.1: *The Cartesian grid used for discretizing a structure. The indication follows the notation of equations (2.2.17) and (2.2.18).*

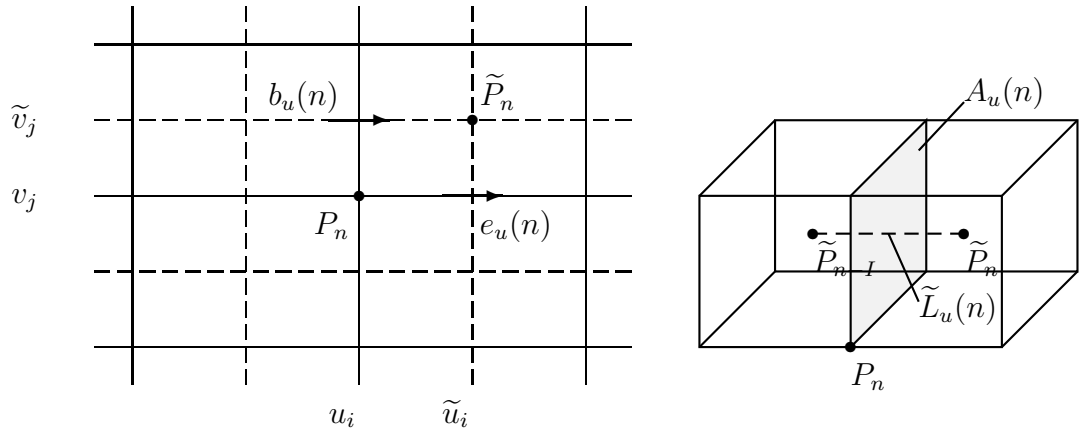


Figure 2.2: *Depiction of the primary grid $\mathcal{G}$ relative to the dual grid $\widetilde{\mathcal{G}}$; the numbering has been chosen such, that every dual mesh edge intersects the primary mesh surface with the same number.*

The dual edges, surfaces and volumes are defined analogously to (2.2.18).

Figure 2.2 shows the primary grid relative to its dual. In a first instance, homogeneous material is assumed in every resulting elementary volume. Other cases will be considered later on.

2.2.3 The Maxwell Grid Equations

2.2.3.1 Discretization of Faraday's law

Focusing on the discretization of Faraday's law and Gauss' law of magnetism using the primary grid, the FIT makes use of the electric voltage $\widehat{\mathrm{e}}_p$ along edge L_p, $p \in \{u, v, w\}$ and the surface integral of the magnetic flux over the surface A_p, denoted

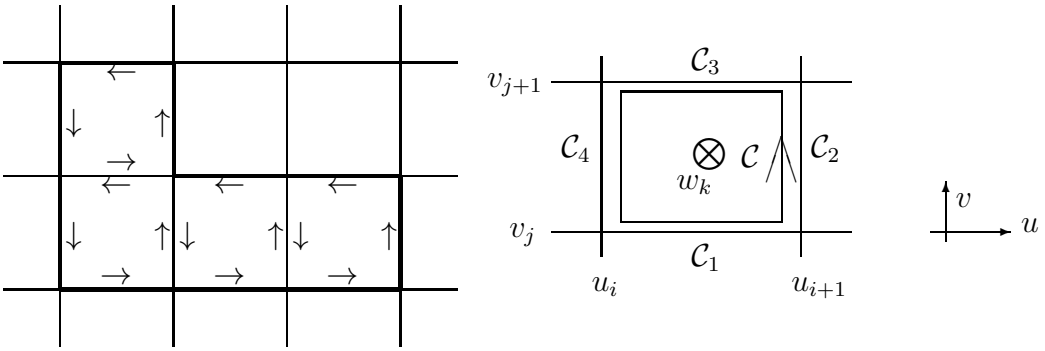


Figure 2.3: *Integration path for the derivation of the first Maxwell grid equation.*

by $\widehat{\widehat{\mathrm{b}}}_p$, respectively:

$$\widehat{\mathrm{e}}_p(i,j,k) = \int_{L_p(i,j,k)} \vec{E} \cdot \mathrm{d}s, \tag{2.2.22a}$$

$$\widehat{\widehat{\mathrm{b}}}_p(i,j,k) = \int_{A_p(i,j,k)} \vec{B} \cdot \mathrm{d}\vec{A}. \tag{2.2.22b}$$

Then, for one mesh cell, e. g. the cell depicted in figure 2.3, holds:

$$\widehat{\mathrm{e}}_u(i,j,k) + \widehat{\mathrm{e}}_v(i+1,j,k) - \widehat{\mathrm{e}}_u(i,j+1,k) - \widehat{\mathrm{e}}_v(i,j,k) = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathrm{b}}}_w(i,j,k). \tag{2.2.23}$$

According to the same figure, it is sufficient to consider only one mesh cell, as the line-integrals of directly neighbored cells are mutually canceling. Thus, if the induction law holds for one elementary surface, it holds for all surfaces. All electric voltages are grouped in a vector $\widehat{\mathbf{e}}$ and all magnetic fluxes in a vector $\widehat{\widehat{\mathbf{b}}}$, leading to the following equation

$$\mathbf{C}_{\mathrm{FIT}}\widehat{\mathbf{e}} = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathbf{b}}}, \tag{2.2.24}$$

which, as already mentioned, is an exact representation of Faraday's law. The matrix $\mathbf{C}_{\mathrm{FIT}}$ is sparse and singular and as apparent from equation (2.2.23), each row has two entries 1 and two -1. It is, thus, a topology matrix which plays the role of the curl operator in the discrete equations:

$$\mathbf{C}_{\mathrm{FIT}} = \begin{pmatrix} \emptyset & +1 & -1 & -1 & +1 & \\ -1 & +1 & \emptyset & +1 & -1 & \\ +1 & -1 & -1 & +1 & \emptyset & \end{pmatrix}. \tag{2.2.25}$$

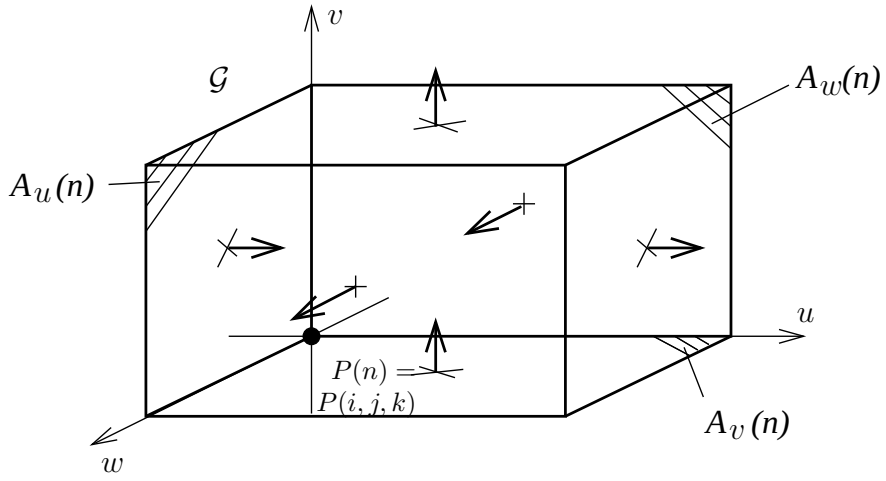


Figure 2.4: *In order to derive the discrete Gauss' law of magnetism, the surface integral of the magnetic flux over a primary grid cell shown above is split in six surface integrals over its cell-surfaces.*

2.2.3.2 Discretization of Gauss' law of magnetism

If in Gauss' law of magnetism the surface integral of the magnetic flux over a mesh cell is split in six surface integrals over the cell-surfaces, the discrete Gauss' law of magnetism is obtained. Again, as shown in figure 2.4, the surface-integrals of directly neighbored cells are mutually canceling, so that only one cell has to be considered.

The integrals over the six surfaces shown in figure 2.4 give:

$$\begin{aligned}\widehat{\widehat{\mathrm{b}}}_u(i,j,k) - \widehat{\widehat{\mathrm{b}}}_u(i+1,j,k) + \widehat{\widehat{\mathrm{b}}}_v(i,j,k) - \widehat{\widehat{\mathrm{b}}}_v(i,j+1,k) \\ + \widehat{\widehat{\mathrm{b}}}_w(i,j,k) - \widehat{\widehat{\mathrm{b}}}_w(i,j,k+1) = 0.\end{aligned} \tag{2.2.26}$$

The resulting grid equation is:

$$\mathbf{S}_{\mathrm{FIT}}\, \widehat{\widehat{\mathbf{b}}} = 0, \tag{2.2.27}$$

where $\mathbf{S}_{\mathrm{FIT}}$ is:

$$\mathbf{S}_{\mathrm{FIT}} = \left(\begin{array}{ccccccc} \overbrace{\quad}^{I} & & \overbrace{\quad}^{J} & & \overbrace{\quad}^{K} & & \\ & +1 & & +1 & & +1 & \\ -1 & & -1 & & -1 & & \end{array} \right), \tag{2.2.28}$$

a sparse topological matrix which corresponds to the divergence operator.

2.2.3.3 Discretization of Ampère's law and Gauss' law

For the discretization of the first Maxwell equation, the electric voltages $\widehat{\mathbf{e}}$ are mapped to the mesh edges while the components of the magnetic flux $\widehat{\widehat{\mathbf{b}}}$ to the surfaces of $\widetilde{\mathcal{G}}$. In order to evaluate the line and surface integral in the second Maxwell

equation, the components of $\widehat{\mathbf{h}}$ should be mapped to the grid edges and the components of $\widehat{\widehat{\mathbf{d}}}$ to the grid surfaces. At this point the introduction of the dual grid is required, with which the discretization of Ampère's and Gauss' law, i. e. the second and third Maxwell equation are accomplished. The magnetic voltage $\widehat{\mathrm{h}}_p$ along the dual edge $\widetilde{L}_p$, $p \in \{u, v, w\}$ and the surface integral of the dielectric flux $\widehat{\widehat{\mathrm{d}}}_p$ over the dual surface $\widetilde{A}_p$ are used.

Summarizing, the Maxwell grid equations read:

$$\mathbf{C}_{\mathrm{FIT}}\widehat{\mathbf{e}} = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathbf{b}}}, \tag{2.2.29a}$$

$$\widetilde{\mathbf{C}}_{\mathrm{FIT}}\widehat{\mathbf{h}} = \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathbf{d}}} + \widehat{\widehat{\mathbf{j}}}_e, \tag{2.2.29b}$$

$$\widetilde{\mathbf{S}}_{\mathrm{FIT}}\widehat{\widehat{\mathbf{d}}} = \mathbf{q}, \tag{2.2.29c}$$

$$\mathbf{S}_{\mathrm{FIT}}\widehat{\widehat{\mathbf{b}}} = 0, \tag{2.2.29d}$$

where the matrices $\widetilde{\mathbf{S}}_{\mathrm{FIT}}$ and $\widetilde{\mathbf{C}}_{\mathrm{FIT}}$ are the topology matrices representing the source (divergence) and the curl operator, respectively, corresponding to the dual grid $\widetilde{\mathcal{G}}$.

2.2.4 Properties of the Matrices

The solutions of the continuous Maxwell equations fulfill certain vector analytic relations, in particular

$$\nabla \cdot \nabla \times \vec{E} = 0 \quad \text{and} \quad \nabla \times \nabla\varphi = 0. \tag{2.2.30}$$

An important feature of the FIT is that discrete analogues of these equations exist for the grid spaces, i. e. they hold for both the primary and the dual grid. Furthermore, they guarantee that important physical properties, for instance energy or charge conservation and the validity of the continuity equation, are maintained in the Maxwell grid equations. They are derived in [73, 71] and read:

$$\mathbf{S}_{\mathrm{FIT}}\mathbf{C}_{\mathrm{FIT}} = \widetilde{\mathbf{S}}_{\mathrm{FIT}}\widetilde{\mathbf{C}}_{\mathrm{FIT}} = 0, \tag{2.2.31a}$$

$$\mathbf{C}_{\mathrm{FIT}}\widetilde{\mathbf{S}}^{\mathrm{T}}_{\mathrm{FIT}} = \widetilde{\mathbf{C}}_{\mathrm{FIT}}\mathbf{S}^{\mathrm{T}}_{\mathrm{FIT}} = 0. \tag{2.2.31b}$$

Finally, with respect to the matrices $\mathbf{C}_{\mathrm{FIT}}$ and $\widetilde{\mathbf{C}}_{\mathrm{FIT}}$ the duality condition of $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ can be expressed by

$$\widetilde{\mathbf{C}}_{\mathrm{FIT}} = \mathbf{C}^{\mathrm{T}}_{\mathrm{FIT}}. \tag{2.2.32}$$

2.2.5 Material Discretization

As mentioned earlier, the coupling of the primary and the dual quantities as well as the consideration of the intrinsic model data is accomplished by the discretized forms of equations (2.1.4). The calculation is accomplished in a two-step procedure.

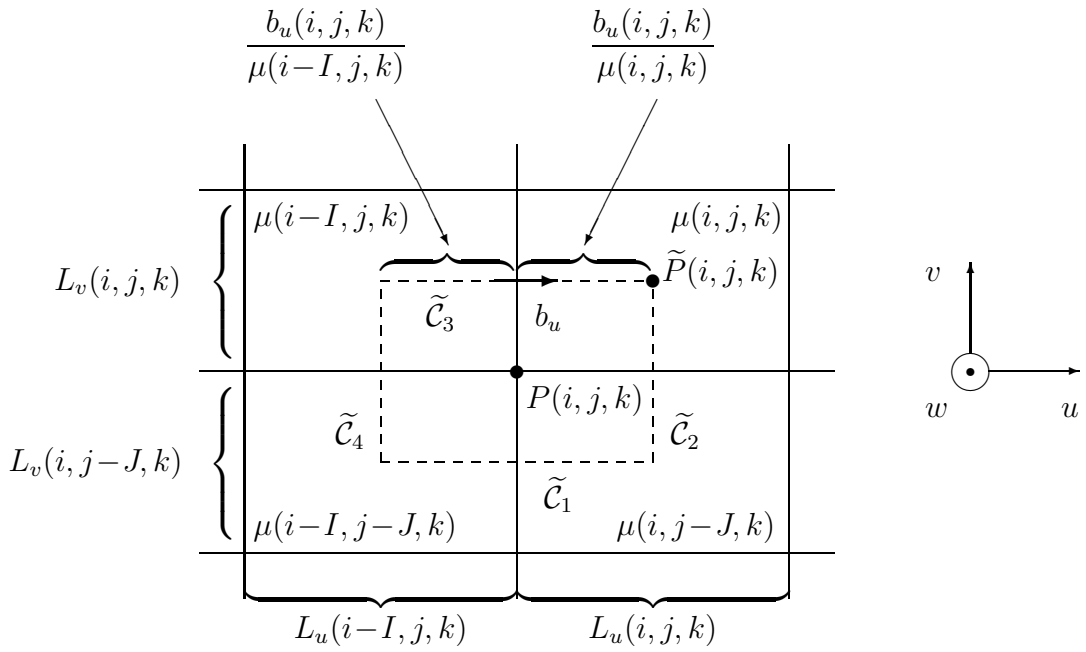


Figure 2.5: *Integration path for determination of the relation between the magnetic field voltage* $\widehat{\mathbf{h}}$ *defined on the dual grid and the magnetic flux* $\widehat{\widehat{\mathbf{b}}}$ *defined on the primary grid.*

On one hand, the electric voltage $\widehat{\mathbf{e}}$ and the flux quantity $\widehat{\widehat{\mathbf{b}}}$ have to be connected with the field density $\widehat{\widehat{\mathbf{d}}}$ and the field strength $\widehat{\mathbf{h}}$, respectively. On the other hand, a suitable averaging has to be assumed for the generally discontinuous, location-dependent material properties at the intersection points of edges and surfaces of both grids.

The procedure will be shown by using the example of the relation between $\widehat{\mathbf{h}}$ and $\widehat{\widehat{\mathbf{b}}}$. Figure 2.5 shows four primary cells intersected by one dual cell. The intersection points will be the integration path of $\int_{\partial A} \vec{H}(\vec{r},t) \cdot \mathrm{d}\vec{s}$. Splitting the integral into four parts corresponding to the primary cells, each of which may be filled with different material, and averaging the permeability along this part lines leads to the desired relation.

More precisely, every dual edge section $\mathcal{C}_1 \ldots \mathcal{C}_4$ traverses two primary cells, which may have different permeability values μ. Only the normal component of the magnetic flux is continuous at the intersection point. With the assumption that its variation in the vicinity of the intersection point is small, it is:

$$\widehat{\widehat{\mathrm{b}}}_u(i,j,k) = \int_{A_u(i,j,k)} \vec{B} \cdot \mathrm{d}\vec{A} \approx b_u A_u(i,j,k), \tag{2.2.33}$$

with b_u giving an approximation at the edge center.

The line integral of $\vec{H}$ along $\mathcal{C}_3$ is then:

$$\widehat{\mathrm{h}}_u(i,j,k) = \int_{\mathcal{C}_3} \vec{H}(\vec{r},t)\cdot \mathrm{d}\vec{s} \approx \frac{b_u}{\mu(i-I,j,k)}\widetilde{L}_u(i-I,j,k) + \frac{b_u}{\mu(i,j,k)}\widetilde{L}_u(i,j,k). \tag{2.2.34}$$

The averaged permeability can be defined as follows:

$$\overline{\mu^{-1}}(i,j,k) := \frac{\int(1/\mu)}{\int \mathrm{d}s}. \tag{2.2.35}$$

With (2.2.33) and (2.2.34) the relation between $\widehat{\widehat{\mathrm{b}}}$ and $\widehat{\mathrm{h}}$ follows:

$$\widehat{\widehat{\mathrm{b}}}_u(i,j,k) = \frac{A_u(i,j,k)}{\widetilde{L}_u(i,j,k)\overline{\mu^{-1}}(i,j,k)}\widehat{\mathrm{h}}_u(i,j,k). \tag{2.2.36}$$

In matrix notation the discrete analogue of (2.1.4) is obtained:

$$\widehat{\mathbf{h}} = \mathbf{M}_{\mu^{-1}}\widehat{\widehat{\mathbf{b}}}. \tag{2.2.37}$$

The material equations for $\mathbf{M}_\varepsilon$ and $\mathbf{M}_\kappa$ can be derived with an analogous consideration using dual faces and primary edges. In this case, the averaged permittivity $\frac{\iint \varepsilon \mathrm{d}A}{\iint \mathrm{d}A}$ and the averaged conductivity $\frac{\iint \kappa \mathrm{d}A}{\iint \mathrm{d}A}$ are used.
Summarizing, in case of linear, frequency-independent material, the discrete material relations read:

$$\widehat{\widehat{\mathbf{d}}} = \mathbf{M}_\varepsilon \widehat{\mathbf{e}}, \tag{2.2.38a}$$
$$\widehat{\mathbf{h}} = \mathbf{M}_{\mu^{-1}}\widehat{\widehat{\mathbf{b}}}, \tag{2.2.38b}$$
$$\widehat{\widehat{\mathbf{j}}} = \mathbf{M}_\kappa \widehat{\mathbf{e}} + \widehat{\widehat{\mathbf{j}}}_e. \tag{2.2.38c}$$

2.2.6 Material Matrices

The averaged inverse permeability values corresponding to each primary mesh cell are grouped in a diagonal matrix:

$$\mathbf{D}_{\mu^{-1}} = \mathrm{diag}\{\overline{\mu_u^{-1}}(1),\dots\overline{\mu_u^{-1}}(n_P),\overline{\mu_v^{-1}}(1),\dots\overline{\mu_v^{-1}}(n_P),\overline{\mu_w^{-1}}(1),\dots\overline{\mu_w^{-1}}(n_P)\}. \tag{2.2.39}$$

Analogously, the matrices $\mathbf{D}_\varepsilon$ and $\mathbf{D}_\kappa$ which contain the permittivity and conductivity values, respectively, of each mesh cell are defined.

The edge lengths and surface areas of the primary grid can also be grouped in diagonal matrices $\mathbf{D}_S$ and $\mathbf{D}_A$ respectively:

$$\mathbf{D}_S = \mathrm{diag}\{L_u(1),\dots,L_u(n_P),L_v(1),\dots,L_v(n_P),L_w(1),\dots,L_w(n_P)\}, \tag{2.2.40a}$$
$$\mathbf{D}_A = \mathrm{diag}\{A_u(1),\dots,A_u(n_P),A_v(1),\dots,A_v(n_P),A_w(1),\dots,A_w(n_p)\}. \tag{2.2.40b}$$

For the dual mesh, the diagonal matrices $\widetilde{\mathbf{D}}_S$ and $\widetilde{\mathbf{D}}_A$ are defined which contain the dual mesh edge lengths and surface areas.

With these definitions, the material matrices are given by:

$$\mathbf{M}_{\varepsilon} = \tilde{\mathbf{D}}_{\mathrm{A}}\mathbf{D}_{\varepsilon}\mathbf{D}_{\mathrm{S}}^{-1}, \tag{2.2.41a}$$

$$\mathbf{M}_{\mu^{-1}} = \tilde{\mathbf{D}}_{\mathrm{S}}\mathbf{D}_{\mu^{-1}}\mathbf{D}_{\mathrm{A}}^{-1}, \tag{2.2.41b}$$

$$\mathbf{M}_{\kappa} = \tilde{\mathbf{D}}_{\mathrm{A}}\mathbf{D}_{\kappa}\mathbf{D}_{\mathrm{S}}^{-1}. \tag{2.2.41c}$$

Thus, the Maxwell equations in terms of $\widehat{\mathbf{e}}$ and $\widehat{\mathbf{h}}$ with the material equations (2.2.41) read:

$$\mathbf{C}_{\mathrm{FIT}}\widehat{\mathbf{e}} = -\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}_{\mu^{-1}}^{-1}\widehat{\mathbf{h}}, \tag{2.2.42a}$$

$$\widetilde{\mathbf{C}}_{\mathrm{FIT}}\widehat{\mathbf{h}} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}_{\varepsilon}\widehat{\mathbf{e}} + \widehat{\widetilde{\mathbf{j}}}, \tag{2.2.42b}$$

$$\widetilde{\mathbf{S}}_{\mathrm{FIT}}\mathbf{M}_{\varepsilon}\widehat{\mathbf{e}} = \mathbf{q}, \tag{2.2.42c}$$

$$\mathbf{S}_{\mathrm{FIT}}\mathbf{M}_{\mu^{-1}}^{-1}\widehat{\mathbf{h}} = 0. \tag{2.2.42d}$$

2.2.7 Material Interface Modeling

2.2.7.1 Staircase approximation

Each primary mesh cell has so far assumed to be homogenously filled with one material. Material interfaces have thus always coincided with mesh cell interfaces. In this case, modeling material interfaces of two different media with an arbitrary interface shape, is straightforward. The percentage of the primary grid cell that is filled with material 1 decides whether the mesh cell is considered to be filled with material 1 or with material 2. If the percentage is less than 50% the cell is considered to be filled with material 1, otherwise, it is considered to be filled with material 2. Figure 2.6 shows the trivial case of the interface to perfectly electric conducting (PEC) material, as both the tangential electric voltages and the normal magnetic fluxes are zero. Thus, implicitly the boundary conditions are fulfilled. Furthermore, all entries of $\mathbf{M}_{\varepsilon}^{\mathrm{staircase}}$ corresponding to primary grid edges of PEC-filled cells are set to zero, while there is no need to modify $\mathbf{M}_{\mu}^{\mathrm{staircase}}$. Nevertheless, this method is inaccurate and therefore alternative methods are prefered.

2.2.7.2 Partially Filled Cells

Alternatively to the homogenously filled mesh cells described above, partial fillings in the mesh cells, with each volume-part possessing different material properties, can be considered. This feature was already incorporated in the classical FIT [69] where the partially filled cells (PFC) were tetrahedrons. In [63] an approach with arbitrary partial filled cells was presented on a hexahedral mesh. In the following, only the case of PEC material is considered.

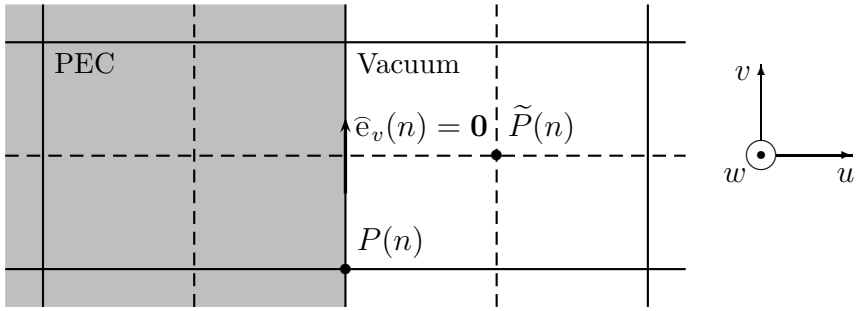


Figure 2.6: *Two adjacent primary cells in 2D. The left is filled with PEC material and the right with vacuum. Due to the boundary conditions, the tangential electric grid voltage* $\widehat{\mathbf{e}}_v$ *is zero.*

Perfectly electric conducting material

Figure 2.7, shows three constellations of the way an interface between PEC material and vacuum cuts the primary faces and edges. If the PEC region defines a small part of the mesh cell, then the magnetic flux which is allocated at the dual mesh center still lies in the vacuum region, otherwise it lies in the PEC region.

In 2D, Faradays' law can be applied with the reduced mesh edges and surfaces:

$$-\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\widehat{\mathrm{b}}}_m = \widehat{\mathrm{e}}_{i_1} + \widehat{\mathrm{e}}_{i_2} - \widehat{\mathrm{e}}_{i_3} - \widehat{\mathrm{e}}_{i_4}. \tag{2.2.43}$$

The factor $f_{A,m}$ describes the proportion of the mesh-cell surface A_m which is filled with perfectly electric conducting material. Analogously, $f_{L,i}$ gives the proportion of the mesh edges L_i which are inside the electric conducting material. The magnetic flux $\widehat{\widehat{\mathbf{b}}}_m$ and the electric voltage $\widehat{\mathbf{e}}_i$ are respectively:

$$\widehat{\widehat{\mathrm{b}}}_m = \int_{A_m f_{A_m}} \underline{\vec{B}}(x,y,z) \cdot \mathrm{d}\vec{A}, \tag{2.2.44a}$$

$$\widehat{\mathrm{e}}_i = \int_{L_i f_{L_i}} \underline{\vec{E}}(x,y,z)\mathrm{d}\vec{s}. \tag{2.2.44b}$$

Formally, the consideration of partially filled meshes simply corresponds to a modification of the material matrices:

$$\mathbf{D}_\mu(m,m)^{PFC} = \mathbf{D}_\mu(m,m)(1 - f_{A_m}), \tag{2.2.45a}$$

$$\mathbf{D}_\varepsilon(i,i)^{PFC} = \mathbf{D}_\varepsilon(i,i)\frac{1}{1 - f_{L_i}}. \tag{2.2.45b}$$

In case that the allocation points of some non-vanishing field vectors, which are necessary for calculating the material matrices, lie inside the perfect conductor, no physical interpretation can be applied to the approach of considering the partial fillings. Nevertheless, the equations are still valid.

As this approach to approximate the partial fillings results just in a modification of the material matrices, without changing their structure, there is no impact on the properties of the FIT.

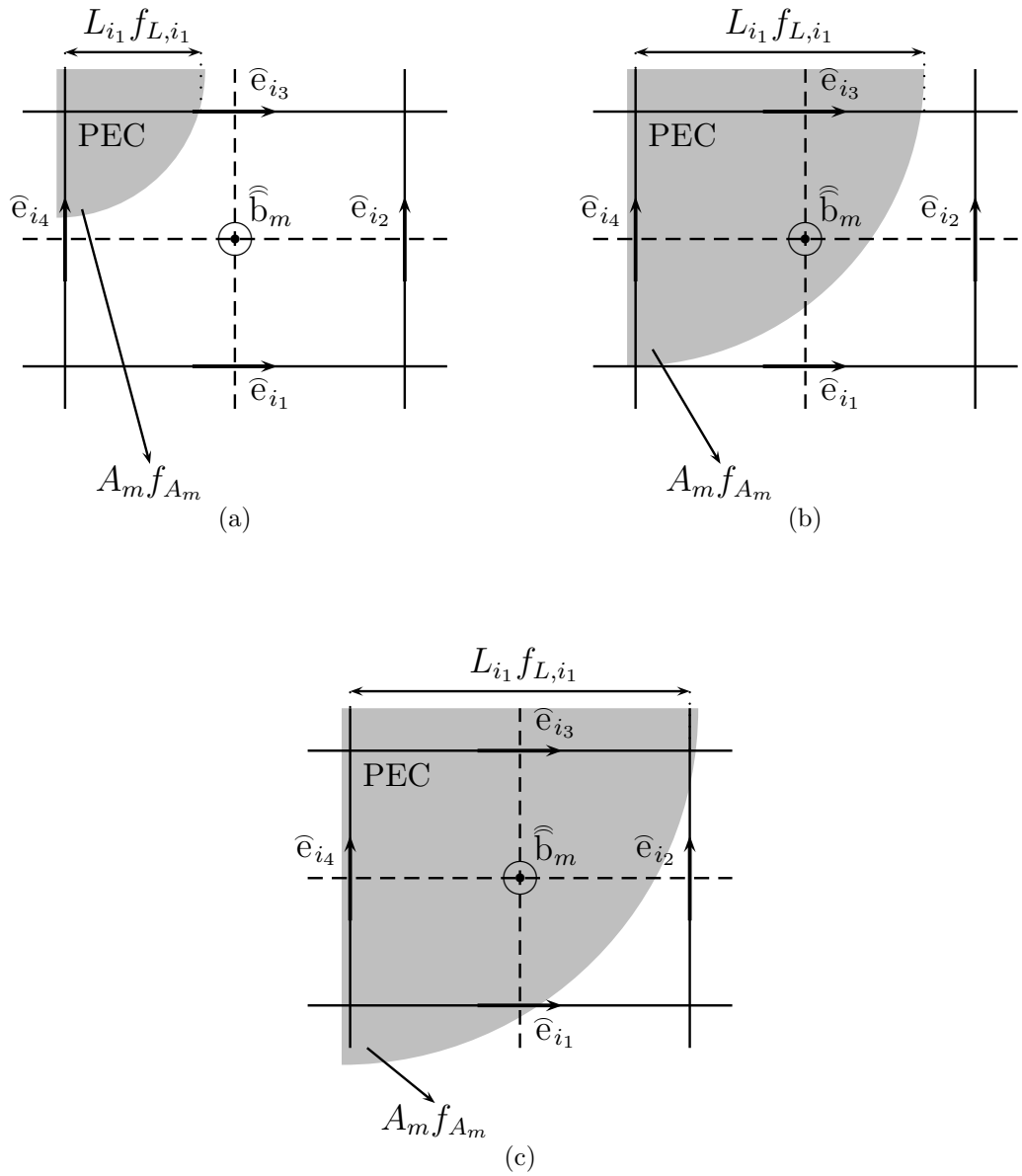


Figure 2.7: *The figure shows different constellations of primary cells which are partially filled with PEC material. In 2.7a the filling is less than* 50% *of the cell area. The dual edge is thus assigned to the vacuum region. In 2.7b and 2.7c the filling is more than* 50%*. In this case the magnetic grid voltage and the magnetic grid induction are assigned to the PEC region. In both cases the PEC-vacuum interface is associated with a reduced cell area* $A_m f_{A_m}$ *and shortened cell edges* $L_{i_1} f_{L_{i_1}}, i = 1 \ldots 4$.

2.2.8 Discrete Wave Equation

Analogously to the continuous wave equation (2.1.14), in the discrete Maxwell equations one of the vectors $\widehat{\mathbf{e}}$ or $\widehat{\mathbf{h}}$ can be eliminated, leading to the discrete wave

equation. For example, elimination of $\widehat{\mathbf{h}}$ leads to

$$\mathbf{M}_{\varepsilon} s^2 \, \widehat{\mathbf{e}} + \mathbf{M}_{\kappa} s \, \widehat{\mathbf{e}} + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}} \mathbf{M}_{\mu^{-1}}^{-1} \mathbf{C}_{\mathrm{FIT}} \widehat{\mathbf{e}} = s \, \widehat{\widehat{\mathbf{j}}}_s. \tag{2.2.46}$$

As the discrete curl operators $\mathbf{C}_{\mathrm{FIT}}, \widetilde{\mathbf{C}}_{\mathrm{FIT}}$ appears twice, this equation is known also as discrete curl-curl equation.

2.2.9 Boundary Conditions and Excitation of the Computational Domain

Analogously to the continuous space, boundaries are used to separate the computational domain from its surrounding. Boundary conditions are defined in order to determine the local fields and in order to model the influence of the surrounding space on the computational domain. The boundary conditions are as diverse as the situations they model.

In principle, depending on whether the energy exchange with the surroundings can take place or not, it is distinguished between open and closed boundary conditions. Most common closed boundary conditions are the Dirichlet and the Neumann boundary conditions. While the first forces the field solution to specific values on the boundary, the latter sets the normal field derivative at the boundary to a defined value. Two special cases are obtained in structures with perfectly electric or magnetic surrounding, i. e. the ideal electric or magnetic boundary conditions, respectively.

Lossy electric boundaries are assigned to the closed boundary conditions, even though energy leaks to the surroundings, yet it is not a free energy exchange. One example is the impedance boundary condition which yields from the ideal electric boundary with some additional assumptions. It is used to model metals with finite conductivity.

Of special interest are the boundaries modeling waveguides. In this case, a guided energy exchange in form of electromagnetic waves is desired. The waveguide port forms both a boundary condition as well as an excitation way.

Finally, a way to excite the computational domain from inside is given by the discrete ports. They are implemented by impressing a voltage at a primary edge or a current density at a dual surface.

2.2.9.1 Electric boundary condition

The electric boundary condition models structures surrounded by perfectly electric conducting material. It is implemented by using the Dirichlet boundary condition. All tangential electric voltages and all normal magnetic fluxes vanish. As these quantities are defined on the edges and surfaces of the primary grid, respectively, the corresponding entries in the matrices $\mathbf{C}_{\mathrm{FIT}}$ and $\mathbf{S}_{\mathrm{FIT}}$ can be set to zero. In this way though, the topological structure of these matrices is destroyed and therefore, preferably, the corresponding entries in the material matrices are set to zero. As

they are diagonal matrices, due to the zero entries the matrices get singular. In this case, the pseudo-inverse has to be used instead. Except of modeling the physical implementation of a perfectly conducting surrounding the electric boundary is often used as a symmetry condition. If, for example, symmetry considerations yield that the electric field of a grid plane contains only normal components, then this plane can be replaced by the electric boundary. In this way, the calculation domain is reduced, thus time and storage capacity is saved.

2.2.9.2 Magnetic boundary condition

Analogously, the magnetic boundary condition formally models structures which are surrounded by material with an infinite permeability. It is implemented by using the Neumann boundary condition. In this case, all tangential magnetic voltages and all normal electric fluxes vanish. Though, as no normal $\widehat{\mathbf{e}}$ components and no tangential $\widehat{\mathbf{h}}$ components are defined at the boundary, in this case, the corresponding entries in the material matrices cannot be set to zero. As shown in [73, p. 56], this boundary condition is implemented by considering reduced dual lengths and surfaces at the boundary, thus no modification of the material matrices is necessary.

Magnetic boundaries are mainly used as symmetry conditions in structures with symmetric fields following the same considerations as stated above for the electric boundaries.

2.2.9.3 Impedance boundary condition

The assumption of PEC material is not always possible. The impedance boundary condition gives the opportunity to consider losses in metallic structures. Making use of the complex permittivity (2.1.7), the wave vector of time-harmonic plane waves propagating in metals with usual frequencies $\kappa \gg \omega\varepsilon$ can be simplified as follows

$$\underline{k} = \omega\sqrt{\mu\varepsilon} \approx \sqrt{\frac{\omega\mu\kappa}{2}}(1-j) = \frac{1-j}{\delta}, \tag{2.2.47}$$

where $\delta = \sqrt{\frac{2}{\omega\mu\kappa}}$ is the skin depth. The wave length

$$\lambda = \frac{2\pi}{\mathrm{Re(k)}} \approx 2\pi\delta, \tag{2.2.48}$$

is much smaller than the corresponding wave length in free space, i. e. the wave is highly damped in the metal. Therefore, an extremely fine mesh would be required to capture this wavelength.

The implementation in the FIT is based on a surface impedance model proposed in [63]. In [75, p. 15 f.], a modification of this implementation is proposed that improves the bad matrix conditioning in the initial approach which is disadvantageous for model order reduction. As a result, frequency-dependent matrices $\mathbf{M}_\mu$ are obtained [75, p. 18].

2.2.9.4 Open boundary condition

There are basically three different ways to model free space in the FIT. One way is to make use of a boundary element method in combination with the FIT. Alternatively, a two-dimensional eigenvalue problem can be set-up on the boundary. The mode patterns solving this problem can be then impressed on the boundary. In this way an infinite extension of the two-dimensional structure as a waveguide can be modeled. Finally, a special absorbing material can be used to surround the computational domain. Each incident wave can penetrate without reflections the material and is then absorbed completely. An ideal boundary condition can be then applied to the absorbing material. This idea forms the basis of the perfectly matched layer (PML) boundaries. Details about the implementation of this approach in the FIT as well as general references to the topic can be found in [75, p.18 f.].

2.2.9.5 Waveguide boundary condition

An important issue is how to transfer the boundaries of waveguide structures, as described in section 2.1.5, to the discrete FIT model. Apparently, a special treatment is needed, as none of the previously described cases applies. The resulting boundary conditions are called waveguide ports and can be used for both terminating a computational domain as well as for exciting it.

Basically two options are possible: either a really open boundary is constructed through which energy can exit the computational domain, but also enter it for excitation. Alternatively, the infinite continuation of the waveguide is modeled by an equivalent closed problem with an impressed boundary current. This corresponds to the usage of the surface current density $\vec{J} = \vec{n} \times \vec{H}$ in continuous electrodynamics.

Both approaches, which were initially described in [70] and more detailed in [47], are based on the calculation of the two-dimensional transversal components $\widehat{\mathbf{e}}_t$ and $\widehat{\mathbf{h}}_t$ of $\widehat{\mathbf{e}}$ and $\widehat{\mathbf{h}}$, which, in analogy to the continuous case, can be formed by superposition of all discrete modes

$$\widehat{\mathbf{e}}_t(w) = \sum_{m=1}^{M} \widehat{\mathbf{e}}_{t,m}(w)(a_m \, \mathrm{e}^{-jk_w w} + b_m \, \mathrm{e}^{jk_w w}), \tag{2.2.49a}$$

$$\widehat{\mathbf{h}}_t(w) = \sum_{m=1}^{M} \widehat{\mathbf{h}}_{t,m}(w)(a_m \, \mathrm{e}^{-jk_w w} - b_m \, \mathrm{e}^{jk_w w}), \tag{2.2.49b}$$

where a_m denotes the $m-$th mode traveling in positive coordinate direction and b_m the corresponding part traveling in negative direction. Furthermore, the phasors $\widehat{\mathbf{e}}_{t,m}$ and $\widehat{\mathbf{h}}_{t,m}$ are two-dimensional vectors, which contain only transversal elements $\widehat{\mathrm{e}}, \widehat{\mathrm{h}}$ of the mode m. The quantities $\widehat{\mathbf{e}}_{t,m}(w), \widehat{\mathbf{h}}_{t,m}(w)$, $\widehat{\mathbf{e}}_t(w)$ and $\widehat{\mathbf{h}}_t(w)$ denote the $w-$th element of the respective matrices. The two-dimensional modes $\widehat{\mathbf{e}}_{t,m}, \widehat{\mathbf{h}}_{t,m}$ are the solutions to an eigenvalue problem that results by letting $\Delta w \to 0$. This is described in detail in [47].

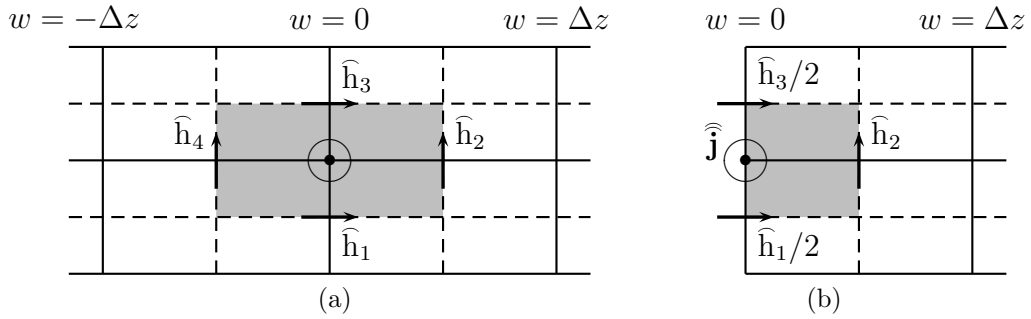


Figure 2.8: *The original and the equivalent structures for modeling the waveguide boundary condition with impressed current densities.*

It is made use of the discrete operator

$$\mathbf{N}_w = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}, \quad \text{with} \quad \mathbf{N}_w \widehat{\mathbf{h}}_t = \begin{pmatrix} -\widehat{\mathbf{h}}_v \\ \widehat{\mathbf{h}}_u \end{pmatrix}. \tag{2.2.50}$$

This is completely analogous to the continuous space:

$$\vec{n}_w \times \vec{H} \to \mathbf{N}_w \widehat{\mathbf{h}}_t, \tag{2.2.51a}$$

$$\vec{n}_w \cdot (\vec{E} \times \vec{H}) = \vec{E} \cdot (-(\vec{n}_w \times \vec{H})) \widehat{\mathbf{e}}_t^{\mathrm{T}} \to (-\mathbf{N}_w) \widehat{\mathbf{h}}_t. \tag{2.2.51b}$$

The amplitude of the phasor is arbitrary and is therefore usually chosen such, that the power in the waveguide takes the value 1, i. e.

$$\widehat{\mathbf{e}}_{t,m}^{\mathrm{T}} (-\mathbf{N}_w) \widehat{\mathbf{h}}_{t,m} = 1. \tag{2.2.52}$$

This relation corresponds to the integral over the Poynting-vector $\int \vec{n}_w (\vec{E} \times \vec{H}^*) \mathrm{d}A$.

Both vectors $\widehat{\mathbf{e}}_{t,m}$ and $\widehat{\mathbf{h}}_{t,m}$ can then be used to model the open boundary, as it is described in [16, 47, 63]. However, of relevance in this work is the alternative approach of modeling the closed problem, described in the following.

The standardization defined above (2.2.52) is valid for one single frequency, or equivalently, for a frequency-independent port impedance $Z = E^{\mathrm{T}}/H^{\mathrm{T}}$. This applies to the TEM modes and monofrequent solutions, but not to the TE- and TM-modes. Therefore, these cases have to be treated separately. Nevertheless, as the port impedance for both the TE- and the TM-modes is known analytically, a simple correction of the values calculated for the TEM and monofrequent case can be accomplished with the help of respective correcting factors. Therefore, first the TEM and monofrequent case will be considered.

TEM modes and mono-frequent solution

As $\widehat{\mathbf{h}}$ is allocated on the dual grid and therefore not known in the boundary surface, the original discrete structure shown in figure 2.8a is replaced by an equivalent structure depicted in figure 2.8b.

A comparison of the shadowed dual grids in both cases gives a relation for $\widehat{\hat{\mathrm{j}}}$:

$$\widehat{\hat{\mathrm{j}}} = \frac{1}{2} (\widehat{\mathrm{h}}_2 + \widehat{\mathrm{h}}_4). \tag{2.2.53}$$

For the $w-$th element of the vector $\widehat{\widehat{\mathbf{j}}}_e$ containing all $\widehat{\mathrm{j}}$ the change in the orientation corresponds to the application of the cross-product $\vec{J}_F = \vec{n} \times \vec{H}$, which can be expressed by the discrete operator $\mathbf{N}_w$. Thus, $\widehat{\widehat{\mathbf{j}}}_e$ is given by:

$$\widehat{\widehat{\mathbf{j}}}_e = \frac{1}{2}\mathbf{N}_w \left(\widehat{\mathbf{h}} \left(\frac{\Delta w}{2} \right) + \widehat{\mathbf{h}} \left(\frac{-\Delta w}{2} \right) \right). \tag{2.2.54}$$

With $\mathbf{a} = (a_1, a_2, \ldots, a_M)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_M)$ the matrices containing the mode-coefficients from (2.2.49) and $\mathbf{Z}_L^{1/2}$ the diagonal matrix containing $\sqrt{Z_{L,m}}$, the generalized current and voltage are defined according to:

$$\mathbf{u} = \mathbf{Z}_\mathrm{L}^{1/2}(\mathbf{a} + \mathbf{b}), \qquad \mathbf{a} = \frac{1}{2}(\mathbf{Z}_\mathrm{L}^{-1/2}\mathbf{u} + \mathbf{Z}_\mathrm{L}^{1/2}\mathbf{i}), \tag{2.2.55a}$$

$$\mathbf{i} = \mathbf{Z}_\mathrm{L}^{-1/2}(\mathbf{a} - \mathbf{b}), \qquad \mathbf{a} = \frac{1}{2}(\mathbf{Z}_\mathrm{L}^{-1/2}\mathbf{u} - \mathbf{Z}_\mathrm{L}^{1/2}\mathbf{i}). \tag{2.2.55b}$$

With this definition of $\mathbf{Z}_L^{1/2}$, in case of no reflections ($a_m = 1$, $b_m = 0$) at the ports, the port power is also $u_m \cdot i_m = 1$. With equations (2.2.55), the w-th element of $\widehat{\mathbf{h}}_{m,t}(a_m e^{(-jk_w w)} - b_m e^{(jk_w w)})$ in (2.2.49b) reads:

$$\begin{aligned}
&\widehat{\mathbf{h}}_{m,t}(a_m \,\mathrm{e}^{-jk_w w} - b_m \,\mathrm{e}^{jk_w w}) \\
=&\frac{1}{2}\widehat{\mathbf{h}}_{m,t} \left(\left(\frac{u}{\sqrt{Z_L}} + i\sqrt{Z_L} \right) \mathrm{e}^{(-jk_w w)} - \left(\frac{u}{\sqrt{Z_L}} - i\sqrt{Z_L} \right) \mathrm{e}^{(jk_w w)} \right) \\
=&\widehat{\mathbf{h}}_{m,t} \left(i\sqrt{Z_L \cos(k_w w)} - j\frac{u}{\sqrt{Z_L}} \sin(k_w w) \right).
\end{aligned} \tag{2.2.56}$$

Combined with equation (2.2.54), this result leads to an expression for the $w-$th element of the excitation vector $\widehat{\widehat{\mathbf{j}}}_{e,m}$ with respect to the $w-$th element of the two-dimensional vector $\widehat{\mathbf{h}}_{t,m}$, which has been calculated before

$$\widehat{\widehat{\mathbf{j}}}_{e,m} = i_m \sqrt{Z_{L,m}} \cos(k_{w,3D}\Delta w/2) \mathbf{N}_w \widehat{\mathbf{h}}_{t,m}, \tag{2.2.57a}$$

$$= i_m \sqrt{Z_{L,m}} \cos(k_{w,3D}\Delta w/2) \mathbf{b}_m^{2D}, \tag{2.2.57b}$$

where i_m is usually set to one and

$$k_{w,3D} = \frac{2}{\Delta w} \arcsin \left(\frac{k_w \Delta w}{2} \right). \tag{2.2.58}$$

This constant results from the adaptation of the two-dimensional propagation constant k_w, to the three-dimensional mesh, to capture the dispersion-effect and thus to avoid an energy error. The procedure is shown in [47]. The term $\cos(k_{w,3D}\Delta w/2)$ thus compensates the energy error made by performing the port-mode standardization (2.2.52) in the same plane $w = 0$, whereas the port-modes are allocated in spatially separated meshes. For fine discretizations its value is very close to one and it is therefore often neglected.

The vector $\widehat{\widehat{\mathbf{j}}}_e$ is two-dimensional. The third dimension is considered with the help of an operator $\mathbf{L}_k$ depending on the ports. The three-dimensional vectors $\widehat{\widehat{\mathbf{j}}}'_{e,m}$ and $\mathbf{b}_m$ are defined as follows:

$$\widehat{\widehat{\mathbf{j}}}'_{e,m} = \mathbf{L}_k \widehat{\widehat{\mathbf{j}}}_{e,m}, \quad \mathbf{b}_m = \mathbf{L}_k \mathbf{b}_m^{2D}. \tag{2.2.59}$$

As the modes are orthogonal, from the definitions (2.2.55) and equation (2.2.52), the generalized port voltage can be given

$$u_m = \sqrt{Z_{L,m}}(a_m + b_m) = \sqrt{Z_{L,m}}\mathbf{b}_m^{\mathrm{T}}\widehat{\mathbf{e}}. \tag{2.2.60}$$

TE- and TM- modes

The port impedance for TE- and TM- modes is respectively

$$Z_{w,\mathrm{TE}} = \frac{\omega\mu}{k_{w,m}(\omega)}, \quad Z_{w,\mathrm{TM}} = \frac{k_{w,m}(\omega)}{\omega\varepsilon}. \tag{2.2.61}$$

The frequency dependence of $k_{w,m}(\omega)$ is given by the dispersion equation

$$k_{w,m}(\omega) = \sqrt{\frac{(\omega^2 - \omega_{c,m}^2)k_{0,w,m}^2}{(\omega^2 - \omega_{c,m}^2)}}, \tag{2.2.62}$$

where $\omega_{c,m}$ is the cut-off frequency. The correcting factors for both mode types are given by:

$$F_{\mathrm{TE}} = \frac{\omega_0 k_{w,m}}{\omega k_{0,w,m}}, \quad F_{\mathrm{TM}} = \frac{\omega k_{0,w,m}}{\omega_0 k_{w,m}} = \frac{1}{F_{\mathrm{TE}}}. \tag{2.2.63}$$

These correcting factors have to be considered for each frequency in equations (2.2.57) and (2.2.60).

2.2.9.6 Discrete ports

While the waveguide ports are a convenient tool to terminate waveguides, sometimes one wishes to excite the computational domain at single internal points, e. g. to model feeding point sources for antennas, current or voltage sources. In this case, discrete ports are used, which impress a voltage at a primary edge or a current density at a dual surface. As for dual orthogonal grids the normal vector to the dual surface lies on the primary edge, the current density can be interpreted as flowing through the primary edge. This edge forms the discrete port path, to which an impedance is associated. As in practice the current density impression is preferred, the discrete port principally consists of a current source with an internal resistor.

Sometimes, e. g. for transmission lines at very low frequencies, discrete ports are more convenient than waveguide ports, despite the fact that the latter is more accurate. However, caution is advised, as at higher frequencies, for instance if the discrete port length is larger than a tenth of a wavelength, the S-parameters may be different than those resulting from using waveguide ports. This may happen due to higher levels of reflections at the port position, that is, the port and the structure are not matched properly.

Chapter 3

Dynamical Systems

In the introduction the notion of a dynamical system has been defined. The systems resulting from the Maxwell grid equations described in the previous chapter belong to a special class of dynamical systems, namely systems described by linear differential equations. In this chapter basic results concerning this class of systems will be given. Basically, two different description approaches of these systems exist: the external and the internal description. The former is associated with the system quantities one is interested in. This leads to the definition of input and output variables as well as a function connecting those two, the transfer function. The internal description is associated with variables which are not directly accessible. These variables are called auxiliary variables. One important special class of auxiliary variables is formed by state variables. Both descriptions are equivalent. It is rather the viewpoint that changes. In this chapter both approaches are discussed. Furthermore, important system concepts, such as stability and passivity, reachability and observability, are summarized. This information will be valuable in the next chapter.

3.1 Dynamical Systems defined by Linear Differential Equations

In the system defined by (2.2.42), which will be denoted by $\mathbf{\Sigma}_{\mathrm{FIT}}$ in the following, only the first two Maxwell equations are considered. The remaining equations play the role of auxiliary conditions. Due to the conditions in (2.2.31), they are satisfied provided that appropriate methods are used, exact calculations are performed and the excitation is consistent. Thus, $\mathbf{\Sigma}_{\mathrm{FIT}}$ can be written as:

$$\mathbf{\Sigma}_{\mathrm{FIT}} : \left\{ \begin{array}{ll} \mathbf{C}_{\mathrm{FIT}} \widehat{\mathbf{e}} & = -\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{M}_{\mu^{-1}}^{-1} \widehat{\mathbf{h}}, \\ \widetilde{\mathbf{C}}_{\mathrm{FIT}} \widehat{\mathbf{h}} & = \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{M}_{\varepsilon} \widehat{\mathbf{e}} + \mathbf{M}_{\kappa} \widehat{\mathbf{e}} + \widehat{\widehat{\mathbf{j}}}_{e}. \end{array} \right. \tag{3.1.1}$$

Based on this formulation, by using the curl-curl equation or by considering losses, higher-order formulations result. This will be subject to the subsequent chapter. Apparently, (3.1.1) is a special case of the general formulation given in the following.

With $\mathbf{R}$ a polynomial of degree k with real coefficients $a_i, i = 1 \dots k$:

$$\mathbf{R}\left(\frac{\partial}{\partial t}\right)\mathbf{w} = \sum_{i=0}^{k} a_i(t)\frac{\partial^i \mathbf{w}(t)}{\partial t^i}, \tag{3.1.2}$$

the general formulation is:

$$\mathbf{R}\left(\frac{\partial}{\partial t}\right)\mathbf{w} = \mathbf{0}. \tag{3.1.3}$$

Without going into detail how $\mathbf{w}$ looks like in particular, here general results about this class of systems will be given.

The admissible trajectories of (3.1.3) lie in the set of locally integrable functions $\mathfrak{L}_1^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^l)$, i. e. those functions $\mathbf{w} : \mathbb{R} \to \mathbb{R}^l$, which for all $a, b \in \mathbb{R}$:

$$\int_a^b |\mathbf{w}(t)|\mathrm{d}t < \infty. \tag{3.1.4}$$

Here, $|.|$ is the Euclidean norm on $\mathbb{R}^l$: $|\mathbf{w}| = \sqrt{\sum_{i=1}^{l} w_i^2}$.

Equation (3.1.3) defines the dynamical system $\mathbf{\Sigma} = (\mathbb{R}, \mathbb{R}^l, \mathfrak{B})$, with

$$\mathfrak{B} = \{\mathbf{w} \in \mathfrak{L}_1^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^l) | \mathbf{w} \text{ is a weak solution of } \mathbf{R}\left(\frac{\partial}{\partial t}\right)\mathbf{w} = 0\}, \tag{3.1.5}$$

where a weak solution is, simplified, a solution that satisfies (3.1.3) for almost all t, that is, except for a set of zero measure. For the exact definition see for instance [41, p. 34].

The systems described by (3.1.3) are linear, time-invariant (LTI). Linearity and time-invariance are important notions in applications. A dynamical system $\mathbf{\Sigma} = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is called linear if $\mathbb{W}$ is a vector space over $\mathbb{R}$ or $\mathbb{C}$ and $\mathfrak{B}$ is a linear subspace of $\mathbb{W}^{\mathbb{T}}$. A dynamical system $\mathbf{\Sigma} = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, $\mathbb{T} = \mathbb{R}$ or $\mathbb{Z}$, is called time-invariant if $\mathfrak{B}$ is shift-invariant. Simply speaking, in a time-invariant system, if a trajectory is in $\mathfrak{B}$, then the shifted trajectory is also in $\mathfrak{B}$. The proofs are straightforward and given for instance in [41, p. 42].

Given the system of Maxwell grid equations (3.1.1), it is not a priori clear which system description is best. Therefore, instead of viewing a system as a map between inputs and outputs, as it is done in the classical approach, here it is more appropriate to define the system by its behavior. The behavioral framework of mathematical systems is introduced thoroughly and comprehensively in the book by Willems and Poldeman [41]. Moreover, when introducing dependencies on material and geometry parameters, the resulting equations do not have the form used in the classical framework (state-space form). Therefore, again the behavioral approach seems to be more appropriate. Nevertheless, this short introductory chapter is not stuck strictly to the behavioral approach. It is rather used additionally for comprehension.

3.1.1 Input and Output Variables

If the polynomial matrix $\mathbf{R}$ is square and has full row rank, then two trajectories with the same past are identical. If on the other hand the rank of the polynomial matrix $\mathbf{R}$ is less than the number of its columns, parts of $\mathbf{w}$ are not uniquely determined by their past, i. e. the trajectories contain free components. This leads to the input/output representation of the system. Simply speaking, the trajectories $\mathbf{w}$ can be partitioned in two components $\mathbf{i}$ and $\mathbf{u}$. One component can be chosen freely, let it be $\mathbf{i}$ and be called the input, while the other, called the output $\mathbf{u}$, is completely determined by its past and the input. The formal definition is given in [41, p. 81]. The partition of $\mathbf{w}$ in input and output is not unique. Consider for example the case in the voltage/current behavior of a resistor, where either of the two variables can be chosen to be the input or the output.

In the previous chapter ports and boundary conditions have been used to describe the interactions of the FIT-system with its environment. Without restricting the general case, in this way naturally in- and output variables are defined. How they will enter the system description is subject to the next chapter.

3.1.2 Transfer Function

Every system in the form $\mathbf{R}\left(\frac{\partial}{\partial t}\right)\mathbf{w} = 0$, $\mathbf{R}$ being full row rank, admits a representation with input and output variables in the sense defined above [41, p.83 f.]. They have the form:

$$\mathbf{P}\left(\frac{\partial}{\partial t}\right)\mathbf{i} = \mathbf{Q}\left(\frac{\partial}{\partial t}\right)\mathbf{u}, \tag{3.1.6}$$

where $\det\mathbf{P}(\cdot) \neq 0$ and $\mathbf{P}^{-1}\mathbf{Q}$ is a matrix of proper rational functions, i. e. the degree of the numerator does not exceed the degree of the denominator. If furthermore, in each entry the degree of the numerator is strictly smaller that the degree of the denominator, the matrix is called strictly proper. The definition of the behavior is modified appropriately. The matrix $\mathbf{H} = \mathbf{P}^{-1}\mathbf{Q}$ is called the transfer function of the behavior/the system.

3.1.3 Convolution Systems and Impulse Response

Under certain conditions, it can be shown that systems defined by differential equations are equivalently described by an in- output map $\mathbf{h}$:

$$\mathbf{u}(\tau) = \int_{-\infty}^{\infty} \mathbf{h}(t,\tau)\mathbf{i}(\tau)\mathrm{d}\tau, \tag{3.1.7}$$

which is known as the impulse response of the system. The name is justified as for an input that is zero for all time instances except for $t = 0$, where it is a pulse, the output is exactly the function $\mathbf{h}$. Assuming that the system is causal, i. e. $\mathbf{h} = \mathbf{0}$ for

$t \leq 0$, and time-invariant, i. e. the function $\mathbf{h}(t, \tau)$ depends only on the difference of t and τ, the convolution sum is:

$$\mathbf{u}(\tau) = \int_{-\infty}^{t} \mathbf{h}(t - \tau)\mathbf{i}(\tau)\mathrm{d}\tau. \tag{3.1.8}$$

The conditions for the equivalence are that the admissible trajectories are restricted to those that $\mathbf{w} = \mathbf{0}, \forall t \leq t', t' \in \mathbb{R}$ (the systems are then called initially at rest) and furthermore that $\det(\mathbf{P}(\cdot)) \neq 0$ and $\mathbf{P}^{-1}\mathbf{Q}$ is strictly proper [41, p. 97]. In this case, the transfer function $\mathbf{H}$ is the Laplace transform of the impulse response [41, p. 281 f.], with the Laplace frequency parameter s, i. e.

$$\mathbf{H}(s) = \mathcal{L}(\mathbf{h}) = \mathbf{P}(s)^{-1}\mathbf{Q}(s). \tag{3.1.9}$$

As already mentioned in the introduction, details about this transformation, namely the Fourier and Laplace transforms, can be found e. g. in [67].

The transfer function plays an important role in system description and therefore also in MOR methods, as these methods use the terms of its Laurent series expansion either around $s = 0$, $s = \sigma$, ∞ or multiple expansion points. This will be explained in detail in the following chapters.

3.1.4 The Moments of a Function

Given a matrix valued function of t, $\mathbf{h} : \mathbb{R}^m \mapsto \mathbb{R}^{p \times m}$, its kth (generalized) moment is defined as:

$$\eta_k = \int_0^\infty t^k \mathbf{h}(t)\, \mathrm{e}^{-s_0 t}\, dt, \quad k = 0, 1, 2, \ldots. \tag{3.1.10}$$

If this function has a Laplace transform, the kth moment of $\mathbf{h}$ evaluated at $s = s_0$ reads:

$$\eta_k = (-1)^k \left.\frac{\mathrm{d}^k}{\mathrm{d}s^k}\mathbf{H}(s)\right|_{s=s0}. \tag{3.1.11}$$

As will be shown in the next chapter, this definition can be generalized to the multi-parameter case where besides s other parameters are considered.

The description presented above is related to the external description of the system, as the input and output are usually the variables that are measurable and that one is interested in. In most cases, other variables, called auxiliary variables, need to be introduced in addition to those intended to be modeled. As the auxiliary variables are not those in which one is particularly interested in and they are usually not directly measurable, the following description is referred to as internal description.

3.2 State-Space Models

3.2.1 Auxiliary Variables

The auxiliary variables, denoted by $\mathbf{x}$, lie in the space $\mathbb{U}_\mathrm{x}$. Together with $\mathbb{U}$ it defines $\mathbb{U} \times \mathbb{U}_\mathrm{x}$, the space in which $(\mathbf{w}, \mathbf{x})$ take values. Then, with $(\mathbb{U}, \mathbb{U}_\mathrm{x}, \mathfrak{B}_\mathrm{x})$ a mathematical system with auxiliary variables is defined, which is called an auxiliary representation of the system $\mathbf{\Sigma}$. The set $\mathfrak{B}$ is often called external behavior, as its elements are the directly measurable variables, while the set $\mathfrak{B}_\mathrm{x}$ is called internal behavior, as often its variables are implicit and not directly measurable. With this definition two descriptions of the system are possible, called the internal and the external description. Further details about these descriptions can be found in [1, p. 59]. For dynamical systems $\mathbb{U} = (\mathbb{T}, \mathbb{W})$ and $\mathbb{U}_\mathrm{x} = (\mathbb{T}, \mathbb{X})$, while the internal behavior is $\mathfrak{B}_\mathrm{x} \subset (\mathbb{W} \times \mathbb{X})^\mathrm{T}$.

With the auxiliary variable $\mathbf{x}$ the generalization of equation (3.1.3) is:

$$\mathbf{R}\left(\frac{\partial}{\partial t}\right)\mathbf{w} = \mathbf{M}\left(\frac{\partial}{\partial t}\right)\mathbf{x}, \tag{3.2.12}$$

where $\mathbf{M}$ is a polynomial of degree k' with real coefficients $b_i, i = 1 \ldots k'$:

$$\mathbf{M}\left(\frac{\partial}{\partial t}\right)\mathbf{x} = \sum_{i=0}^{k'} b_i(t)\frac{\partial^i \mathbf{x}(t)}{\partial t^i}. \tag{3.2.13}$$

The behavior corresponding to equation (3.2.12) is defined:

$$\mathfrak{B}_f = \{(\mathbf{w}, \mathbf{x}) \in \mathfrak{L}_1^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^l \times \mathbb{R}^n) \mid (\mathbf{w}, \mathbf{x}) \text{ satisfies 3.2.12 weakly}\}, \tag{3.2.14a}$$

$$\mathfrak{B} = \{\mathbf{w} \in \mathfrak{L}_1^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^l) \mid \mathbf{x} \in \mathfrak{L}_1^{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n) \text{ such that } (\mathbf{w}, \mathbf{x}) \in \mathfrak{B}_f\}. \tag{3.2.14b}$$

3.2.2 State-Space Models

One important special class of auxiliary variables is formed by state-variables. The concept of state plays an important role for dynamical systems. State variables, loosely speaking, parameterize the "memory" of the system, i. e. they split the past and the future of the behavior. The formal definition is given in [41, p. 119].

As has been mentioned, the variable $\mathbf{w}$ can always be partitioned in input and output variables. This way of viewing can be combined with the state concept leading to systems in input/state/output form. It can be shown, see for instance the notes on [41, p. 122], that every system described by (3.2.12) and having the property of state as defined in [41, p. 119] has a representation of the form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{i}, \tag{3.2.15a}$$

$$\mathbf{u} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{i}, \tag{3.2.15b}$$

which is the well-known classical state-space formulation. The variable $\mathbf{x}$ is the state which takes its values in $\mathbb{R}^n$, n being the order of the state-space representation, and $\mathbb{R}^n$ is the state-space. The state-space model is governed by a set of differential equations that are of first order in $\mathbf{x}$ and of zero order in $\mathbf{w}$ (therefore also of zero order in $\mathbf{i}, \mathbf{u}$). The first equation describes the dynamics of the system as well as the input. The second equation, also called the output equation, is a set of linear algebraic equations and describes the system output. The matrix $\mathbf{D}$ is called the feed-through term. Figure 3.1 illustrates the notion of a dynamical system in state-space form.

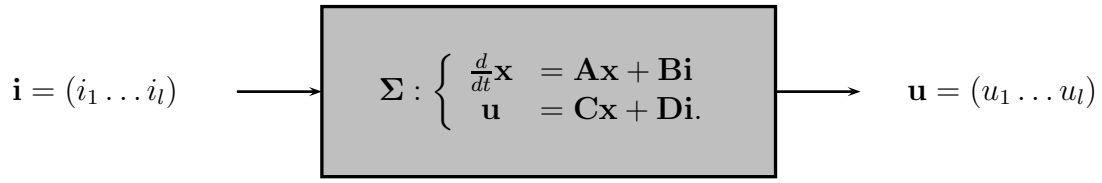


Figure 3.1: *Functional diagram used to describe dynamical systems in state-space form.*

In chapter 4, several representations stemming from the Maxwell grid equations, in particular $\mathbf{\Sigma}_{\mathrm{FIT}}$ from equation (3.1.1), will be given. According to that, there are also several possibilities to define the auxiliary variable $\mathbf{x}$. One possible representation is indeed the classical state-space form, yet other representations are possible. In principal, the same system is defined, but each representation exhibits different numerical properties, as the system matrices and the behavioral equations differ. Thus, although a transformation in the classical state-space form is possible for the FIT-system, it is important to notice, that other representations may be numerically more efficient, which turns true as will be shown in later chapters.

3.2.3 Solution of State-Space Equations

An explicit expression for the solution of the state equations is obtained by first considering the case $\mathbf{i} = \mathbf{0}$ and then the general case $\mathbf{i} \neq \mathbf{0}$. Let $x(\mathbf{i}, \mathbf{x}_0, t)$ be this solution, which depends on the input $\mathbf{i}$ and the initial state $\mathbf{x}_0$. This yields to the expression for $\mathbf{x}$:

$$\mathbf{x} = x(\mathbf{i}, \mathbf{x}_0, t) = \mathrm{e}^{\mathbf{A}(t-t_0)}\,\mathbf{x}_0 + \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}(t-\tau)}\,\mathbf{B}\mathbf{i}(\tau)\mathrm{d}\tau, \quad t \geq t_0, \tag{3.2.16}$$

and for the output:

$$\mathbf{u}(t) = \mathbf{C}x(\mathbf{i}, \mathbf{x}_0, t) + \mathbf{D}\mathbf{i}(t) = \mathbf{C}x(\mathbf{0}, \mathbf{x}_0, t) + \mathbf{C}x(\mathbf{i}, \mathbf{0}, t) + \mathbf{D}\mathbf{i}(t). \tag{3.2.17}$$

3.2.4 Reachability and Observability

There are a number of important concepts related to LTI systems which will be described in this section. First the concept of reachability will be introduced. It is related to the state equations, thus in this case only the matrices $\mathbf{A}$ and $\mathbf{B}$ are of importance.

A state $\bar{\mathbf{x}}$ is called reachable, if from an arbitrary zero state $\mathbf{x}_0$ with a finite input function $\mathbf{i}$ in a finite time T holds:

$$\bar{\mathbf{x}} = x(\mathbf{i}, \mathbf{x}_0, T). \tag{3.2.18}$$

The subspace containing all reachable states of $\mathbf{\Sigma}$ is called the reachable subspace $\mathbb{X}^{\text{reach}} \subset \mathbb{X}$ of the system, while it is called completely reachable if equality holds. The reachability matrix is defined as:

$$\mathcal{R}(\mathbf{A}, \mathbf{B}) = [\mathbf{B}\ \mathbf{AB}\ \mathbf{A}^2\mathbf{B}\ \ldots\ \mathbf{A}^{n-1}\ldots]. \tag{3.2.19}$$

As its column span is determined by the first n terms [1, p. 67], for computational purposes the finite reachability matrix

$$\mathcal{R}_n(\mathbf{A}, \mathbf{B}) = [\mathbf{B}\ \mathbf{AB}\ \mathbf{A}^2\mathbf{B}\ \ldots \mathbf{A}^{n-1}\,] \tag{3.2.20}$$

is of importance. The fundamental relation concerning reachability is that the reachability subspace is given by the image of the linear map corresponding to the reachability matrix, i. e. $\mathbb{X}^{\text{reach}} = \operatorname{im} \mathcal{R}(\mathbf{A}, \mathbf{B})$. An important concept is that of the reachability gramian:

$$\mathcal{P}(t) = \int_0^t \mathrm{e}^{\mathbf{A}\tau}\, \mathbf{B}\mathbf{B}^{\mathrm{T}}\, \mathrm{e}^{\mathbf{A}^{\mathrm{T}}\tau}\, \mathrm{d}\tau. \tag{3.2.21}$$

The reachability gramians are positive semidefinite, i. e. their eigenvalues are nonnegative and their columns span the reachability subspace, $\operatorname{im} \mathcal{P}(t) = \mathbb{X}^{\text{reach}} = \operatorname{im} \mathcal{R}(\mathbf{A}, \mathbf{B})$. In [1, p. 72] equivalent reachability conditions are summarized and proven.

Another important concept is that of observability. In this case, only the matrices $\mathbf{A}$ and $\mathbf{C}$ are of relevance. A state $\bar{\mathbf{x}}$ is called unobservable, if for all $t \geq 0$ it is $\mathbf{u}(t) = x(\mathbf{0}, \bar{\mathbf{x}}, t) = \mathbf{0}$. The subspace $\mathbb{X}^{\text{unobs}}$ containing all unobservable states of $\mathbf{\Sigma}$ is called the unobservable subspace of $\mathbf{\Sigma}$. The system is called completely observable if $\mathbb{X}^{\text{unobs}} = \mathbf{0}$. The observability matrix is defined as:

$$\mathcal{O}(\mathbf{C}, \mathbf{A}) = [\mathbf{C}^{\mathrm{T}}\ \mathbf{A}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}})^2\mathbf{C}^{\mathrm{T}}\ldots]^{\mathrm{T}}, \tag{3.2.22}$$

where again only the first n terms are of importance, yielding the finite observability matrix $\mathcal{O}_n(\mathbf{C}, \mathbf{A})$. The main result concerning observability is that $\mathbb{X}^{\text{unobs}} = \ker \mathcal{O}(\mathbf{C}, \mathbf{A}) = \{\mathbf{x} \in \mathbb{X} : \mathbf{C}\mathbf{A}^{i-1}\mathbf{x} = \mathbf{0}, i > 0\}$. The observability gramian is given by:

$$\mathcal{Q} = \int \mathrm{e}^{\mathbf{A}^{\mathrm{T}}\tau}\, \mathbf{C}^{\mathrm{T}}\mathbf{C}\, \mathrm{e}^{\mathbf{A}\tau}\, \mathrm{d}\tau, \quad t \in \mathbb{R}_+. \tag{3.2.23}$$

A main theorem for linear systems [1, p. 76] states that, for the systems $\mathbf{\Sigma}$ and $\mathbf{\Sigma}^{\mathrm{T}}$,

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}, \quad \mathbf{\Sigma}^{\mathrm{T}} = \begin{pmatrix} -\mathbf{A}^{\mathrm{T}} & -\mathbf{C}^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{D}^{\mathrm{T}} \end{pmatrix}, \tag{3.2.24}$$

the orthogonal complement of the reachable subspace of $\mathbf{\Sigma}$ equals the unobservable subspace of $\mathbf{\Sigma}^{\mathrm{T}}$, i. e. $\mathbb{X}_{\mathbf{\Sigma}}^{\mathrm{reach}} = \mathbb{X}_{\mathbf{\Sigma}^{\mathrm{T}}}^{\mathrm{unobs}}$. Furthermore, a system $\mathbf{\Sigma}$ is reachable if and only if $\mathbf{\Sigma}^{\mathrm{T}}$ is observable [1, p. 77].

Both the reachability and the observability subspaces are very important tools for MOR, as they equal to the Krylov subspaces of $\mathcal{K}_q(\mathbf{A}, \mathbf{B})$ and $\mathcal{K}_q(\mathbf{A}^{\mathrm{T}}, \mathbf{C}^{\mathrm{T}})$, respectively. The reduction techniques based on Krylov subspaces will be described in chapter 5.

3.2.5 S-Parameters

The voltage and current are often chosen as input and output variables in circuit-based systems describing applications in electrodynamics. Nevertheless, in practical high-frequency applications voltages, currents and impedances cannot be measured in a direct way and are therefore considered as secondary or derived quantities. What can be measured directly are the reflection and transmission coefficients [12, p. 248], that is, the relation of the amplitudes and phase angles of the reflected or scattered waves from a junction to the amplitudes and phase angles of the incident wave. As the field equations and most microwave devices are linear, there is a linear relation between the scattered-wave amplitudes and the incident-wave amplitudes. This relationship is described by the scattering matrix $\mathbf{S}$:

$$\mathbf{b} = \mathbf{S}\mathbf{a} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \tag{3.2.25}$$

where $\mathbf{a}$ and $\mathbf{b}$ are the incident-wave amplitudes and reflected-wave amplitudes respectively. Then, the diagonal entries represent the reflection coefficients, while the other entries are the transmission coefficients between two ports. The scattering matrix is symmetric, i. e. $\mathbf{S}^{\mathrm{T}} = \mathbf{S}$, for a reciprocal system. For loss-less systems it is unitary, i. e. $\mathbf{S}^{\mathrm{T}} = \mathbf{S}^{-1}$.

As shown in [75, p. 45] the quantities $\mathbf{a}$ and $\mathbf{b}$ can be chosen as input and output variables for FIT systems leading to the formulations shown therein. These formulations could be used for the subsequent MOR techniques. However, better numerical properties of the resulting system matrices are obtained by defining the input and output variables as explained in the next chapter by means of the current density vector $\widehat{\widehat{\mathbf{j}}}_s$ and the electric voltage $\bar{\mathbf{e}}$ at the ports and then calculating the scattering parameters from the transfer function (which is the same as the impedance). The relation between the standardized impedance $\bar{\mathbf{H}}$ and the scattering matrix $\mathbf{S}$ is given by:

$$\mathbf{S} = (\bar{\mathbf{H}} - \mathbf{I})(\bar{\mathbf{H}} + \mathbf{I})^{-1}. \tag{3.2.26}$$

How the errors occurring naturally to the impedance calculation by means of the FIT are propagated to the S-Parameters is investigated in [75, p. 44] for the single-port case.

3.2.6 Stability

In principal, stability means that small inputs produce small outputs, where the definition of "small" depends on the norm. Various norms are defined for LTI systems, for example summarized in [1].

A distinction is made between systems without external influence and systems with excitation. A system without external influence is called stable if all elements of its behavior $\mathfrak{B}$ are bounded, i. e. $|\mathbf{w}| \leq M$, $\mathbf{w} \in \mathfrak{B}$, for $t > 0$. It is called asymptotically stable if all elements of $\mathfrak{B}$ approach to zero, i. e. $\mathbf{w} \in \mathfrak{B} \to 0$, for $t \to 0$. If one or more elements do not converge, but are bounded for all time-instances, it is called marginally stable.

For systems with external influence, i. e. systems with in- and output variables as defined by equation (3.1.6), a distinction is made between internal stability and bounded-input, bounded-output (BIBO) stability. A system is internally stable, if the corresponding system resulting for $\mathbf{i} = 0$ is asymptotically stable. It is internally marginally stable, if the corresponding system for $\mathbf{i} = 0$ is marginally stable.

A system is called BIBO stable if any bounded input $\mathbf{i}$ results in a bounded output $\mathbf{u}$:

$$\sup_{t\geq 0}\|\mathbf{i}(t)\| < \infty \Rightarrow \sup_{t\geq 0}\|\mathbf{u}(t)\| < \infty, \quad (\mathbf{i}, \mathbf{u}) \in \mathfrak{B}. \tag{3.2.27}$$

Let the description of $\mathbf{\Sigma}$ with the convolution integral be $\mathbf{u} = \int_{-\infty}^{\infty} \mathbf{h}(t-\tau)\mathbf{i}(\tau)\mathrm{d}\tau$, where $\mathbf{h}$ is the impulse response. Then, the system is BIBO stable if and only if [1, p. 151]:

$$\int_0^{\infty} |\mathbf{h}(t)|\mathrm{d}t < \infty. \tag{3.2.28}$$

Poles and eigenvalues of a system are strongly related to the stability of the system. In the following, stability conditions are given:

1. A system is internally asymptotically stable, if all eigenvalues of $-\mathbf{A}$ have a negative real part.
2. A system is internally marginally stable, if all eigenvalues of $-\mathbf{A}$ have a non-positive real part. The eigenvalues on the imaginary axis have to be simple roots of the characteristic polynomial $\det(s\mathbf{I} + \mathbf{A})$.
3. A system is BIBO stable, if all eigenvalues of $-\mathbf{A}$, that is the roots of the numerator, which are not canceling with the roots of the denominator, have a negative real part. If no roots are canceling, the system is asymptotically stable.

3.2.7 Passivity

While stability is a natural property of physical models, in some cases a stronger concept is required. For instance, source-free electronic structures that are stable

can become instable if non-linear components are connected to them. Therefore, the concept of passivity is introduced. In particular, while at the composition of stable elements, the stability of the overall system is not guaranteed in general, the composition of passive elements always leads to passive overall systems.

A system is called passive, if it does not generate energy. As described in [1, p. 163], the passivity concept is related to the so-called positive realness of a system. A rational function $\mathbf{H}(s)$ is called positive real if:

$$\mathrm{Re}\{s\} \geq 0 \Rightarrow \mathrm{Re}\{\mathbf{H}(s)\} \geq 0, \quad s \in \mathbb{C}, s \text{ not a pole.} \tag{3.2.29}$$

Conditions for the positive realness of $\mathbf{H}$ can be found in [1, p. 164 f.]. As described already in [63] and more generally in [75, p. 41], due to the relation $\widetilde{\mathbf{C}} = \mathbf{C}^{\mathrm{T}}$ and the positive definite material matrices, systems resulting from the FIT are usually positive real and therefore passive. This property is not taken for granted, as for many other discretization methods it is not given.

It is a main goal for MOR techniques to preserve stability and passivity in the reduced order model.

Chapter 4

The FIT System

This chapter is devoted to the various system formulations that result from the Maxwell grid equations. After specifying the in- and output variables of the FIT system, the derivation of the state-space and other relevant formulations from the Maxwell grid equations is provided. Along with each system representation, properties of the respective matrices are given. Special attention is paid to parametric systems, where the focus is set to both material and geometry parameters. A crucial difference between these parameter types exists. In particular, while the matrices obtained by the FIT exhibit an explicit linear dependence on the material parameters, this is not the case for geometry parameters, which show up an implicit dependence. However, in order to obtain the multivariate polynomial form introduced in section 1 which is required for the subsequent MOR, an explicit dependence also on the geometry parameters is necessary. To this purpose, a method is developed for obtaining a -nonlinear- explicit dependence on the geometry parameters.

4.1 Input and Output Variables

Starting point for all formulations is the FIT-system $\mathbf{\Sigma}_{\mathrm{FIT}}$, (2.2.42), which is given here in the frequency domain:

$$\mathbf{\Sigma}_{\mathrm{FIT}} : \begin{cases} \mathbf{C}_{\mathrm{FIT}}\widehat{\mathbf{e}} &= -s\mathbf{M}_{\mu^{-1}}^{-1}\widehat{\mathbf{h}} \\ \widetilde{\mathbf{C}}_{\mathrm{FIT}}\widehat{\mathbf{h}} &= s\mathbf{M}_{\varepsilon}\widehat{\mathbf{e}} + \mathbf{M}_{\kappa}\widehat{\mathbf{e}} + \widehat{\widehat{\mathbf{j}}}_{e}. \end{cases} \tag{4.1.1}$$

The structure described by this system interacts with its environment through the ports that let energy enter and leave the structure. The variables describing this interaction in case of waveguide ports or discrete ports are the current density vector $\widehat{\widehat{\mathbf{j}}}_{e}$ and the voltage $\widehat{\mathbf{e}}$ across the external terminal. From physical considerations it follows that $\widehat{\widehat{\mathbf{j}}}_{e}$ at the ports can be set to any time function and thus be considered as a free variable. Furthermore, given $\widehat{\widehat{\mathbf{j}}}_{e}$, i. e. $\mathbf{i}$, the variable $\mathbf{u}$ is determined by the input.

The generalized current is associated by $\widehat{\widehat{\mathbf{j}}}_{e} = \mathbf{R}\mathbf{i}$ with the current density vector $\widehat{\widehat{\mathbf{j}}}_{e}$ at the excitation ports, as described in 2.2.57. The generalized voltage is $\mathbf{u} =$

$\mathbf{L}\widehat{\mathbf{e}}$. These are the external variables, denoted by $\mathbf{w} = (\mathbf{i}, \mathbf{u})$, while the remaining variables in (3.1.1) are the auxiliary variables. According to section 3.1.1 in the previous chapter, $\mathbf{w}$ can be partitioned into in- and output variables. Thus, from an intuitive point of view, $\mathbf{\Sigma}_{\mathrm{FIT}}$ can be considered as an input/output system with $\mathbf{i}$ as input and $\mathbf{u}$ as output. Furthermore, the auxiliary variables are state variables in case of a state-space model.

In general, it is possible to define the input ports to be different from the output ports, with dimension l and l' respectively, but in practice this is rather unusual, hence, $l = l'$.

4.2 Single-Parameter Systems

4.2.1 The State-Space Formulation

In the following, the state-space formulation of $\mathbf{\Sigma}_{\mathrm{FIT}}$ will be derived. In [41] a general procedure is proposed, nevertheless here an intuitive approach is chosen. With the in- and output variables defined in the previous section, the auxiliary variable $\mathbf{x} = (\widehat{\mathbf{e}}, \widehat{\mathbf{h}})^{\mathrm{T}}$ turns out to form also the state vector. From (4.1.1) it directly follows:

$$\begin{pmatrix} \mathbf{M}_{\varepsilon} & 0 \\ 0 & \mathbf{M}_{\mu^{-1}}^{-1} \end{pmatrix} s \begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix} = - \begin{pmatrix} \mathbf{M}_{\kappa} & -\widetilde{\mathbf{C}}_{\mathrm{FIT}} \\ \mathbf{C}_{\mathrm{FIT}} & 0 \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix} - \begin{pmatrix} \widehat{\widehat{\mathbf{j}}}_e \\ 0 \end{pmatrix}. \quad (4.2.2)$$

With the inversion of the material matrices $\mathbf{M}_{\varepsilon}, \mathbf{M}_{\mu}$, which is trivial as they are diagonal, the system equals to:

$$s \begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix} = - \begin{pmatrix} \mathbf{M}_{\varepsilon} & 0 \\ 0 & \mathbf{M}_{\mu^{-1}}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{M}_{\kappa} & -\widetilde{\mathbf{C}}_{\mathrm{FIT}} \\ \mathbf{C}_{\mathrm{FIT}} & 0 \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix} - \begin{pmatrix} \mathbf{M}_{\varepsilon} & 0 \\ 0 & \mathbf{M}_{\mu^{-1}}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\widehat{\mathbf{j}}}_e \\ 0 \end{pmatrix}. \quad (4.2.3)$$

Applying the relation $\widehat{\widehat{\mathbf{j}}}_e = \mathbf{R}\mathbf{i}$ and simplifying the notation of the inverse matrices results in:

$$s \underbrace{\begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix}}_{\mathbf{x}} = \underbrace{- \begin{pmatrix} \mathbf{M}_{\varepsilon}^{-1}\mathbf{M}_{\kappa} & -\mathbf{M}_{\varepsilon}^{-1}\widetilde{\mathbf{C}}_{\mathrm{FIT}} \\ \mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}} & 0 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix}}_{\mathbf{x}} + \underbrace{\begin{pmatrix} \mathbf{M}_{\varepsilon}^{-1}\mathbf{R} \\ 0 \end{pmatrix}}_{\mathbf{B}} \mathbf{i}, \quad (4.2.4)$$

while for the output variable $\mathbf{u} = \mathbf{L}\widehat{\mathbf{e}}$ holds:

$$\mathbf{u} = \underbrace{\begin{pmatrix} \mathbf{L} & \mathbf{0} \end{pmatrix}}_{\mathbf{C}} \begin{pmatrix} \widehat{\mathbf{e}} \\ \widehat{\mathbf{h}} \end{pmatrix}. \quad (4.2.5)$$

In the loss-free case ($\mathbf{M}_\kappa = \mathbf{0}$), the system matrix of (4.2.4) can be transformed in a skew-symmetric matrix. To this purpose, new variables

$$\widehat{\mathbf{e}}' = \mathbf{M}_\varepsilon^{1/2}\widehat{\mathbf{e}} \quad \text{and} \quad \widehat{\mathbf{h}}' = \mathbf{M}_\mu^{-1/2}\widehat{\mathbf{h}}, \quad \text{where} \quad \mathbf{M}_\varepsilon^{1/2}\mathbf{M}_\varepsilon^{1/2} = \mathbf{M}_\varepsilon, \mathbf{M}_\mu^{-1/2}\mathbf{M}_\mu^{-1/2} = \mathbf{M}_{\mu^{-1}}, \tag{4.2.6}$$

are defined. As $\mathbf{C}_{\mathrm{FIT}} = \widetilde{\mathbf{C}}_{\mathrm{FIT}}^{\mathrm{T}}$, the resulting system matrix is skew-symmetric. For the in- and output variables it is: $\mathbf{M}_\varepsilon^{1/2}\widehat{\mathbf{j}}_e = \mathbf{M}_\varepsilon^{1/2}\mathbf{R}\mathbf{i} = \mathbf{R}'\mathbf{i}$ and analogously $\mathbf{L}' = \mathbf{L}\mathbf{M}_\varepsilon^{-1/2}$. The symmetrized system reads:

$$s \underbrace{\begin{pmatrix} \widehat{\mathbf{e}}' \\ \widehat{\mathbf{h}}' \end{pmatrix}}_{\mathbf{x}'} = \underbrace{-\begin{pmatrix} \mathbf{M}_\varepsilon^{-1/2}\mathbf{M}_\kappa\mathbf{M}_\varepsilon^{-1/2} & -\mathbf{M}_\varepsilon^{-1/2}\mathbf{C}_{\mathrm{FIT}}\mathbf{M}_\mu^{-1/2} \\ \mathbf{M}_\mu^{-1/2}\mathbf{C}_{\mathrm{FIT}}\mathbf{M}_\varepsilon^{-1/2} & \mathbf{0} \end{pmatrix}}_{\mathbf{A}'} \underbrace{\begin{pmatrix} \widehat{\mathbf{e}}' \\ \widehat{\mathbf{h}}' \end{pmatrix}}_{\mathbf{x}'} + \underbrace{\begin{pmatrix} \mathbf{R}' \\ 0 \end{pmatrix}}_{\mathbf{B}'} \mathbf{i}. \tag{4.2.7}$$

In both cases, the resulting system is a first-order system in the classical state-space form, here exemplary given for the non-symmetrized form:

$$\mathbf{\Sigma}_{\mathrm{FIT,linear}} \begin{cases} s\mathbf{x} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{i}, \\ \mathbf{u} &= \mathbf{C}\mathbf{x}. \end{cases} \tag{4.2.8}$$

The system can be shifted about a frequency s_0, yielding:

$$\mathbf{\Sigma}_{\mathrm{FIT,linear,shift}} \begin{cases} (s - s_0)\mathbf{x} &= (\mathbf{A} - s_0\mathbf{I})\mathbf{x} + \mathbf{B}\mathbf{i}, \\ \mathbf{u} &= \mathbf{C}\mathbf{x}. \end{cases} \tag{4.2.9}$$

For the systems considered here, $\mathbf{D} = \mathbf{0}$, i. e. no direct coupling between in- and output exists.

4.2.1.1 Transfer function and moments of the state-space system

The transfer function $\mathbf{H}$, with $\mathbf{u} = \mathbf{H}\mathbf{i}$, is:

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}. \tag{4.2.10}$$

For the shifted system (4.2.9) the transfer function reads:

$$\mathbf{H}(s - s_0) = \mathbf{C}(\mathbf{I} + (s - s_0)\bar{\mathbf{A}})^{-1}\bar{\mathbf{B}}, \tag{4.2.11}$$

with

$$\bar{\mathbf{A}} = (s_0\mathbf{I} - \mathbf{A})^{-1} \quad \text{and} \quad \bar{\mathbf{B}} = \bar{\mathbf{A}}\mathbf{B}. \tag{4.2.12}$$

The Laurent series expansion of the transfer function $\mathbf{H}(s)$ in the neighborhood of s_0 reads:

$$\mathbf{H}(s-s_0) = \sum_{k=0}^{\infty} \underbrace{\frac{1}{k!}\frac{\mathrm{d}\mathbf{H}(s)}{\mathrm{d}s^k}\bigg|_{s=s_0}}_{\boldsymbol{\eta}_k} (s-s_0)^k \tag{4.2.13}$$

$$= \boldsymbol{\eta}(s_0) + \boldsymbol{\eta}_1(s_0)\frac{(s-s_0)}{1!} + \cdots + \boldsymbol{\eta}_k(s_0)\frac{(s-s_0)^k}{k!} + \ldots. \tag{4.2.14}$$

The moments (provided that $\mathbf{D} = \mathbf{0}$) around $s_0 = 0$ are:

$$\boldsymbol{\eta}_k(s_0 = 0) = \mathbf{C}\mathbf{A}^{-(k+1)}\mathbf{B}, \quad k \geq 0, \tag{4.2.15}$$

and those at s_0:

$$\boldsymbol{\eta}_k(s_0) = \mathbf{C}(s_0\mathbf{I} - \mathbf{A})^{-(k+1)}\mathbf{B}, \quad k \geq 0. \tag{4.2.16}$$

4.2.1.2 Reduced state-space system

The reduced systems $\hat{\boldsymbol{\Sigma}}_{\mathrm{FIT,linear}}$ and $\hat{\boldsymbol{\Sigma}}_{\mathrm{FIT,linear,shift}}$ are given by:

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{FIT,linear}} \left\{ \begin{array}{rl} s\hat{\mathbf{x}} & = \hat{\mathbf{A}}\mathbf{x} + \hat{\mathbf{B}}\mathbf{i}, \\ \mathbf{u} & = \hat{\mathbf{C}}\hat{\mathbf{x}}. \end{array} \right. \quad \text{and} \tag{4.2.17}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{FIT,linear,shift}} \left\{ \begin{array}{rl} (s-s_0)\hat{\mathbf{x}} & = (\hat{\mathbf{A}} - s_0\mathbf{I})\hat{\mathbf{x}} + \hat{\mathbf{B}}\mathbf{i}, \\ \mathbf{u} & = \hat{\mathbf{C}}\hat{\mathbf{x}}. \end{array} \right. \tag{4.2.18}$$

Analogously to (4.2.13), the moments of the reduced systems are given by:

$$\hat{\boldsymbol{\eta}}_k(s_0) = \hat{\mathbf{C}}(s_0\mathbf{I} - \hat{\mathbf{A}})^{-(k+1)}\hat{\mathbf{B}}, \quad k \geq 0. \tag{4.2.19}$$

4.2.2 The Curl-Curl Formulation

The discrete curl-curl equation (2.2.46), where the auxiliary variable $\widehat{\mathbf{h}}$ has been eliminated, reads:

$$\mathbf{M}_\varepsilon s^2\, \widehat{\mathbf{e}} + \mathbf{M}_\kappa s\, \widehat{\mathbf{e}} + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\widehat{\mathbf{e}} \;=\; s\, \widehat{\widehat{\mathbf{j}}}_s. \tag{4.2.20}$$

Again, the in- and output variables as defined above, with $\widehat{\widehat{\mathbf{j}}}_s = \mathbf{B}\mathbf{i}$ and $\mathbf{u} = \mathbf{C}\widehat{\mathbf{e}}$, respectively, are used, resulting in the system:

$$\boldsymbol{\Sigma}_{\mathrm{FIT,curl}} : \left\{ \begin{array}{rl} \mathbf{M}_\varepsilon s^2\widehat{\mathbf{e}} + \mathbf{M}_\kappa s\widehat{\mathbf{e}} + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\widehat{\mathbf{e}} & = s\mathbf{B}\mathbf{i}, \\ \mathbf{u} & = \mathbf{C}\widehat{\mathbf{e}}, \end{array} \right. \tag{4.2.21}$$

which is a second-order linear system in $\widehat{\mathbf{e}}$, as described in (1.2.2). The vector $\widehat{\mathbf{e}}$ plays the role of the auxiliary vector $\mathbf{x}$. The number of unknowns is about half of those in (4.2.8).

Equation (4.2.21) can again be symmetrized. With $\mathbf{C}'^{\mathrm{T}} = \mathbf{B}'^{\mathrm{T}} = \mathbf{M}_\varepsilon^{-1/2}\mathbf{B} = \mathbf{R}'$ and the auxiliary vector $\mathbf{x}' = \mathbf{M}_\varepsilon^{1/2}\widehat{\mathbf{e}}$ it is:

$$\boldsymbol{\Sigma}_{\mathrm{FIT,curl,symm}} : \begin{cases} (s^2\mathbf{I} + \mathbf{M}_\varepsilon^{-1}\mathbf{M}_\kappa s + \mathbf{M}_\varepsilon^{-1/2}\mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\mathbf{M}_\varepsilon^{-1/2})\mathbf{x}' & = s\mathbf{B}'\mathbf{i}, \\ \mathbf{u} & = \mathbf{C}'\mathbf{x}'. \end{cases} \tag{4.2.22}$$

With

$$(\mathbf{F}, \mathbf{K}, \mathbf{A}_{\mathrm{CC}}, \mathbf{B}, \mathbf{C}, \mathbf{x}) = \begin{cases} (\mathbf{M}_\varepsilon, \mathbf{M}_\kappa, \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}, \mathbf{B}, \mathbf{C}, \widehat{\mathbf{e}}) & , \text{unsymmetric} \\ (\mathbf{I}, \mathbf{M}_\varepsilon^{-1}\mathbf{M}_\kappa, \mathbf{M}_\varepsilon^{-1/2}\mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\mathbf{M}_\varepsilon^{-1/2}, \mathbf{B}', \mathbf{C}', \widehat{\mathbf{e}}') & , \text{symmetric} \end{cases} \tag{4.2.23}$$

a general formulation for both systems can be given:

$$\boldsymbol{\Sigma}_{\mathrm{FIT,curl,general}} : \begin{cases} \mathbf{F}s^2\mathbf{x} + \mathbf{K}s\mathbf{x} + \mathbf{A}_{\mathrm{CC}}\mathbf{x} & = s\mathbf{B}\mathbf{i}, \\ \mathbf{u} & = \mathbf{C}\mathbf{x}. \end{cases} \tag{4.2.24}$$

The matrix $\mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}$ (and its symmetrized version) is singular, as $\mathbf{C}_{\mathrm{FIT}}$ is singular. Therefore, in order to make use of its moments, it is necessary to shift the expression about a frequency s_0. In such cases, a variable transformation of s is accomplished in (4.2.21):

$$\left(\mathbf{M}_\varepsilon(s-s_0)^2 + 2\mathbf{M}_\varepsilon s s_0 - \mathbf{M}_\varepsilon s_0^2 + \mathbf{M}_\kappa s + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\right)\widehat{\mathbf{e}} = (s-s_0)\,\mathbf{B}\mathbf{i} + s_0\mathbf{B}\mathbf{i}. \tag{4.2.25}$$

The term $2\mathbf{M}_\varepsilon s s_0 - \mathbf{M}_\varepsilon s_0^2$, can be written as follows:

$$\begin{aligned} 2\mathbf{M}_\varepsilon s s_0 - \mathbf{M}_\varepsilon s_0^2 &= 2\mathbf{M}_\varepsilon s s_0 - 2\mathbf{M}_\varepsilon s_0 s_0 + 2\mathbf{M}_\varepsilon s_0 s_0 - \mathbf{M}_\varepsilon s_0^2 \\ &= 2\mathbf{M}_\varepsilon s_0(s-s_0) + \mathbf{M}_\varepsilon s_0^2, \end{aligned} \tag{4.2.26}$$

and equation (4.2.20) results in:

$$\left(\mathbf{M}_\varepsilon(s-s_0)^2 + (2\mathbf{M}_\varepsilon s_0 + \mathbf{M}_\kappa)(s-s_0) + \mathbf{M}_\varepsilon s_0^2 + \mathbf{M}_\kappa s_0 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\right)\widehat{\mathbf{e}} = (s-s_0)\,\mathbf{B}\mathbf{i} + s_0\mathbf{B}\mathbf{i}. \tag{4.2.27}$$

If no losses are considered ($\mathbf{M}_\kappa = \mathbf{0}$), the loss-free, shifted, non-symmetrized curl-curl system is given by:

$$\boldsymbol{\Sigma}_{\mathrm{FIT,curl,shift}} : \begin{cases} \left(\mathbf{M}_\varepsilon(s-s_0)^2 + (2\mathbf{M}_\varepsilon s_0)(s-s_0) + \mathbf{M}_\varepsilon s_0^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\right)\mathbf{x} & = s\mathbf{B}\mathbf{i}, \\ \mathbf{u} & = \mathbf{C}\mathbf{x}. \end{cases} \tag{4.2.28}$$

4.2.2.1 Transfer function and moments of the curl-curl system

The transfer function of the non-symmetrized loss-free curl-curl system is:

$$\mathbf{H}(s) = \mathbf{C}(\mathbf{M}_\varepsilon s^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}})^{-1}\mathbf{B}. \tag{4.2.29}$$

while for the shifted version holds:

$$\mathbf{H}(s-s_0) = \mathbf{C}\left(\mathbf{M}_\varepsilon(s-s_0)^2 + 2\mathbf{M}_\varepsilon s s_0 - \mathbf{M}_\varepsilon s_0^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}}\right)^{-1}\mathbf{B}. \tag{4.2.30}$$

The moments are again the Laurent series coefficients defined in (4.2.13) and read:

$$\boldsymbol{\eta}_k(s_0) = \mathbf{C}(\mathbf{M}_\varepsilon(s-s_0)^2 + 2\mathbf{M}_\varepsilon s s_0 - \mathbf{M}_\varepsilon s_0^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}\mathbf{C}_{\mathrm{FIT}})^{-(k+1)}\mathbf{B}, \quad k \geq 0. \tag{4.2.31}$$

4.2.2.2 Reduced curl-curl system

A reduced curl-curl system of (4.2.28) is defined by:

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{FIT,curl,shift}} : \begin{cases} \left(\hat{\mathbf{M}}_\varepsilon(s-s_0)^2 + 2\hat{\mathbf{M}}_\varepsilon s_0(s-s_0) + \hat{\mathbf{M}}_\varepsilon s_0^2 + \hat{\mathbf{A}}_{\mathrm{CC}}\right)\hat{\mathbf{x}} & = s\hat{\mathbf{B}}\mathbf{i}, \\ \mathbf{u} & = \hat{\mathbf{C}}\hat{\mathbf{x}}, \end{cases} \tag{4.2.32}$$

and its transfer function and moments are defined analogously to (4.2.30) and (4.2.31), respectively.

4.2.3 Higher-Order Systems

In general, due to atomic and molecular interactions caused by external electric fields, the material parameters are frequency-dependent. This has already mentioned in section (2.1.2). The Maxwell equations applied for these cases result in higher-order systems, as introduced in (1.2.3):

$$\boldsymbol{\Sigma}_{\mathrm{higher-order}} : \begin{cases} \sum_{k=0}^{\mathrm{N_A}} \mathbf{A}_k\, s^k\mathbf{x} & = \sum_{k=0}^{\mathrm{N_B}} \mathbf{B}_k\, s^k\mathbf{i}, \\ \mathbf{u} & = \sum_{k=0}^{\mathrm{N_B}} \mathbf{C}_k\, s^k\mathbf{x}. \end{cases} \tag{4.2.33}$$

The specification of $\mathbf{A}_k, \mathbf{B}_k, N_{\mathrm{A}}$ and N_{B} depends on the material properties.

In [75, p. 32 f.] two examples of FIT systems in this form are derived, namely systems resulting from structures with dispersive dielectrica and systems resulting from structures with PML-boundary conditions.

The transfer function of $\boldsymbol{\Sigma}_{\mathrm{higher-order}}$ is:

$$\mathbf{H}(s) = \sum_{k=0}^{\mathrm{N_B}} \mathbf{C}_k\, s^k \left(\sum_{k=0}^{\mathrm{N_A}} \mathbf{A}_k s^k\right)^{-1} \sum_{k=0}^{\mathrm{N_B}} \mathbf{B}_k s^k. \tag{4.2.34}$$

A reduced system of (4.2.33) reads:

$$\hat{\boldsymbol{\Sigma}}_{\text{higher-order}} : \begin{cases} \sum_{k=0}^{\mathrm{N_A}} \hat{\mathbf{A}}_k \, s^k \hat{\mathbf{x}} &= \sum_{k=0}^{\mathrm{N_B}} \hat{\mathbf{B}}_k \, s^k \mathbf{i}, \\ \mathbf{u} &= \sum_{k=0}^{\mathrm{N_B}} \hat{\mathbf{C}}_k \, s^k \hat{\mathbf{x}}. \end{cases} \tag{4.2.35}$$

4.3 Multi-Parameter Systems

In the curl-curl equation (4.2.21) only the frequency parameter s appears explicitly. Other parameters, e. g. the material parameters ε, μ and κ, or geometry parameters are implicitly included in the material matrices $\mathbf{M}_{\varepsilon}(\boldsymbol{\xi}), \mathbf{M}_{\mu^{-1}}(\boldsymbol{\xi})$ and $\mathbf{M}_{\kappa}(\boldsymbol{\xi})$ in a nonlinear dependence. Thus, the general formulation of parametric systems, resulting from the non-symmetric curl-curl system is:

$$\boldsymbol{\Sigma}_{\text{FIT,curl}} : \begin{cases} \mathbf{M}_{\varepsilon}(\boldsymbol{\xi}) s^2 \widetilde{\mathbf{e}} + \mathbf{M}_{\kappa}(\boldsymbol{\xi}) s \widetilde{\mathbf{e}} + \mathbf{C}_{\text{FIT}}^{\mathrm{T}} \mathbf{M}_{\mu^{-1}}(\boldsymbol{\xi}) \mathbf{C}_{\text{FIT}} \widetilde{\mathbf{e}} &= s\mathbf{R}\mathbf{i}, \\ \mathbf{u} &= \mathbf{L}\widetilde{\mathbf{e}}. \end{cases} \tag{4.3.36}$$

The in- and output matrices are considered parameter-independent. Analogously, the parameter dependence can be considered in the linear and the higher-order systems, both in their normal as well as in their shifted versions.

Some order reduction techniques, however, as will be described in chapter 6, require an explicit dependence on the parameters. In the following, it is therefore aimed to find a formulation with explicit parameter specification by linearizing (4.3.36).

4.3.1 Linearization of the Curl-Curl Equation

In the following, the method is restricted to frequency and normalized geometry variation, s and $\boldsymbol{\nu} = (\nu_x, \nu_y, \nu_z)$, respectively. The mesh is considered as being stretched rectilinearly according to the geometry variation. In this way, the mesh topology is maintained. Let $\mathcal{G}_1$ be the original mesh and $\mathcal{G}_{2x,y,z}$ the maximally modified mesh in $x-$, $y-$ and $z-$ direction, respectively, then the following equation

$$\mathcal{G}_u = \mathcal{G}_1 + (\mathcal{G}_{2u} - \mathcal{G}_1)\nu_u \quad \text{with} \quad \nu_u \in [0,1], u = x, y, z, \tag{4.3.37}$$

describes pictorially how the points, edges and surfaces of the modified mesh $\mathcal{G}_u$ in each direction result from the respective points, edges and surfaces of $\mathcal{G}_1$ and $\mathcal{G}_{2x,y,z}$. It is assumed that the variation in each direction is not larger than the respective edge lengths of the original mesh. Figure 4.1 shows the original and the stretched mesh in horizontal direction for a simple test model.

The mesh modifications result in a variation of the geometry dependent matrices $\mathbf{D}_{\mathrm{S}}, \mathbf{D}_{\mathrm{A}}, \tilde{\mathbf{D}}_{\mathrm{S}}, \tilde{\mathbf{D}}_{\mathrm{A}}$ which contribute to the material matrices $\mathbf{M}_{\varepsilon}, \mathbf{M}_{\mu}$ and $\mathbf{M}_{\kappa}$, given in (2.2.41). In the following, problems without losses are considered. In these cases

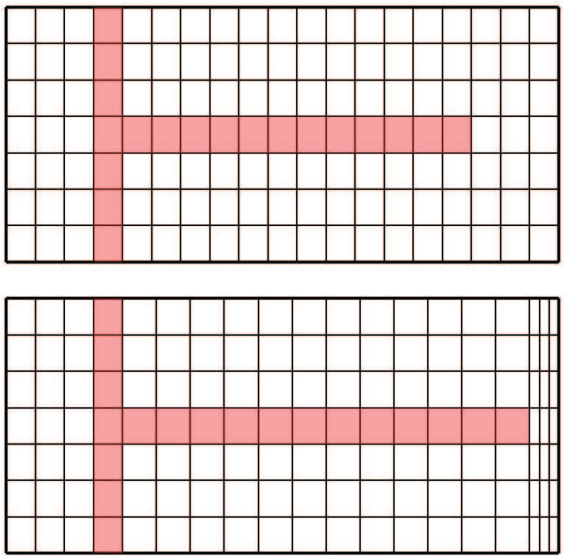


Figure 4.1: *Geometry variation based on a rectilinear stretched mesh with a fixed mesh topology.*

$\mathbf{M}_\kappa = \mathbf{0}$. Special care has to be taken to set up the inverted matrices $\mathbf{D}_{\mathrm{S}}^{-1}$ and $\mathbf{D}_{\mathbf{A}}^{-1}$. Concentrating on the diagonal matrix of edge related contributions, it is:

$$\mathbf{D}_{\mathrm{S}} = \underbrace{\mathbf{D}_{\mathrm{S}_1}}_{\mathbf{M}_0} + \sum_{u=x,y,z} \underbrace{(\mathbf{D}_{\mathrm{S}_{2u}} - \mathbf{D}_{\mathrm{S}_1})}_{\mathbf{M}_u} \nu_u, \quad u = x, y, z. \tag{4.3.38}$$

It becomes clear at this point, that due to

$$\mathbf{D}_{\mathrm{S}}^{1/2} = \mathbf{D}_{\mathrm{S}_1}^{1/2} + \left(\sum_{u=x,y,z} (\mathbf{D}_{\mathrm{S}_{2u}} - \mathbf{D}_{\mathrm{S}_1}) \right)^{1/2} \neq \mathbf{D}_{\mathrm{S}_1}^{1/2} + \sum_{u=x,y,z} (\mathbf{D}_{\mathrm{S}_{2u}} - \mathbf{D}_{\mathrm{S}_1})^{1/2} \tag{4.3.39}$$

the symmetric curl-curl system cannot be made use of and therefore the non-symmetric curl-curl system (4.2.21) has to be used for the linearization.

Each element $\mathbf{D}_{\mathrm{S}}(ii), i = 1 \ldots n$ can be explicitly stated as:

$$\mathbf{D}_{\mathrm{S}}^{-1}(ii) = \frac{1}{\mathbf{M}_0(ii) + \mathbf{M}_x(ii)\nu_x + \mathbf{M}_y(ii)\nu_y + \mathbf{M}_z(ii)\nu_z}. \tag{4.3.40}$$

The elements for the inverse matrix $\mathbf{D}_S^{-1}$ are given by:

$$\mathbf{D}_{\mathrm{S}}^{-1}(ii) = \frac{1}{\mathbf{M}_0(ii)} \frac{1}{1 - \left(\sum_{u=x,y,z} -\frac{\mathbf{M}_u(ii)}{\mathbf{M}_0(ii)} \nu_u \right)}, \quad u = x, y, z. \tag{4.3.41}$$

With the help of the geometric series

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \ |q| \leq 1, \tag{4.3.42}$$

(4.3.41) can be written in the form:

$$\mathbf{D}_{\mathrm{S}}^{-1}(ii) = \frac{1}{\mathbf{M}_0(ii)} \sum_{k=0}^{\infty} \left(\sum_{u=x,y,z} -\frac{\mathbf{M}_u(ii)}{\mathbf{M}_0(ii)} \nu_u \right)^k . \tag{4.3.43}$$

Naturally, the condition

$$\left| \left(-\frac{\mathbf{M}_x(ii)}{\mathbf{M}_0(ii)} \nu_x - \frac{\mathbf{M}_y(ii)}{\mathbf{M}_0(ii)} \nu_y - \frac{\mathbf{M}_z(ii)}{\mathbf{M}_0(ii)} \nu_z \right) \right| \leq 1 \tag{4.3.44}$$

has to be satisfied. Notice that for each fraction holds:

$$\left| \frac{\mathbf{M}_{x,y,z}(ii)}{\mathbf{M}_0(ii)} \right| \leq 1 \quad \text{and} \quad \nu_{x,y,z} \leq 1. \tag{4.3.45}$$

Due to the diagonality feature of the matrices $\mathbf{M}_0$ and $\mathbf{M}_x, \mathbf{M}_y, \mathbf{M}_z$, it is:

$$\mathbf{D}_{\mathrm{S}}^{-1} = \mathbf{M}_0^{-1} \sum_{k=0}^{\infty} \left(\sum_{u=x,y,z} -\mathbf{M}_u \mathbf{M}_0^{-1} \nu_u \right)^k , \; u = x, y, z. \tag{4.3.46}$$

Analogously, the remaining matrices can be formulated following the same idea in the form:

$$\tilde{\mathbf{D}}_{\mathrm{S}} = \mathbf{N}_0 + \mathbf{N}_x \nu_x + \mathbf{N}_y \nu_y + \mathbf{N}_z \nu_z, \tag{4.3.47a}$$

$$\mathbf{D}_A = \mathbf{K}_0 + \mathbf{K}_x \nu_x + \mathbf{K}_y \nu_y + \mathbf{K}_z \nu_z, \tag{4.3.47b}$$

$$\tilde{\mathbf{D}}_A = \mathbf{L}_0 + \mathbf{L}_x \nu_x + \mathbf{L}_y \nu_y + \mathbf{L}_z \nu_z. \tag{4.3.47c}$$

Substitution of the material matrices with $\mathbf{D}_{\mathrm{S}}, \tilde{\mathbf{D}}_{\mathrm{S}}, \mathbf{D}_{\mathrm{A}}, \tilde{\mathbf{D}}_{\mathrm{A}}$ from (4.3.46) and (4.3.47) in (4.2.28),(no losses are considered), leads to an expression with explicit dependence on the parameters s, ν_x, ν_y and ν_z:

$$\begin{aligned} &\left(\mathbf{L}_0 + \sum_{u=x,y,z} \mathbf{L}_u \nu_u \right) \mathbf{D}_\varepsilon \, \mathbf{M}_0^{-1} \sum_{k=0}^{\infty} \left(\sum_{u=x,y,z} -\mathbf{M}_u \mathbf{M}_0^{-1} \nu_u \right)^k \\ &\qquad \left(\mathbf{I} \, (s - s_0)^2 + \mathbf{I} \, s_0 (s - s_0) + \mathbf{I} \, s_0^2 \right) \widehat{\mathbf{e}} + \\ &+ \left(\mathbf{N}_0 + \sum_{u=x,y,z} \mathbf{N}_u \nu_u \right) \mathbf{D}_{\mu^{-1}} \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}} \, \mathbf{L}_0^{-1} \sum_{k=0}^{\infty} \left(\sum_{u=x,y,z} -\mathbf{L}_u \mathbf{L}_0^{-1} \nu_u \right)^k \mathbf{C}_{\mathrm{FIT}} \widehat{\mathbf{e}}. \end{aligned} \tag{4.3.48}$$

Notice, that the system matrix of the resulting system is a polynomial of time-invariant matrices. The general form of this kind of problems is given in section 4.3.3.

The error caused by this linearization, i. e. the relative error of truncating the series expressing $\mathbf{D}_S^{-1}$ and $\mathbf{D}_A^{-1}$ after k terms, has been analyzed in [55]. The results therein and further investigations showed that a value $k = 2 \ldots 4$ is a good trade-off between accuracy and system complexity.

4.3.2 Material Parameter Variations

In the previous subsection attention was paid to geometry variations. If material parameters play a role, the diagonal matrices $\mathbf{D}_{\varepsilon}, \mathbf{D}_{\mu^{-1}}$ and $\mathbf{D}_{\kappa}$ are also affected. Whereas $\mathbf{D}_{\varepsilon}$ and $\mathbf{D}_{\kappa}$ depend linearly on the material parameters ε and κ, notice that for $\mathbf{D}_{\mu^{-1}}$, it is its inverse that enters (4.2.21). Let $\mathbf{D}_{\varepsilon_1}$ be the material matrix corresponding to the parameter ε_1 and $\mathbf{D}_{\varepsilon_2}$ the matrix corresponding to the material parameter ε_2. Then, for the matrix corresponding to any ε holds:

$$\mathbf{D}_{\varepsilon} = \mathbf{D}_{\varepsilon,1} + \underbrace{(\mathbf{D}_{\varepsilon,2} - \mathbf{D}_{\varepsilon,1})}_{\Delta\mathbf{D}_{\varepsilon}} \nu_{\varepsilon}, \quad \nu_{\varepsilon} \in [0 \ldots 1]. \tag{4.3.49}$$

Thus, the modified equation (4.2.28) leads to:

$$\begin{aligned}\big((\mathbf{M}_{\varepsilon,1} + \mathbf{\Delta M}_{\varepsilon}\nu_{\varepsilon})(s-s_0)^2 + 2(\mathbf{M}_{\varepsilon,1} + \mathbf{\Delta M}_{\varepsilon}\nu_{\varepsilon})s_0(s-s_0) + \\ (\mathbf{M}_{\varepsilon,1} + \mathbf{\Delta M}_{\varepsilon}\nu_{\varepsilon})s_0^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu}^{-1}\mathbf{C}_{\mathrm{FIT}}\big)\,\widehat{\mathbf{e}} = s\mathbf{Bi},\end{aligned} \tag{4.3.50}$$

with $\mathbf{M}_{\varepsilon,1} = \tilde{\mathbf{D}}_{\mathrm{A}}\mathbf{D}_{\varepsilon,1}\mathbf{D}_{\mathrm{S}}^{-1}$ and $\mathbf{\Delta M}_{\varepsilon} = \tilde{\mathbf{D}}_{\mathrm{A}}\mathbf{\Delta D}_{\varepsilon}\mathbf{D}_{\mathrm{S}}^{-1}$. The matrix on the left hand side of the equation can also be written in the form:

$$\begin{aligned}\mathbf{\Delta M}_{\varepsilon}\nu_{\varepsilon}(s-s_0)^2 + \mathbf{M}_{\varepsilon,1}(s-s_0)^2 + 2\mathbf{\Delta M}_{\varepsilon}s_0\nu_{\varepsilon}(s-s_0) + 2\mathbf{M}_{\varepsilon,1}s_0(s-s_0) + \\ \mathbf{\Delta M}_{\varepsilon}s_0^2\nu_{\varepsilon} + \mathbf{M}_{\varepsilon,1}s_0^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu}^{-1}\mathbf{C}_{\mathrm{FIT}}.\end{aligned} \tag{4.3.51}$$

For the permeability, given $\mathbf{D}_{\mu_1^{-1}}, \mathbf{D}_{\mu_2^{-1}}$ corresponding to the material parameters μ_1 and μ_2, respectively, the matrix $\mathbf{D}_{\mu^{-1}}$ is given by:

$$\mathbf{D}_{\mu^{-1}} = \mathbf{D}_{\mu_1^{-1}} + \underbrace{(\mathbf{D}_{\mu_2^{-1}} - \mathbf{D}_{\mu_1^{-1}})}_{\mathbf{\Delta D}_{\mu^{-1}}} \nu_{\mu}, \quad \nu_{\mu} \in [0 \ldots 1]. \tag{4.3.52}$$

If, additionally to the geometry parameters, the material parameters μ and ε are varied, then (4.3.48) has to be combined with (4.3.49) and (4.3.52), respectively.

4.3.3 Multi-Parameter Polynomial Systems

Equations (4.3.48) and (4.3.51) derived in the previous sections can be generalized with the help of systems which are set up by multivariate matrix polynomials. As more than one parameter is varied, the introduction of the notion of a multi-index is required. A multi-index is an r-tuple $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)$ of non-negative integers $\alpha_1, \ldots, \alpha_r$, i. e. $\boldsymbol{\alpha} \in \mathbb{N}_0^r$. With $\mathbf{s} = (s_1, s_2, \ldots, s_r)^{\mathrm{T}} \in \mathbb{R}^r$ the following definitions

hold [17]:

$$|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_r, \tag{4.3.53a}$$

$$\boldsymbol{\alpha}! = \prod_{k=1}^{r} \alpha_k!, \tag{4.3.53b}$$

$$\mathbf{s}^{\boldsymbol{\alpha}} = \prod_{k=1}^{r} s_k^{\alpha_k}, \tag{4.3.53c}$$

$$D^{\boldsymbol{\alpha}} = \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial s_1^{\alpha_1} \cdots \partial s_r^{\alpha_r}}, \tag{4.3.53d}$$

$$\mathbf{x}_{\boldsymbol{\alpha}} = \mathbf{0}, \quad \text{if } \exists\, i \in \{1, \ldots, r\} : \alpha_i < 0. \tag{4.3.53e}$$

With these definitions, and with $\mathbf{A}_{\boldsymbol{\alpha}} \in \mathbb{R}^{n\times n}, \mathbf{x}, \mathbf{B}_{\boldsymbol{\alpha}}, \mathbf{C}_{\boldsymbol{\alpha}} \in \mathbb{R}^{n\times 1}$, the general form of polynomial systems of degree $|\boldsymbol{\alpha}|_{\max}$ and r scalar variables $s_1, s_2, \ldots, s_r \in \mathbb{C}$ is:

$$\boldsymbol{\Sigma}_{\text{multi-param}} : \begin{cases} \left(\sum\limits_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{A}_{\boldsymbol{\alpha}} \right) \mathbf{x}(\mathbf{s}) &= \left(\sum\limits_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{B}_{\boldsymbol{\alpha}} \right) \mathbf{i}(\mathbf{s}), \\ \mathbf{u}(\mathbf{s}) &= \left(\sum\limits_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{C}_{\boldsymbol{\alpha}} \right) \mathbf{x}(\mathbf{s}). \end{cases} \tag{4.3.54}$$

As already referred to in the introduction, these systems are a special case of (1.2.4). While the parameter dependence in the general form (1.2.4) is not specified, what results here is an implicit, nonlinear dependence on $\mathbf{s}$.

4.3.3.1 The transfer function and moments of multi-parameter systems

From (4.3.54) it follows for $\mathbf{x}$:

$$\mathbf{x}(\mathbf{s}) = \left(\sum_{|\boldsymbol{\alpha}=0|}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{A}_{\boldsymbol{\alpha}}^{\boldsymbol{\alpha}} \right)^{-1} \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{B}_{\boldsymbol{\alpha}} \right) \mathbf{i}(\mathbf{s}), \tag{4.3.55}$$

which gives for $\mathbf{u}(\mathbf{s})$:

$$\mathbf{u}(\mathbf{s}) = \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{C}_{\boldsymbol{\alpha}} \right) \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{A}_{\boldsymbol{\alpha}}^{\boldsymbol{\alpha}} \right)^{-1} \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{B}_{\boldsymbol{\alpha}} \right) \mathbf{i}(\mathbf{s}). \tag{4.3.56}$$

Thus, the transfer function of system $\boldsymbol{\Sigma}_{\text{multi-param}}$ (4.3.54) is given by:

$$\mathbf{H}(s) = \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{C}_{\boldsymbol{\alpha}} \right) \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{A}_{\boldsymbol{\alpha}}^{\boldsymbol{\alpha}} \right)^{-1} \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{B}_{\boldsymbol{\alpha}} \right). \tag{4.3.57}$$

The moment $\boldsymbol{\eta_\gamma}$ of the transfer function $\mathbf{H}(s)$ in $\mathbf{s} = \mathbf{0}$ is given by:

$$\boldsymbol{\eta_\gamma}(\mathbf{0}) = \sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{max}} \mathbf{C}_{\boldsymbol{\alpha}} \mathbf{x}_{\boldsymbol{\gamma}-\boldsymbol{\alpha}}. \tag{4.3.58}$$

This relation can be easily shown by setting the Taylor series expansion of $\mathbf{x}$

$$\mathbf{x} = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \mathbf{x}_{\boldsymbol{\beta}} \mathbf{s}^{\boldsymbol{\beta}} \tag{4.3.59}$$

in the first equation of (4.3.54) and assuming without loss of generality that $\mathbf{i} = 1$. The transfer function then is:

$$\mathbf{H}(\mathbf{s}) = \left(\sum_{|\boldsymbol{\alpha}|=\mathbf{0}}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{max}} \mathbf{s}^{\boldsymbol{\alpha}} \mathbf{C}_{\boldsymbol{\alpha}} \right) \left(\sum_{|\boldsymbol{\beta}|=0}^{\infty} \mathbf{x}_{\boldsymbol{\beta}} \mathbf{s}^{\boldsymbol{\beta}} \right) \tag{4.3.60}$$

$$= \sum_{|\boldsymbol{\beta}|=0}^{\infty} \sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}_{max}|} \mathbf{s}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} \mathbf{C}_{\boldsymbol{\alpha}} \mathbf{x}_{\boldsymbol{\beta}}. \tag{4.3.61}$$

Application of the differential operator D^γ leads to:

$$D^\gamma \, \mathbf{H}(\mathbf{s})\big|_{\mathbf{s}=\mathbf{0}} = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}_{max}|} \frac{(\boldsymbol{\alpha}+\boldsymbol{\beta})!}{(\boldsymbol{\alpha}+\boldsymbol{\beta}-\boldsymbol{\gamma})!} \mathbf{s}^{\boldsymbol{\alpha}+\boldsymbol{\beta}-\boldsymbol{\gamma}} \mathbf{C}_{\boldsymbol{\alpha}} \mathbf{x}_{\boldsymbol{\beta}}. \tag{4.3.62}$$

After evaluation at $\mathbf{s} = \mathbf{0}$ only terms with $\boldsymbol{\alpha} + \boldsymbol{\beta} - \boldsymbol{\gamma} = \mathbf{0}$ remain which leads to the result:

$$\mathbf{H}_{\boldsymbol{\gamma}}(\mathbf{s} = \mathbf{0}) = \frac{1}{\boldsymbol{\gamma}!} \, D^{\boldsymbol{\gamma}} \mathbf{H}(\mathbf{s})\big|_{\mathbf{s}=\mathbf{0}} = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}_{max}|} \mathbf{C}_{\boldsymbol{\alpha}} \mathbf{x}_{\boldsymbol{\gamma}-\boldsymbol{\alpha}}. \tag{4.3.63}$$

4.3.4 The Multi-Parameter Curl-Curl System

The material matrices in the systems (4.2.8), (4.2.21) or (4.2.33) may be considered as depending on a parameter vector $\boldsymbol{\xi}$ containing for example material or geometry parameters. Their general form is described by (1.2.5).

The loss-free shifted curl-curl system (4.2.21) reads then:

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\xi},\text{curl,shift}} : \Big\{ \; & (\mathbf{M}_\varepsilon(\boldsymbol{\xi})(s-s_0)^2 \\ & +2\mathbf{M}_\varepsilon(\boldsymbol{\xi}) s_0 (s-s_0) + \mathbf{M}_\varepsilon(\boldsymbol{\xi}) s_0^2 + \mathbf{A}_{\text{CC}}(\boldsymbol{\xi})) \, \mathbf{x} \quad &= s\mathbf{B}\mathbf{i}, \\ & \mathbf{u} \quad &= \mathbf{C}\mathbf{x}. \end{aligned} \tag{4.3.64}$$

The transfer function of the non-symmetrized loss-free shifted curl-curl system is:

$$\begin{aligned} \mathbf{H}((s-s_0), \boldsymbol{\xi}) = \mathbf{C} \big(& \mathbf{M}_\varepsilon(\boldsymbol{\xi})(s-s_0)^2 + 2\mathbf{M}_\varepsilon(\boldsymbol{\xi}) s s_0 - \mathbf{M}_\varepsilon(\boldsymbol{\xi}) s_0^2 \\ & + \mathbf{C}_{\text{FIT}}^{\text{T}} \mathbf{M}_{\mu^{-1}}(\boldsymbol{\xi}) \mathbf{C}_{\text{FIT}} \big)^{-1} \mathbf{B}. \end{aligned} \tag{4.3.65}$$

The moments are again the Laurent series coefficients defined in (4.2.13) and read:

$$\boldsymbol{\eta}_k(s_0) = \mathbf{C}(\mathbf{M}_\varepsilon(\boldsymbol{\xi})(s-s_0)^2 + 2\mathbf{M}_\varepsilon(\boldsymbol{\xi})ss_0 - \mathbf{M}_\varepsilon(\boldsymbol{\xi})s_0^2 + \mathbf{C}_{\mathrm{FIT}}^{\mathrm{T}}\mathbf{M}_{\mu^{-1}}(\boldsymbol{\xi})\mathbf{C}_{\mathrm{FIT}})^{-(k+1)}\mathbf{B}, \quad k \geq 0. \tag{4.3.66}$$

A reduced curl-curl system of (4.3.64) is defined by:

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},\mathrm{curl,shift}} : \left\{ \begin{array}{rl} \left(\hat{\mathbf{M}}_\varepsilon(\boldsymbol{\xi})(s-s_0)^2 + 2\hat{\mathbf{M}}_\varepsilon(\boldsymbol{\xi})s_0(s-s_0) + \hat{\mathbf{M}}_\varepsilon(\boldsymbol{\xi})s_0^2 + \hat{\mathbf{A}}_{\mathrm{CC}}(\boldsymbol{\xi})\right)\hat{\mathbf{x}} & = s\hat{\mathbf{B}}\mathbf{i}, \\ \mathbf{u} & = \hat{\mathbf{C}}\hat{\mathbf{x}}. \end{array} \right. \tag{4.3.67}$$

Chapter 5

Model Order Reduction for Single-Parameter Systems

MOR techniques based on projection have been described as the problem of determining matrices $\mathbf{V}$ *and* $\mathbf{W}$ *that form an oblique projection of the original system, such that the reduced system approximates the original one in a special sense. How to determine* $\mathbf{V}$ *and* $\mathbf{W}$ *is the objective of this chapter. Most model reduction techniques are based on the retention of some invariant system properties, which are for instance the Hankel singular values, the eigenvalues or the moments of the system. In order to accomplish this retention and determine the corresponding projection matrices* $\mathbf{V}, \mathbf{W}$*, basically, three different types of methods exist: methods based on the singular value decomposition (SVD), methods based on Krylov subspaces and finally methods that are a mixture of SVD- and Krylov based methods. While the SVD methods will be briefly addressed in the introductory part, the focus of this chapter will be set on Krylov based methods. These methods are based on the invariance of some power series coefficients of the transfer impedance, in other words they rely on the moment matching of the original and the reduced system impedances. The approach to those methods will be twofold: first, a way to approximate the impedance of* $\mathbf{\Sigma}$ *by means of power series expansions (partial realization and Padé approximation, which are special cases of rational interpolation) will be given. Then, the connection with Krylov subspaces will be shown.*

5.1 Overview of Basic Approximation Methods

Krylov based reduction techniques are based on the invariance of some power series coefficients of the transfer impedance. In 5.3 it will be shown how they are associated with special Krylov subspaces of the system matrices. Before going into detail in 5.2 and 5.3, a survey of the other type of methods, the SVD-based methods, will be presented.

5.1.1 Balanced Truncation

In balanced truncation, which was introduced in [36], the invariant properties are the Hankel singular values, which are based on the reachability and observability subspaces of the system. More precisely, the method is based on the observation that states that require a large amount of energy to reach are in the span of eigenvectors of the reachability gramian corresponding to small eigenvalues. States that require a large amount of energy in order to be observed lie in the span of eigenvectors of the observability gramian corresponding to small eigenvalues. Thus one way to reduce the original system is to eliminate those states which are difficult to reach or to observe. As the concepts of reachability and observability are basis-dependent, in principal, the difficult-to-reach states are not the same as the difficult-to-observe states. Nevertheless, with an appropriate transformation, a basis can be determined, such that the transformed gramians $\mathcal{P}_t$ and $\mathcal{Q}_t$ are equal, i. e. that the states addressed above are precisely the same.

A reachable, observable and stable system is called balanced, if $\mathcal{P} = \mathcal{Q}$. Furthermore, it is called principal-axis-balanced if $\mathcal{P} = \mathcal{Q} = \operatorname{diag}(\sigma_1 \dots \sigma_n)$, with $\sigma_1 \dots \sigma_n$ being the Hankel singular values. With the Cholesky factor $\mathbf{U}$ of the reachability gramian $\mathcal{P}$ and the eigenvalue decomposition of $\mathbf{U}^{\mathrm{T}}\mathcal{Q}\mathbf{U}$

$$\mathcal{P} = \mathbf{U}\mathbf{U}^{\mathrm{T}}, \quad \mathbf{U}^{\mathrm{T}}\mathcal{Q}\mathbf{U} = \mathbf{K}\Sigma^2\mathbf{K}^{\mathrm{T}}, \tag{5.1.1}$$

the balancing transformation, with which the transformed gramians $\mathcal{P}_t = \mathbf{T}\mathcal{P}\mathbf{T}^{\mathrm{T}}$ and $\mathcal{Q}_t = \mathbf{T}^{-*}\mathcal{Q}\mathbf{T}^{-1}$ are both equal to Σ, is given by $\mathbf{T} = \Sigma^{1/2}\mathbf{K}^{\mathrm{T}}\mathbf{U}^{-1}$. Further details can be found in [1, p. 210].

For the order reduction by balanced truncation, let the linear system $\boldsymbol{\Sigma}$ with the system matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ be balanced with the gramians $\mathcal{P} = \mathcal{Q} = \operatorname{diag}(\sigma_1 \dots \sigma_n) = \Sigma$. Then, the following partitioning is considered:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{pmatrix}, \tag{5.1.2}$$

where $\Sigma_1 \in \mathbb{R}^{p \times p}$ and Σ_2 contains the negligible Hankel singular values. The main advantages of balanced truncation are the preservation of stability and passivity, as well as the existence of a global error bound, i. e. :

$$|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_1|_\infty \leq 2 \sum_{i=k+1}^{q} \sigma_i. \tag{5.1.3}$$

These two properties are proven both for continuous-time as well as for discrete-time systems in [1, p. 212 f.]. Therein, also numerical issues of the balanced truncation as well other types of balancing can be found. However, due to its association with the solution of Lyapunov equations and the SVD, balanced truncation requires $\mathcal{O}(n^2)$ operations and $\mathcal{O}(n^3)$ storage and is thus inefficient for large-scale settings.

There is ongoing work in the field of balanced truncation, e. g. in [42]. Also in [7, 44] balanced truncation of symmetric second-order systems is investigated. Balanced truncation received attention also with respect to multi-parameter systems in [4]. This issue will be revisited in chapter 6.

5.1.2 Proper Orthogonal Decomposition

Closely related to the SVD and balanced truncation is the proper orthogonal decomposition (POD). The starting point is a collection of snapshots:

$$\mathbf{x}_i = \mathbf{x}(t_i), \quad i = 1 \dots N. \tag{5.1.4}$$

With $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_N]$, this state trajectory is written in terms of an orthonormal basis $\mathbf{U}$ and a coefficient matrix $\mathbf{\Gamma}$:

$$\mathbf{X} = \mathbf{U}\mathbf{\Gamma}. \tag{5.1.5}$$

The POD of (5.1.5) seeks to find truncated snapshots which are reconstructed from only k basis vectors:

$$\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \dots \hat{\mathbf{x}}_N] = \mathbf{U}_k \mathbf{\Gamma}_{kk}, \quad k < N, \tag{5.1.6}$$

where $\mathbf{U}_k = [\mathbf{u}_1 \dots \mathbf{u}_k]$ and $\mathbf{\Gamma}_{kk} = [\gamma_1(1:k,1) \dots \gamma_N(1:k,1)]$, such that $\hat{\mathbf{X}}$ approximates $\mathbf{X}$ in some averaged sense. Usually, the second induced norm of $|\mathbf{X} - \hat{\mathbf{X}}|$ is minimized, which is equivalent to the minimization of the second induced norms of the gramians:

$$\mathcal{P} = \sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}} \in \mathbb{R}^{n \times n} \quad \text{and} \quad \hat{\mathcal{P}} = \sum_{i=1}^{N} \hat{\mathbf{x}}_i \hat{\mathbf{x}}_i^{\mathrm{T}} \text{ with } k = \operatorname{rank}(\hat{\mathcal{P}}) < \operatorname{rank}(\mathcal{P}) \tag{5.1.7}$$

of the original and reduced data, respectively. This problem can be solved with the help of the Hankel singular values of $\mathbf{X}$ and a fundamental theorem called Schmidt-Eckart-Young-Mirsky theorem [1, p. 37]. Thus, for the order reduction, the SVD of $\mathbf{X}$ is computed $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}}$. The leading k left singular vectors of $\mathbf{X}$ form the projection matrices, that is $\mathbf{U}_k = \mathbf{V} = \mathbf{W}$.

In [1, p. 279] the similarities between POD and balanced truncation are examined. In particular, if instead of a Galerkin projection, as used above, a Petrov-Galerkin projection ($\mathbf{V} \neq \mathbf{W}$) is applied, and if the input is the impulse response of the system, then the balanced truncation of linear systems is shown to be a POD. It is shown, that in this case the projection on the dominant eigenspace of the product of two gramians leads to a global error bound.

The POD is widely used especially for nonlinear PDEs or nonlinear ODEs. An extensive literature exists for the POD. A selection of contributions can be found e. g. in [1, p. 282].

5.1.3 Modal Approximation

A further truncation method is the modal approximation, which is derived from the eigenvalue decomposition (EVD) of the system matrix $\mathbf{A}$, in contradiction to the previously mentioned techniques which are based on the EVD and balancing of the reachability and observability gramians. A diagonalizable matrix $\mathbf{A}$, e. g. the

curl-curl system matrix in FIT, can be transformed in a basis composed of its eigenvectors. The transfer function $\mathbf{H}$ is decomposed in terms of partial fractions associated with the respective eigenvalues. The reduced system is then obtained by retaining only those partial fractions that are related to poles which have a dominant impact on the time-behavior of the system. Normally, eigenvalues that lie far left of the imaginary axis (all eigenvalues lie on the left half-plane, as otherwise the system would be unstable) corresponding to fast decaying time-instances are of low importance, thus for the reduction the eigenvalues closest to the imaginary axis are chosen. The invariant properties for this reduction method are thus the k dominant poles, i. e. the k eigenvalues with largest real part. The partial fractions, which are related to the remaining poles, are omitted and form the approximation error $\mathbf{H}_{corr}$.

However, as described in [19, 216], and illustrated by an example in [1, p. 283], the nearness of the poles to the imaginary axis is not necessarily related to the dominant behavior of the system, i. e. to the location of the resonances of the system. This difficulty to a priori identify the truly dominant eigenvalues forms a disadvantage of this approach. Nevertheless, in [19, p. 455] a measure of dominance for eigenvalues is defined, which is associated with the reachability and observability of the system. In this way, a connection to the balanced truncation is established, as it is based on the same principle of evaluating the states with respect to their reachability and observability. The order reduction is obtained by projecting on the matrices formed by the eigenvectors defined above.

5.2 Partial Realization, Padé Approximation and Rational Interpolation

5.2.1 Moment Matching

The single-parameter reduction techniques in this work are based on the invariance of some power series coefficients of the transfer impedance. As described in chapter 3, the given original systems are uniquely determined by their impulse response $\mathbf{h}$ or equivalently by the transfer function $\mathbf{H}$, which is the Laplace transform of the impulse response. The transfer functions $\mathbf{H}$ considered here are rational and can be therefore expanded in power series. Thus, for a given system $\mathbf{\Sigma}$, with transfer function $\mathbf{H}$, an approximating lower-dimensional system $\hat{\mathbf{\Sigma}}$, with transfer function $\hat{\mathbf{H}}$, is determined, that matches the leading power series coefficients of $\mathbf{H}$. In the following, applications with several possible power series will be shown.

5.2.1.1 Partial realization

The Laurent series expansions of $\mathbf{H}$ and $\hat{\mathbf{H}}$ around infinity read:

$$\mathbf{H} = \sum_{k=0}^{\infty} \boldsymbol{\eta}_k(\infty) s^{-k} \quad \text{and} \quad \hat{\mathbf{H}} = \sum_{k=0}^{\infty} \hat{\boldsymbol{\eta}}_k(\infty) s^{-k}. \tag{5.2.8}$$

The coefficients $\boldsymbol{\eta}_k(\infty), \hat{\boldsymbol{\eta}}_k(\infty)$ are the Markov parameters of $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Sigma}}$, respectively. The approximation of $\boldsymbol{\Sigma}$ by $\hat{\boldsymbol{\Sigma}}$ in this approach is accomplished by matching the l leading Markov parameters of $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Sigma}}$:

$$\boldsymbol{\eta}_k(\infty) = \hat{\boldsymbol{\eta}}_k(\infty), \quad k = 0, 1, \ldots, l. \tag{5.2.9}$$

This problem is known as partial realization [23].

5.2.1.2 Padé approximation

Alternatively, the power series expansion can be performed around $s = 0$, yielding

$$\mathbf{H} = \sum_{k=0}^{\infty} \boldsymbol{\eta}_k(0) s^k, \quad \text{and} \quad \hat{\mathbf{H}} = \sum_{k=0}^{\infty} \hat{\boldsymbol{\eta}}_k(0) s^k, \tag{5.2.10}$$

respectively. The coefficients $\boldsymbol{\eta}_k(0), \hat{\boldsymbol{\eta}}_k(0)$ are the moments of $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Sigma}}$, respectively. A reduced order model that matches the leading moments of the original model is known as its Padé approximation [26].

5.2.1.3 Shifted Padé approximation

The power series expansions around an arbitrary complex point s_0 yields:

$$\mathbf{H} = \sum_{k=0}^{\infty} \boldsymbol{\eta}_k(s_0) \, (s - s_0)^k, \quad \text{and} \quad \hat{\mathbf{H}} = \sum_{k=0}^{\infty} \hat{\boldsymbol{\eta}}_k(s_0) \, (s - s_0)^k, \tag{5.2.11}$$

respectively. In this case the coefficients $\boldsymbol{\eta}_k(s_0), \hat{\boldsymbol{\eta}}_k(s_0)$ are called shifted moments of $\boldsymbol{\Sigma}$ and $\hat{\boldsymbol{\Sigma}}$, respectively. The reduced order model matching the leading shifted moments of the original model is called a shifted Padé approximation.

5.2.1.4 Multi-point Padé approximation or rational interpolation

It is also possible to use an expansion at multiple points $s_{0_1}, s_{0_2}, \ldots s_{0_i}$. At each interpolation point a different number of moments can be matched. The approxition resulting from this approach is known as multi-point Padé approximation. In case that only one moment per point is matched, it is known as rational interpolation. The power series expansion reads:

$$\mathbf{H} = \sum_{k_i=0}^{\infty} \boldsymbol{\eta}_{k_i}(s_{0_i}) \, (s - s_{0_i})^{k_i}, \quad \text{and} \quad \hat{\mathbf{H}} = \sum_{k_i=0}^{\infty} \hat{\boldsymbol{\eta}}_{k_i}(s_{0_i}) \, (s - s_{0_i})^{k_i}. \tag{5.2.12}$$

5.2.2 Explicit Moment Matching

In the early development of moment matching techniques, the reduced order model was formed from an explicit knowledge of the moments. The respective methods are known as explicit moment matching methods [48, 8] and have been successfully used in control theory since the 19-seventies.

5.2.2.1 Asymptotic waveform evaluation (AWE)

Especially an approach known as Asymptotic Waveform Evaluation (AWE) [40] received large attention. It is based on an iterative calculation of the moments $\boldsymbol{\eta}(0)$, as it is given for the general polynomial multi input multi output (MIMO) case in (5.4.29). The calculation leads to power iterations which are of limited use. This is due to several reasons, one of which being that they can only find the eigenvector corresponding to the largest eigenvalue [65, p. 204 f.]. In practice, this limits the resulting order to a value between 10 and 15 [17].

One idea to overcome this drawback is to use Padé approximations at multiple interpolation points leading to a method called complex frequency hopping (CFH) [10]. Nevertheless, explicit moment matching methods were soon replaced by Krylov subspace methods which will be introduced in the following.

5.3 Krylov Subspace Iterations

The shortcomings of explicit moment matching methods were overcome by using methods based on Krylov subspaces. The Krylov subspace $\mathcal{K}_q(\mathbf{A}, \mathbf{b})$ of a matrix $\mathbf{A}$ with respect to a vector $\mathbf{b}$ is defined as:

$$\mathcal{K}_q(\mathbf{A}, \mathbf{b}) = \{\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \ldots, \mathbf{A}^{q-1}\}. \tag{5.3.13}$$

Replacing $\mathbf{b}$ by $\mathbf{B}$ gives the Krylov subspace $\mathcal{K}_q(\mathbf{A}, \mathbf{B})$ of $\mathbf{A}$ with respect to multiple vectors contained as columns in $\mathbf{B}$ (MIMO case). Obviously, the Krylov subspace of $\mathbf{A}$ is identical to the reachability subspace $\mathcal{R}_q(\mathbf{A})$ defined in chapter 3. Methods based on these subspaces are called Krylov-based methods.

A vast number of names and acronyms have proliferated in association with Krylov subspaces. Nevertheless, all these methods are based on a few fundamental ideas. Different avenues of describing this foundation exist. Following the reasoning of [65], the Arnoldi algorithm [2] will be considered to describe this foundation. Directly connected to the Arnoldi algorithm is the Lanczos algorithm [30], which can be viewed as a specialization of the Arnoldi algorithm to the case where $\mathbf{A}$ is hermitian. Only the basic algorithms and the basic uses will be given. An overview of variants of Krylov-based methods and their uses are contained e. g. in [43, 1, 22, 65].

5.3.1 The Arnoldi Algorithm

The method was introduced in 1951, initially devised to reduce a nonhermitian matrix to Hessenberg form. Starting from the first column, and proceeding column by column, this is done by orthogonal similarity transformations, i. e. $\mathbf{A} = \mathbf{V}\mathbf{H}\mathbf{V}^{\mathrm{T}}$. This equation represents the complete reduction of $\mathbf{A}$ to the Hessenberg matrix $\mathbf{H}$. However, in practical applications the size n of $\mathbf{A}$ is huge, thus only a partial reduction to Hessenberg form is proceeded, i. e. the first k columns of $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{H}$ are considered. Let $\mathbf{V}_k$ be the matrix consisting of the first k columns of $\mathbf{V}$, i. e. the

vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_k$. Let also $\tilde{\mathbf{H}}_k$ be the $(k+1) \times k$ upper-left submatrix of $\mathbf{H}$, which is also a Hessenberg matrix. It follows for $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\tilde{\mathbf{H}}_k$:

$$\begin{bmatrix} & \\ \mathbf{A} & \\ & \end{bmatrix} \underbrace{\left[\begin{array}{c|c|c} & & \\ \mathbf{v}_1 & \ldots & \mathbf{v}_k \\ & & \end{array}\right]}_{\mathbf{V}_k} = \left[\begin{array}{c|c|c} & & \\ \mathbf{v}_1 & \ldots & \mathbf{v}_{k+1} \\ & & \end{array}\right] \underbrace{\begin{bmatrix} h_{11} & \cdots & h_{1k} \\ h_{21} & \cdots & \\ & \ddots & \vdots \\ & & h_{k+1,k} \end{bmatrix}}_{\tilde{\mathbf{H}}_k}. \quad (5.3.14)$$

The last column of this equation gives an $(k+1)$-term recurrence:

$$\mathbf{A}\mathbf{v}_k = h_{1k}\mathbf{v}_1 + \cdots + h_{kk}\mathbf{v}_k + h_{k+1,k}\mathbf{v}_{k+1}, \quad (5.3.15)$$

from which the Arnoldi algorithm follows by simple application of the modified Gram-Schmidt iteration [65, p. 58]. The procedure is given in algorithm 5.1.

Algorithm 5.1 Arnoldi Method

1: $\mathbf{v}_1 = \mathbf{b}/|\mathbf{b}|$, $\mathbf{b}$ arbitrary
2: **for** $k = 1, \ldots,$ **do**
3: $\quad \mathbf{r}_k = \mathbf{A}\mathbf{v}_k$
4: $\quad$ **for** i=1,...,k **do**
5: $\quad\quad h_{jk} = \mathbf{v}^{\mathrm{T}}\mathbf{r}_k$
6: $\quad\quad \mathbf{r}_k = \mathbf{r}_k - h_{jk}\mathbf{v}_j$
7: $\quad$ **end for**
8: $\quad h_{k+1,k} = |\mathbf{r}_k|$
9: $\quad \mathbf{v}_{k+1} = \mathbf{r}_k/h_{k+1,k}$
10: **end for**

At each step of algorithm 5.1, the previous vector $\mathbf{v}_k$ is multiplied with $\mathbf{A}$ and then the resulting vector $\mathbf{r}_k$ is orthogonalized against all previous $\mathbf{v}_i$'s by the Gram-Schmidt procedure. The algorithm stops, if $\mathbf{r}_k = \mathbf{r}_k - h_{jk}\mathbf{v}_j = 0$. Conditions under which this situation occurs can be found, e. g. in [43, p. 148 f.].

A key-result in this context following directly from (5.3.14) is that the $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ form an orthonormal basis of $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$. The proof can be found in [43, p. 146].

Notice, that if $\mathbf{H}_k$ denotes the matrix obtained from $\tilde{\mathbf{H}}_k$ by deleting its last row, the following relations can be shown:

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\mathbf{H}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{r}_k\mathbf{e}_k^{\mathrm{T}}, \quad (5.3.16)$$

where $\mathbf{e}_k$ is the canonical unit vector $\in \mathbb{R}^{k\times 1}$. The term $\mathbf{r}_k\mathbf{e}_k^{\mathrm{T}}$ is the residual $\mathbf{R}_k$, which is by construction orthogonal to the columns of $\mathbf{V}_k$.

It has to be mentioned that for both the Arnoldi and the Lanczos algorithms, block variants, e. g. in [22, p. 485 f.], [43, p .196 f.] and band variants, e. g. [75, p. 53 f.] exist. Nevertheless, here the classical approaches are given.

5.3.2 The Lanczos Algorithm

The Lanczos algorithm was originally proposed as a method for solving linear systems of equations and eigenvalue problems. It can be viewed as a specialization of the Arnoldi algorithm to the case where $\mathbf{A}$ is hermitian. In the following, a further simplification is assumed, that $\mathbf{A}$ is real and symmetric. This implies that $\mathbf{H}_k$ is also real and symmetric and as it is also Hessenberg, it immediately follows that it is tri-diagonal. Thus, the $(k+1)$-recurrence (5.3.15) is replaced by a three-term recurrence. It is accustomed to use different symbols for the Lanczos algorithm. The matrix $\mathbf{H}_k$ is now denoted by $\mathbf{T}_k$, furthermore $h_{kk} = \alpha_k$ and $h_{k+1,k} = h_{k,k+1} = \beta_k$. Then it is:

$$\mathbf{T}_k = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_{k-1} \\ & & & \beta_{k-1} & \alpha_k \end{bmatrix}. \tag{5.3.17}$$

The procedure is stated in algorithm 5.2.

Algorithm 5.2 Basic Lanczos algorithm

1: $\beta_0 = 0, \mathbf{v}_0 = \mathbf{0}, \mathbf{v}_1 = \mathbf{b}/|\mathbf{b}|$, $\mathbf{b}$ arbitrary
2: **for** $k = 1, 2, \ldots$ **do**
3: $\quad \mathbf{r}_k = \mathbf{A}\mathbf{v}_k$
4: $\quad \alpha_k = \mathbf{v}_k^{\mathrm{T}}\mathbf{r}_k$
5: $\quad \mathbf{r}_k = \mathbf{r}_k - \beta_{k-1}\mathbf{v}_{k-1} - \alpha_k\mathbf{v}_k$
6: $\quad \beta_k = |\mathbf{r}_k|$
7: $\quad \mathbf{v}_{k+1} = \mathbf{r}_k/\beta_k$
8: **end for**

5.3.3 The Two-Sided Lanczos Algorithm

If instead of one starting vector, two starting vectors are used, the two-sided Lanczos algorithm is obtained, which is shown in algorithm 5.3.

Other variants, as e. g. the implicitly restarted Arnoldi and Laczos methods (IRAM) can be found in [43, p. 143 f.].

Summarizing the preceding results leads to the observation that all algorithms above rely on a common basic iteration, called the basic Krylov iteration. In the kth step with $\mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{b} \in \mathbb{R}^{n\times 1}$ and the start-vector $\mathbf{v}_1 = \mathbf{b}/|\mathbf{b}|$ it is:

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{r}_k\mathbf{e}_k^{\mathrm{T}}, \tag{5.3.18}$$

where $\mathbf{V}_k = [\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_k] \in \mathbb{R}^{n\times k}$ is orthonormal, $\mathbf{H}_k = \mathbf{V}_k^{\mathrm{T}}\mathbf{A}\mathbf{V}_k \in \mathbb{R}^{k\times k}$, $\mathbf{r}_k \in \mathbb{R}^{n\times k}$ and $\mathbf{e}_k$ is the canonical unit vector $\in \mathbb{R}^{k\times 1}$.

Algorithm 5.3 Two-sided Lanczos algorithm

1: $\beta_1 = \sqrt{|\mathbf{b}^{\mathrm{T}}\mathbf{c}^{\mathrm{T}}|}, \gamma_1 = \mathrm{sign}(\mathbf{b}^{\mathrm{T}}\mathbf{c}^{\mathrm{T}})\beta_1, \quad \mathbf{v}_1 = \mathbf{b}/\beta_1, \mathbf{w}_1 = \mathbf{c}^{\mathrm{T}}/\gamma_1,$
2: **for** $k = 1, 2, \ldots$ **do**
3: $\quad \alpha_k = \mathbf{w}_k^{\mathrm{T}}\mathbf{A}\mathbf{v}_k$
4: $\quad \mathbf{r}_k = \mathbf{A}\mathbf{v}_k - \alpha_k\mathbf{v}_k - \gamma_k\mathbf{v}_{k-1}, \quad \mathbf{q}_k = \mathbf{A}^{\mathrm{T}}\mathbf{w}_k - \alpha_k\mathbf{w}_k - \beta_k\mathbf{w}_{k-1}$
5: $\quad \beta_{k+1} = \sqrt{|\mathbf{r}_k^{\mathrm{T}}\mathbf{q}_k|}, \quad \gamma_{k+1} = \mathrm{sign}(\mathbf{r}_k^{\mathrm{T}}\mathbf{q}_k)\beta_{j+1}$
6: $\quad \mathbf{v}_{k+1} = \mathbf{r}_k/\beta_{k+1}, \quad \mathbf{w}_{k+1} = \mathbf{q}_j/\gamma_{j+1}$
7: **end for**

The Krylov iteration has three main uses: the iterative solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$, the iterative approximation of the eigenvalues of $\mathbf{A}$ and the approximation of linear systems by moment matching.

The Arnoldi algorithm can be used to solve systems of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$. The idea is at each step k to approximate the exact solution $\mathbf{x}_0$ by the vector $\mathbf{x}_k$ that minimizes the norm of the residual $\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_n$. The standard algorithm for this purpose is the GMRES (generalized minimal residuals). Details to this topic can be found e. g. in [65, p. 266 f.]. Other methods can be found in [1, 43].

The second application, the iterative approximation of the eigenvalues of $\mathbf{A}$, is achieved by using the eigenvalues of the projected matrix $\mathbf{H}_k$ as approximations of the most dominant eigenvalues of $\mathbf{A}$. In particular, in each step, the Krylov iteration is used to determine matrices $\mathbf{V}_k$ and $\mathbf{W}_k$, with

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{R}_k, \tag{5.3.19}$$

where $\mathbf{R}_k$ is orthogonal to $\mathbf{W}_k$. This condition is known as Petrov-Galerkin condition. The eigenvalues of the Hessenberg matrix $\mathbf{H}_k$

$$\mathbf{H} = \mathbf{W}^{\mathrm{T}}\mathbf{A}\mathbf{V}, \text{ with } \mathbf{W}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_k, \text{ i. e. } \mathbf{W} \text{ and } \mathbf{V} \text{ are biorthogonal} \tag{5.3.20}$$

can be computed by standard methods and are called Arnoldi or Lanczos estimates, depending on which method is used. They are also known as Ritz values. Of course, as naturally $k \ll n$, not all eigenvalues are computed. However, typically the extreme eigenvalues, i. e. the eigenvalues near the edge of the spectrum $\mathbf{A}$, are identified. In most applications these are exactly the eigenvalues of interest and this feature made the Krylov iteration to one of the most important iterative methods. In [22] and [1, p. 324] the use of Krylov methods with respect to invariant subspace approximation is motivated by using the Rayleigh quotient.

Nevertheless, the application of interest in this chapter is the system approximation by moment matching. In [24, p. 26 f.] a nice brief historical overview of model order reduction by projection is given. The first significant work, in which the mathematic relation between a Krylov-based algorithm and partial realizations was recognized, is [23]. In [68], Padé and shifted Padé approximations were related to variants of Krylov subspaces. In his thesis [24], E. Grimme provides a general framework of Krylov based methods for the general case of rational interpolation. These topics will be subject to the next section.

5.4 Krylov Subspace Projection Methods for Moment Matching

The projection framework described in the introductory section has, so far, not been set in context with the Krylov subspaces described in this section. This will be done in the following, first for the linear and the curl-curl system in section 5.4.1 and then for the general polynomial system in section 5.4.2. If the system is given in state-space data, the partial realization problem, as well as the Padé interpolation and the rational interpolation problems can be solved in a numerically efficient way by projecting on the matrices $\mathbf{V}$ and/or $\mathbf{W}$ produced by the Arnoldi algorithm, the Lanczos algorithm or the two-sided Lanczos algorithm. In this case, the starting vectors are not arbitrary as in the case of eigenvalue approximations, but are determined by the system matrices $\mathbf{B}$ and $\mathbf{C}$.

The single-parameter systems arising in this work are the linear system (4.2.8), which is skew-symmetric, the curl-curl system (4.2.24), and the polynomial system (4.2.33). The projection-based reduction of polynomial systems will be addressed separately in the Padé approximation section. The linear and the curl-curl system constitute two competing approaches which have been subject to an extensive cost comparison, with respect to the required matrix-vector multiplications and orthogonalization steps in [75]. The curl-curl system turns out to be approximately half as expensive as the linear system. Therefore, the curl-curl system will take precedence over the linear system, nevertheless, for simplicity reasons the procedures will be shown mainly using the linear system.

Key properties of Krylov-based methods are primarily that the moment matching is achieved without explicitly computing the moments, as opposed to the explicit moment matching techniques, and that the procedure is implemented iteratively.

5.4.1 Linear and Curl-Curl System

5.4.1.1 Partial realization by Krylov subspaces

Consider the linear system $\mathbf{\Sigma}_{\text{FIT,linear,shift}}$ (4.2.9) and let it be denoted by $\mathbf{\Sigma}$ in this section. Furthermore, consider the Krylov subspaces $\mathcal{K}_q(\mathbf{A}, \mathbf{B})$ and $\mathcal{K}_q(\mathbf{A}^{\mathrm{T}}, \mathbf{C}^{\mathrm{T}})$ and matrices $\mathbf{V}$ and $\mathbf{W}$, such that:

$$\det(\mathbf{W}^{\mathrm{T}}\mathbf{V}) \neq 0, \tag{5.4.21a}$$

$$\mathcal{K}_q(\mathbf{A}, \mathbf{B}) \subseteq \operatorname{colsp}\{\mathbf{V}\}, \tag{5.4.21b}$$

$$\mathcal{K}_q(\mathbf{A}^{\mathrm{T}}, \mathbf{C}^{\mathrm{T}}) \subseteq \operatorname{colsp}\{\mathbf{W}\}. \tag{5.4.21c}$$

A reduced system $\hat{\mathbf{\Sigma}}$ as in (4.2.18) with $\mathbf{V}$ and $\mathbf{W}$ as defined above, is a partial realization of $\mathbf{\Sigma}$ and matches $2q$ Markov parameters (5.2.8), if both equations (5.4.21b) and (5.4.21c) hold. If only one of these relations holds, only q Markov parameters are matched.

In fact, the projections are distinguished between one-sided and two-sided projections. Depending on this feature, the number of values $\boldsymbol{\eta}$ matched can be determined. For one-sided projections q values are matched. If, additionally, relation (5.4.21c) holds, then, twice as many values, i. e. $2q$ values, can be matched with a reduced system of the same dimension as above. However, on the other hand, the choice of $\mathbf{V}$ and $\mathbf{W}$ determines if the stability and passivity are retained in the reduced system. The Lanczos algorithm naturally uses a one sided projection, as only the matrix $\mathbf{V}$ is used. This does not hold true for the Arnoldi and the two-sided Lanczos. Whereas $2q$ moments can be matched, the stability is not retained in these both cases.

For the moment matching derivation, consider the first q Markov parameters of $\boldsymbol{\Sigma}$, $\mathbf{C}\mathbf{A}^k\mathbf{B}$, and the first q Markov parameters of $\hat{\boldsymbol{\Sigma}}$, $\hat{\mathbf{C}}\hat{\mathbf{A}}^k\hat{\mathbf{B}}$. With the choice $\mathbf{V}, \mathbf{W}$ as above, follows:

$$\begin{aligned}\hat{\mathbf{C}}\hat{\mathbf{A}}^k\hat{\mathbf{B}} &= \mathbf{C}\mathcal{R}_q(\mathbf{A},\mathbf{B})\mathbf{W}^{\mathrm{T}}\mathbf{A}^k\mathcal{R}_q(\mathbf{A},\mathbf{B})\mathbf{e}_1 \\ &= \mathbf{C}\mathcal{R}_q(\mathbf{A},\mathbf{B})\mathbf{W}^{\mathrm{T}}\mathbf{A}^k\mathbf{B} \\ &= \mathbf{C}\mathcal{R}_q(\mathbf{A},\mathbf{B})\mathbf{e}_{k+1} \\ &= \mathbf{C}\mathbf{A}^k\mathbf{B}.\end{aligned} \tag{5.4.22}$$

Equality of moments means that $\hat{\boldsymbol{\Sigma}}$ is a partial realization of $\boldsymbol{\Sigma}$ that matches q Markov parameters.

Partial realizations of curl-curl system (4.2.24) can be handled with the transformation $s' = s^2$. In [75, p. 57 f.], partial realizations are investigated with respect to FIT-systems, in particular the linear and the curl-curl system. Issues such as the symmetric case, stop criteria, error calculation, convergence speed-up or conditioning are addressed. Partial realizations received little attention in the model order reduction community in comparison with Padé approximations, which will be described subsequently. This was mainly due to the larger size of reduced models they produce, but also because they are particularly efficient for easy-to-invert material matrices, as they are in FIT systems. For systems resulting from Finite Elements this is not necessarily the case.

5.4.1.2 Padé via Lanczos (PVL)

A Padé approximation $\hat{\boldsymbol{\Sigma}}$ of the linear system $\boldsymbol{\Sigma}$ is obtained by imposing the following conditions:

$$\det(\mathbf{V}^{\mathrm{T}}\mathbf{V}) \neq 0, \tag{5.4.23a}$$

$$\mathcal{K}_q(\bar{\mathbf{A}}, \bar{\mathbf{B}}) \subseteq \mathrm{colsp}\{\mathbf{V}\}, \tag{5.4.23b}$$

$$\mathcal{K}_q(\bar{\mathbf{A}}^{\mathrm{T}}, \mathbf{C}^{\mathrm{T}}) \subseteq \mathrm{colsp}\{\mathbf{V}\}, \tag{5.4.23c}$$

where $\bar{\mathbf{A}} = (s_0\mathbf{I} - \mathbf{A})^{-1}, \bar{\mathbf{B}} = \bar{\mathbf{A}}\mathbf{B}$. The reduced system $\hat{\boldsymbol{\Sigma}}$ matches $2q$ moments if both last relations in (5.4.23) hold. The projection matrix $\mathbf{V}$ is calculated by means of the Lanczos algorithm. This approach was first introduced in [48]. In [18] and [25] they were used independently for large network analysis problems.

The choice of interpolation points has been investigated in [24]. The optimal interpolation point is the barycenter of poles in the respective frequency range. As this is not a priori known, the center of the frequency range is a good alternative. A new work which may be useful in the error analysis and in the selection of expansion points in Krylov-based model order reduction is presented in [78]. It is based on the investigation of the Sylvester equations associated to the Krylov subspaces.

An extensive analysis of Padé approximations with respect to the linear and the curl-curl FIT systems is again contained in [75, p. 72 f.]. In particular, simplifications for symmetric and skew-symmetric matrices (corresponding to the symmetric curl-curl system and the skew-symmetric linear system) are described. Furthermore, as by nature this approach is not based on a symmetric and thus stability preserving projection, modifications that ensure stability and passivity in the reduced system are proposed. In this context, one particular method known as passive reduced-order interconnect macro-modeling algorithm (PRIMA), which was initially introduced in [38], is used. The same method was used in [49], in EMC simulations. A short introduction is also given in [45]. PRIMA is based on the usage of $\mathbf{V}$ both as projection as well as test matrix, while the projection of the matrices is done explicitly. For the symmetrized linear system, the reduced system reads:

$$\hat{\mathbf{H}}(s) = \mathbf{B}^{\mathrm{T}}\mathbf{V}(s\mathbf{I} + \mathbf{V}^{\mathrm{T}}\mathbf{A}\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}\mathbf{B}. \tag{5.4.24}$$

For real s_0, this reduced order model preserves stability and passivity. The moments matched are half of the maximal possible moments that can be matched. The proofs can be found in [38] and in [20].

5.4.1.3 Multiple interpolation points

Multi-point Padé approximations as defined by (5.2.12), can also be performed by projection methods. In [24], methods resembling the Lanczos- and Arnoldi procedures, known as rational Krylov methods, are proposed.

The reduction using the PRIMA approach can also be extended to multiple interpolation points:

$$\mathbf{V}_q = [\mathbf{V}_1, \mathbf{V}_2 \ldots \mathbf{V}_I], \tag{5.4.25}$$

where each matrix $\mathbf{V}_i$ is the projection matrix corresponding to the interpolation point s_{0_i}. The dimension of each $\mathbf{V}_i$ can be chosen separately.

As described in [75, p. 78], for simulations in electrodynamics usually one interpolation point is sufficient. Except for a smoother error distribution in the frequency range, there is no large benefit of the usage of multiple interpolation points. Nevertheless, making use of local subspaces corresponding to distinct points in the frequency range is the key-idea for the neighboring-subspace method introduced for multi-parameter systems in chapter 6.

Table 5.1, similar as in [24], gives an overview of the different approximation methods and the respective power series expansions.

Approximant name	Power series expansion
Partial realization	$\sum_{k=0}^{\infty} \boldsymbol{\eta}_k s^{-k}$
Padé approximation	$\sum_{k=0}^{\infty} \boldsymbol{\eta}_k s^{k}$
Shifted Padé	$\sum_{k=0}^{\infty} \boldsymbol{\eta}_k (s-s_0)^{k}$
Rational interpolation or shifted Padé	$\sum_{k_i=0}^{\infty} \boldsymbol{\eta}_{k_i} (s-s_{0_i})^{k_i}$

Table 5.1: *Overview of the power series expansions used in the different approximation methods.*

5.4.1.4 Two-step Lanczos

The methods described above are either very fast but producing reduced systems with relatively large order (e. g. partial realization), or they are slow but producing small systems (e. g. modal approximation, Padé approximation). For a really fast frequency sweep the numerical cost of a Padé approximation is typically too high. On the other hand, the systems obtained by partial realizations are normally too large for an effective sweep over a large number of frequencies samples. The combination of a method from the first category and a method from the second category leads to an approach where in a first step the original system is reduced for example with a partial realization to a model of moderate size. In a second step, this model is reduced by means of a fast Padé approximation. The overall time is negligibly higher than for a single partial realization, but the model size is the same as that of a direct Padé approximation. This approach was first presented in [76] under the name two-step Lanczos, whereas it is not restricted to the Lanczos algorithm for the reduced system computation. Details on numerical issues can be found in [75, p. 85 f.].

5.4.2 Padé Approximations for Polynomial Systems

The implicit moment matching methods based on the Arnoldi or the Lanczos algorithm are not suitable for polynomial systems of the form (4.2.33). This section is devoted to this class of systems. Consider a polynomial system (4.2.33) which is shifted about a frequency parameter s_0:

$$\sum_{k=0}^{N_A} \left((s-s_0)^k \mathbf{A}'_k \right) \mathbf{X} = \sum_{k=0}^{N_B} \left((s-s_0)^k \mathbf{B}_k \right) \mathbf{i}, \tag{5.4.26}$$

with

$$\mathbf{A}'_k = \sum_{i=k}^{N_A} \begin{pmatrix} n \\ k \end{pmatrix} \mathbf{A}_i, \quad \mathbf{B}'_k = \sum_{i=k}^{N_B} \begin{pmatrix} n \\ k \end{pmatrix} \mathbf{B}_i. \tag{5.4.27}$$

Throughout this section the MIMO case will be considered, the single input single output (SISO) case easily follows for $m = 1$. The Taylor expansion of $\mathbf{X}$ reads:

$$\mathbf{X}(s) = \sum_{k=0}^{\infty} (s - s_0)\mathbf{X}_k, \tag{5.4.28}$$

leading to the following recursion to calculate $\mathbf{X}_n$:

$$\begin{aligned}
\mathbf{X}_0 &= \mathbf{A}_0'^{-1}\mathbf{B}_0', \\
\mathbf{X}_1 &= \mathbf{A}_0'^{-1}(\mathbf{B}_1' - \mathbf{A}_1'\mathbf{X}_0), \\
\mathbf{X}_2 &= \mathbf{A}_0'^{-1}(\mathbf{B}_2' - \mathbf{A}_1'\mathbf{X}_1 - \mathbf{A}_2'\mathbf{X}_0), \\
&\vdots \\
\mathbf{X}_n &= \mathbf{A}_0'^{-1}(\mathbf{B}_n' - \sum_{k=1}^{\min(N_a,n)} \mathbf{A}_k'\mathbf{X}_{n-k}).
\end{aligned} \tag{5.4.29}$$

This recursion motivates to introduce the concept of higher-order Krylov subspaces. For matrices $\mathbf{M}_1, \mathbf{M}_2, \ldots \mathbf{M}_M \in \mathbb{R}^{n\times n}$, a vector $\mathbf{x} \in \mathbb{R}^{n\times 1}$ and the sequence:

$$\begin{aligned}
\mathbf{X}_0 &= \mathbf{X}, \\
\mathbf{X}_1 &= \mathbf{M}_1\mathbf{X}_0, \\
\mathbf{X}_2 &= \mathbf{M}_1\mathbf{X}_1 + \mathbf{M}_2\mathbf{X}_0, \\
&\vdots \\
\mathbf{X}_k &= \mathbf{M}_1\mathbf{X}_{k-1} + \mathbf{M}_2\mathbf{X}_{k-2} + \cdots + \mathbf{M}_n\mathbf{X}_{k-n},
\end{aligned} \tag{5.4.30}$$

a higher-order Krylov subspace is defined as:

$$\mathcal{K}_q(\mathbf{M}_1, \mathbf{M}_2, \ldots \mathbf{M}_M, \mathbf{X}) = \text{span}\{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{q-1}\}. \tag{5.4.31}$$

A very efficient approach to calculate a stable basis of $\mathcal{K}_q(\mathbf{M}_1, \mathbf{M}_2, \ldots \mathbf{M}_M, \mathbf{x}) = \text{span}\{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{q-1}\}$, is known as well conditioned asymptotic waveform evaluation (WCAWE), and was introduced in [51]. Prior to giving the WCAWE iteration, several definitions have to be given, here for block matrices. Let $\mathbf{E}_k$ be matrices consisting of $m \times m$ blocks. Except of the k-th block which is a unity matrix, all the other elements are zero. Furthermore, let $\mathbf{U}$ be an $n \times n$ upper-triangular non-singular matrix. Using $\mathbf{U}$, the following terms can be defined:

$$\mathbf{P}_w(n, k) = \prod_{t=w}^{k} \mathbf{U}(t : n - k + t - 1, t : n - k + t - 1)^{-1}. \tag{5.4.32}$$

The indexing of $\mathbf{U}$ and $\mathbf{P}_w$ follows the notation of MATLAB® [35] and it refers to blocks in case of $m > 1$.

Then, the WCAWE is based on the following iteration, here given in its block-variant which was elaborated in [29]:

$$\begin{aligned}
\tilde{\mathbf{V}}_0 &= \mathbf{A}_0'^{-1}\mathbf{B}_0', \\
\tilde{\mathbf{V}}_1 &= \mathbf{A}_0'^{-1}(\mathbf{B}_1'\mathbf{E}_1^{\mathrm{T}}\mathbf{P}(2,1)\mathbf{E}_1 - \mathbf{A}_1'\mathbf{V}_1), \\
&\vdots \\
\tilde{\mathbf{V}}_n &= \mathbf{A}_0'^{-1}\left(\sum_{k=1}^{\min(n_B,n-1)} \mathbf{B}_k'\mathbf{E}_1^{\mathrm{T}}\mathbf{P}_{\mathbf{U}_1}(n,k)\mathbf{E}_{n-k} - \mathbf{A}_1'\mathbf{V}_{k-1}\right. \\
&\qquad \left.\sum_{k=2}^{\min(n_A,n-1)} \mathbf{A}_k'[\mathbf{V}_0 \dots \mathbf{V}_{n-k}]\mathbf{P}_2(n,k)\mathbf{E}_{n-k}\right).
\end{aligned} \tag{5.4.33}$$

With $[\tilde{\mathbf{V}}_0, \tilde{\mathbf{V}}_1 \dots \tilde{\mathbf{V}}_n]$ from the iteration above and the vectors $[\mathbf{X}_0, \mathbf{X}_1, \dots \mathbf{X}_n]$ given by the iteration (5.4.29), it can be shown by induction [51] that:

$$\mathrm{span}\{\tilde{\mathbf{V}}_0\tilde{\mathbf{V}}_1 \dots \tilde{\mathbf{V}}_n\} = \mathrm{span}\{\mathbf{X}_0\, \mathbf{X}_1\, \dots\, \mathbf{X}_n\}. \tag{5.4.34}$$

Basically, $\mathbf{U}$ can be an arbitrary $n \times n$ upper-triangular non-singular matrix. Nevertheless, the best conditioning for this approach is achieved, if $\mathbf{VU}$ is obtained by orthogonalizing $\mathbf{X}_n$. Therefore, a modified Gram-Schmidt is incorporated in equations (5.4.33). The matrices $\mathbf{V}_k$ are connected with the $\mathbf{X}_k$ by:

$$\tilde{\mathbf{V}}_n = \mathbf{X}_n - [\mathbf{V}_0 \dots \mathbf{V}_{n-1}][\mathbf{V}_0 \dots \mathbf{V}_{n-1}]^{\mathrm{T}}\mathbf{X}_n. \tag{5.4.35}$$

In case of $m > 1$, the new vectors have to be orthogonalized with respect to each other by means of the QR-factorization:

$$\tilde{\mathbf{V}}_n = \mathbf{V}_n\mathbf{U}_n. \tag{5.4.36}$$

The matrix $\mathbf{U}$ can be then expanded by:

$$\mathbf{U}(1 : n-1, n) = [\mathbf{V}_0 \dots \mathbf{V}_{n-1}]^{\mathrm{T}}\mathbf{X}_n \quad \text{and} \quad \mathbf{U}(n,n) = \mathbf{U}_n. \tag{5.4.37}$$

It has to be mentioned, that despite the improvement in stability, similar to Padé approximations, a high numerical cost is required for this approach, which has been implemented and tested for FIT-Systems by [29] and will be a substantial part of the contraction method for multi-parameter systems described in chapter 6.

Chapter 6

Model Order Reduction for Multi-Parameter Systems

The reduction techniques presented so far apply to single-parameter systems. Main subject in this thesis, though, is the development of order reduction techniques for multi-parameter systems, where the system matrices are material and geometry dependent, while particular focus is put on the latter. After giving a brief overview of existing approaches for parametric model order reduction, this chapter presents several methods for FIT-systems. While material parameter variations are easy to implement in all methods, the geometrical variations exhibit some difficulties. For each method this issue is touched on in particular. Furthermore, each method description is accompanied with a discussion about numerical properties, for instance about techniques for choosing the interpolation points, about modeling error-estimations, the stability issue, multiple-input multiple-output systems etc. A simple example is used to demonstrate the methods.

6.1 Methods Overview

It is on hand to extend the reduction principle underlying the single-parameter case to multi-parameter systems. One approach to design parametric reduction methods is therefore based on the retention of some invariant system properties. One property that can be retained, are the moments, as they have been presented in chapter 4 for multi-parameter system representations. The related methods are associated with higher-order Krylov subspaces.

The first work related to parametric order reduction based on Krylov projection methods was [74], where a moment matching approach is presented for a system depending on two parameters. The method calculates a projection matrix $\mathbf{V}$, such that the reduced order model matches the first moments of the transfer function with respect to both parameters. The generalization of this work to the multi-parameter case was presented in [15]. Based on the Taylor series expansion of the transfer function, appropriate higher-order Krylov subspaces are derived. The main drawback of this method is that the reduced model order grows very rapidly even for a

small number of parameters. This method was again revisited in [28] to reduce parametric systems based on matrix-valued spline approximations. A two-sided Arnoldi algorithm featuring a deflation procedure is presented that extends the two-sided Arnoldi variant in [5] to the parametric case. The moments were recursively calculated by a robust method presented in [6]. Nevertheless, no satisfactory reductions were obtained.

In this work two alternative methods based on multi-parameter moment retention will be presented. Both methods are iterative and stable and circumvent the ill-conditioned problem of directly calculating the moments. The first resembles the single-parameter Arnoldi algorithm, as in the calculation iteration only the orthonormal basis vectors are multiplied with the system matrices. Therefore, it is called multi-parameter Arnoldi method. The second algorithm is based on the deduction of the original problem on a sequence of single-parameter models. In [17] this method was called the contraction method.

Alternatively, PMOR approaches that are not based on higher-order Krylov subspaces exist. In [4], an approach is presented that is based on balanced truncation and interpolation of the transfer functions of weighted linear reduced order models evaluated at several expansion points of the parameter-set. The expansion points have been chosen according to the sparse grids method [3]. The method is very sensitive to the choice of the expansion points. The order of the systems corresponding to the chosen expansion points defines the order of the final reduced order model. Although it is a very powerful method, difficulties are encountered in systems with weekly damped modes. Furthermore, in order to apply the sparse grids method, the parameter dependency of the matrices has to be known. This method has not been implemented for FIT-systems.

The methods described above require an explicit dependence on the parameters. They require the systems in the form (4.3.54), in other words, that the system matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$ of the parametric systems depend explicitly on the parameters. According to the classification in the introduction on whether the frequency is considered separately from the other parameters in the parameter-vector or not, the respective systems are denoted by $\mathbf{\Sigma_s}$ and $\mathbf{\Sigma_\xi}$, respectively. In order to be used for FIT systems, the system equations need to be appropriately adapted so that they exhibit an explicit dependence on the parameters. While material parameters enter the curl-curl formulation (4.2.20) in an explicit linear dependence, the system matrices depend implicitly and non-linearly on the geometry parameters. As described in chapter 4, a linearization of the system equations poses the requirement of a constant mesh topology. Geometry parameter variations are thus difficult to handle.

The trend is therefore towards a different category of methods that can handle the implicit parameter dependence of the system matrices. These methods use the FIT systems in the initial form (1.2.4) and (1.2.5). In this case, there is no requirement of a constant mesh topology.

The most straight-forward approach is inspired by the multi-point Padé approximation for the single-parameter case 5.2.1.4. Let $\mathbf{\Sigma}_i$ be the system corresponding

to each parameter-set, $\mathcal{K}_i$ be the respective Krylov subspace and $\mathbf{V}_i$ the projection matrix of this local system. The projection matrix of the parametric system is defined as the composition of the local projection matrices $\mathbf{V}_i$. This method was called neighboring-subspace method in [56], as the subspaces seem to be, simply speaking, close to each other, at least for small parameter variations or for a small sensitivity on a specific parameter.

While this approach offers a greater flexibility concerning geometry variations, there are still some restrictions, as will be explained in detail in section 6.5.

An interesting approach is introduced in [39] which is also based on the reduction of local original models in the parameter space. A parametric reduced order model is calculated by interpolating the system matrices of the local reduced models, while it is aimed to find compatible system representations with optimal interpolation properties. To this purpose, a so-called matrix matching is applied, that modifies the system realizations at some parameters such that the corresponding matrices become as similar as possible. Matrix matching is based only on the reduced models, i. e. it is not related to the full model. This implies that it can even be applied to geometry parameter variations associated to different mesh topologies. The matrix interpolation was successfully used in applications, i. e. in the field of computing optimal configurations of self-optimizing systems, as described in [28]. Nevertheless, this approach is associated with the costly solution of Lyapunov equations. This method has not been applied to FIT systems, it is therefore an open question to compare the obtained accuracy with the accuracy of the neighboring-subspace method and moreover, if the additional effort outweighs an eventual accuracy improvement.

Finally, there are methods that arise from analogous considerations as the PMOR, though they follow different avenues of approaching the problem. One major representative of these methods is the sensitivity analysis, which is a general approach to a priori obtain information about the influence of parameter variations on the system behavior, for instance during the design and optimization process of computer aided engineering.

In order to quantify the influence of parameter variations, it is convenient to use a gradient representation in the parameter space which corresponds to a linearization around the nominal working point. In this way, the system behavior can be predicted for small parameter changes. A general approach leading to such derivative information required for the gradient representation is given by the sensitivity analysis. In case of the so-called adjoint sensitivity analysis, the gradient information results from evaluating the system twice, regardless of the number of parameters. At the same time, for the adjoint sensitivity analysis the derivatives of the system matrix entries with respect to the design parameters have to be computed. For the adjoint sensitivity analysis in high-frequency structures see for instance [37]. Also in [46] a brief introduction is given.

Another method that is connected to model order reduction but is not a reduction method itself is based on the interpolation of the reduced order S-parameters. Analogously to the neighboring-subspace method, local systems $\mathbf{\Sigma}_i$ are considered. However, instead of considering the local projection matrices as in the neighboring-

subspace method or the local transfer impedance, the reduced S-parameters are determined. The interpolation of these reduced S-parameters gives an approximation to the original system S-parameters. As the interpolation step is performed after the reduction of each local model, considerations about how to avoid that the mesh is affected by geometry variations so that the associated $\mathbf{V}$ remains constant, are not of relevance here.

Both the sensitivity analysis and the S-parameter fitting are not covered in the discussion of this work. Nevertheless, their close connection to MOR should always be kept in mind. For practical applications it depends on the specific problem requirements which method is best.

In this chapter, the procedure will be to first discuss the moment matching methods, i. e. the multi-parameter Arnoldi and the contraction method, and then the neighboring-subspace method.

6.2 Moment Matching for Multi-Parameter Systems

Consider systems of the form (4.3.54). For simplicity reasons SISO systems are considered. The projection spaces in the multi-port case consist of the SISO projection spaces, thus, this simplification does not imply a generality restriction. Consider the multi-parameter Taylor series expansion of $\mathbf{x}(\mathbf{s})$ and $\mathbf{H}(\mathbf{s})$ around the expansion point $\mathbf{s}_0$:

$$\mathbf{x}(\mathbf{s}) = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}, \quad \mathbf{H}(\mathbf{s}) = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \boldsymbol{\eta}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}, \tag{6.2.1}$$

where $\mathbf{x}_{\boldsymbol{\beta}}$ and $\boldsymbol{\eta}_{\boldsymbol{\beta}}$ are the moments of $\mathbf{x}$ and $\mathbf{H}$, respectively. In order to calculate the moments $\mathbf{x}_{\boldsymbol{\beta}}$, the Taylor series expansion of $\mathbf{x}(\mathbf{s})$ is substituted in (4.3.54):

$$\left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)\left(\sum_{|\boldsymbol{\beta}|=0}^{\infty} \mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}\right) = \left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s}). \tag{6.2.2}$$

The moments can be iteratively determined by:

$$\mathbf{x}_{\boldsymbol{\beta}} = \mathbf{A}_{\mathbf{0}}^{-1}\left(\mathbf{b}_{\boldsymbol{\beta}} - \sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{A}_{\boldsymbol{\alpha}}\mathbf{x}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}\right). \tag{6.2.3}$$

The equation described above defines a multi-level recursion. In each step all $\mathbf{x}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}$, with $|\boldsymbol{\beta}| \leq q$ (q is the order of the truncated Taylor series) and $1 \leq |\boldsymbol{\alpha}| \leq |\boldsymbol{\alpha}|_{\max}$ are recalled. The moments of $\boldsymbol{\eta}_{\boldsymbol{\beta}}$ can be calculated with the help of $\mathbf{x}_{\boldsymbol{\beta}}$ according to [17, 11]:

$$\boldsymbol{\eta}_{\boldsymbol{\beta}}(\mathbf{s}_0) = \sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{c}_{\boldsymbol{\alpha}}\mathbf{x}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}. \tag{6.2.4}$$

A reduced order system of (4.3.54) is then given by:

$$\hat{\boldsymbol{\Sigma}}: \begin{cases} \left(\sum\limits_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\hat{\mathbf{A}}_{\boldsymbol{\alpha}}\right)\hat{\mathbf{x}}(\mathbf{s}) &= \left(\sum\limits_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\hat{\mathbf{b}}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s}), \\ \mathbf{u}(\mathbf{s}) &= \left(\sum\limits_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\hat{\mathbf{c}}_{\boldsymbol{\alpha}}\right)\hat{\mathbf{x}}(\mathbf{s}). \end{cases} \tag{6.2.5}$$

with the abbreviations

$$\hat{\mathbf{A}}_{\boldsymbol{\alpha}} = \mathbf{W}^{\mathrm{T}}\mathbf{A}_{\boldsymbol{\alpha}}\mathbf{V}, \quad \hat{\mathbf{b}}_{\boldsymbol{\alpha}} = \mathbf{W}^{\mathrm{T}}\mathbf{b}_{\boldsymbol{\alpha}}, \quad \hat{\mathbf{c}}_{\boldsymbol{\alpha}} = \mathbf{c}_{\boldsymbol{\alpha}}\mathbf{V}. \tag{6.2.6}$$

Methods for choosing $\mathbf{V}$ and $\mathbf{W}$ will be discussed in the following sections.

Analogously to the original model, the multi-parameter Taylor series expansion of $\hat{\mathbf{x}}(\mathbf{s})$ and $\hat{\mathbf{H}}(\mathbf{s})$ around the expansion point $\mathbf{s}_0$ is:

$$\hat{\mathbf{x}}(\mathbf{s}) = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \hat{\mathbf{x}}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}, \quad \hat{\mathbf{H}}(\mathbf{s}) = \sum_{|\boldsymbol{\beta}|=0}^{\infty} \hat{\boldsymbol{\eta}}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}, \tag{6.2.7}$$

with $\hat{\mathbf{x}}_{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\eta}}_{\boldsymbol{\beta}}$ representing the moments of $\hat{\mathbf{x}}$ and $\hat{\mathbf{H}}$, respectively. Analogously to the original model, the moments $\hat{\mathbf{x}}_{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\eta}}_{\boldsymbol{\beta}}$ are given by the respective recursive iterations:

$$\hat{\mathbf{x}}_{\boldsymbol{\beta}} = \hat{\mathbf{A}}_{\mathbf{0}}^{-1}\left(\hat{\mathbf{b}}_{\boldsymbol{\beta}} - \sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \hat{\mathbf{A}}_{\boldsymbol{\alpha}}\hat{\mathbf{x}}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}\right), \quad \hat{\boldsymbol{\eta}}_{\boldsymbol{\beta}}(\mathbf{s}_0) = \sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\leq|\boldsymbol{\alpha}|_{\max}} \hat{\mathbf{c}}_{\boldsymbol{\alpha}}\hat{\mathbf{x}}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}. \tag{6.2.8}$$

If $\mathbf{W} = \mathbf{V}$ and $\mathbf{V}$ is chosen such that

$$\det(\mathbf{W}^{\mathrm{T}}\mathbf{A}_{\mathbf{0}}\mathbf{V}) \neq 0, \quad \text{span}\left\{\bigcup_{|\boldsymbol{\beta}|=0}^{|\boldsymbol{\beta}|\leq q} \mathbf{x}_{\boldsymbol{\beta}}\right\} \subseteq \text{colsp}\mathbf{V}, \tag{6.2.9}$$

then the first moments of the multi-parameter Taylor series expansion of $\mathbf{x}(\mathbf{s})$ and $\mathbf{H}(\mathbf{s})$ around an expansion point $\mathbf{s}_0$ match those of $\hat{\mathbf{x}}(\mathbf{s})$ and $\hat{\mathbf{H}}(\mathbf{s})$, respectively, [17, 11], i.e.

$$\boldsymbol{\eta}_{\boldsymbol{\beta}} = \hat{\boldsymbol{\eta}}_{\boldsymbol{\beta}}, \; \forall \boldsymbol{\beta} : |\boldsymbol{\beta}| \leq q. \tag{6.2.10}$$

Analogously to the single-parameter case, the direct calculation of the moments is ill-conditioned and therefore replaced by stable and efficient algorithms, e. g. the two moment matching techniques shown next.

6.3 Multi-Parameter Polynomial Arnoldi Method

Let the first k moments $\mathbf{x}_{\boldsymbol{\beta}}$ be numbered by $i = 1, \ldots, k$, i. e. $\mathbf{x}_{\boldsymbol{\beta}_i}$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be an orthonormal basis of span $\{\mathbf{x}_{\boldsymbol{\beta}_1}, \ldots, \mathbf{x}_{\boldsymbol{\beta}_k}\}$ which is wished to be extended

by $\mathbf{v}_{k+1}$ with the help of the moment $\mathbf{x}_{\boldsymbol{\beta}}$. With the reduced QR-factorization of $\mathbf{X}_k = [\mathbf{x}_{\boldsymbol{\beta}_1} \dots \mathbf{x}_{\boldsymbol{\beta}_k}]$ it is:

$$\underbrace{\left(\begin{array}{c|c|c} \mathbf{x}_{\boldsymbol{\beta}_1} & \dots & \mathbf{x}_{\boldsymbol{\beta}_k} \end{array}\right)}_{\mathbf{X}_k} = \underbrace{\left(\begin{array}{c|c|c} \mathbf{v}_1 & \dots & \mathbf{v}_k \end{array}\right)}_{\mathbf{V_k}} \underbrace{\begin{pmatrix} \mathbf{h}_{\boldsymbol{\beta}_1} & & \\ & \ddots & \\ & & \mathbf{h}_{\boldsymbol{\beta}_k} \end{pmatrix}}_{\mathbf{H}_k}, \tag{6.3.11}$$

with $\mathbf{h}_{\boldsymbol{\beta}_1} \dots \mathbf{h}_{\boldsymbol{\beta}_k}$ being the coordinate vectors. It is thus for each i:

$$\mathbf{x}_{\boldsymbol{\beta}_i} = \mathbf{V}_k \mathbf{h}_{\boldsymbol{\beta}_i}. \tag{6.3.12}$$

The starting vector $\mathbf{v}_1$ and its corresponding coordinate $\mathbf{h}_{\boldsymbol{\beta}_1}$ are determined with the help of the unit vector $\mathbf{e}_1$ according to:

$$\mathbf{v}_1 = \frac{\mathbf{A_0^{-1}} \mathbf{b_0}}{|\mathbf{A_0^{-1}} \mathbf{b_0}|}, \quad \mathbf{h}_{\boldsymbol{\beta}_1} = |\mathbf{A_0^{-1}} \mathbf{b_0}| \mathbf{e}_1. \tag{6.3.13}$$

The recursion (6.2.3) obtains the form:

$$\mathbf{x}_{\boldsymbol{\beta}} \;=\; \mathbf{A_0^{-1}} \left(\mathbf{b}_{\boldsymbol{\beta}} - \sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}| \le |\boldsymbol{\alpha}|_{\max}} \mathbf{A}_{\boldsymbol{\alpha}} \mathbf{V}_k \mathbf{h}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}\right), \tag{6.3.14}$$

which is then separated in terms of $\mathbf{h}_k$ and $\mathbf{r}_k$ into horizontal and perpendicular components of span$\{\mathbf{V}\}$, respectively. For $\mathbf{x}_{\boldsymbol{\beta}}$ it can be written:

$$\mathbf{x}_{\boldsymbol{\beta}} = \mathbf{V}_k \mathbf{h}_k + \mathbf{r}_k = \mathbf{V}_k \mathbf{h}_k + \frac{\mathbf{r}_k}{|\mathbf{r}_k|} |\mathbf{r}_k| = \begin{pmatrix} \mathbf{V}_k & \mathbf{v}_{k+1} \end{pmatrix} \begin{pmatrix} \mathbf{h}_k \\ |\mathbf{r}_k| \end{pmatrix}. \tag{6.3.15}$$

Thus, the new basis vector in the sequence and the corresponding coordinate vector are:

$$\mathbf{v}_{k+1} = \frac{\mathbf{r}_k}{|\mathbf{r}_k|}, \quad \mathbf{h}_{\boldsymbol{\beta}_{k+1}} = \begin{pmatrix} \mathbf{h}_k \\ |\mathbf{r}_k| \end{pmatrix}. \tag{6.3.16}$$

The terms $\mathbf{h}_k$ and $\mathbf{r}_k$ can be calculated by:

$$\mathbf{h}_k = \mathbf{V}_k^{\mathrm{T}} \mathbf{A_0} \left(\mathbf{b}_{\boldsymbol{\beta}} - \sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}| \le |\boldsymbol{\alpha}|_{\max}} \mathbf{H}_{\boldsymbol{\alpha},k} \mathbf{h}_{\boldsymbol{\alpha}-\boldsymbol{\beta}}\right), \tag{6.3.17a}$$

$$\mathbf{r}_k = (\mathbf{I} - \mathbf{V}_k \mathbf{V}_k^{\mathrm{T}}) \mathbf{A_0} \mathbf{b}_{\boldsymbol{\beta}} - \sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}| \le |\boldsymbol{\alpha}|_{\max}} \mathbf{R}_{\boldsymbol{\alpha},k} \mathbf{h}_{\boldsymbol{\beta}-\boldsymbol{\alpha}}, \tag{6.3.17b}$$

where $\mathbf{H}_{\boldsymbol{\alpha},k}$ and $\mathbf{R}_{\boldsymbol{\alpha},k}$ are the horizontal and perpendicular part of $\mathbf{A_0^{-1}} \mathbf{A}_{\boldsymbol{\alpha}}\ \mathbf{V}_k$ in (6.3.14) with respect to colsp $\{\mathbf{V}_k\}$, i.e.:

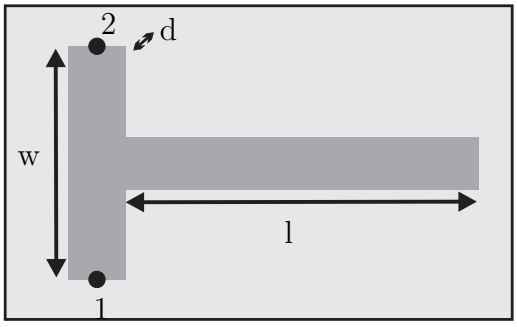
$$\mathbf{A_0^{-1}}\ \mathbf{A}_{\boldsymbol{\alpha}} \mathbf{V}_k = \mathbf{V}_k \mathbf{H}_{\boldsymbol{\alpha},k} + \mathbf{R}_{\boldsymbol{\alpha},k}, \quad \mathbf{V}_k^{\mathrm{T}} \mathbf{R}_{\boldsymbol{\alpha},k} = 0. \tag{6.3.18}$$

Algorithm 6.1 Multi-parameter Arnoldi method

Given $\mathbf{V}_{\boldsymbol{\alpha}}, \mathbf{b}_{\boldsymbol{\alpha}}$
Initialize: $\mathbf{v}_1 = \frac{\mathbf{A_0^{-1}b_0}}{|\mathbf{A_0^{-1}b_0}|}$ and $\mathbf{h}_{\boldsymbol{\beta}_1} = |\mathbf{A_0^{-1}b_0}|\mathbf{e}_1$, $k = 2$.
for $r = 2, r \leq q, r = r + 1$ **do**
  for $\forall \mathbf{b} : |\mathbf{b}| = r$ **do**
    Calculate $\mathbf{H}_{\boldsymbol{\alpha},k}$ and $\mathbf{R}_{\boldsymbol{\alpha},k}$ in (6.3.18)
    Calculate $\mathbf{h}_k$ in (6.3.17a)
    Calculate $\mathbf{h}_r$ in (6.3.17b)
    Set $\mathbf{v}_{k+1} = \frac{\mathbf{r}_k}{|\mathbf{r}_k|}$ and
    $\mathbf{h}_{\boldsymbol{\beta}_{k+1}} = \begin{pmatrix} \mathbf{h}_k \\ |\mathbf{r}_k| \end{pmatrix}$.
    k=k+1
  end for
end for

The procedure is summarized in algorithm 6.1.

The microstrip of figure 6.1 will be used throughout this chapter to test and compare the reduction methods. The length l, the width w and the depth d are considered as parameters. In [55] the length-variation is considered separately for the multi-parameter Arnoldi algorithm, here the situation is analyzed, that all three parameters are varied simultaneously, that is a three-dimensional variation is considered. The maximal admissible variation range for the multi-parameter Arnoldi method and the subsequent contraction method is given in figure 6.1. This chapter concentrates on geometry parameters, material parameters are considered in chapter 7. Furthermore, the examples therein are larger and of more complexity.



Parameter Range	
frequency s [GHz]	$0 \ldots 5$
length l [mm]	$8 \ldots 9$
width w [mm]	$10 \ldots 11$
depth d [mm]	$10 \ldots 10.1$

Figure 6.1: *Geometric structure and parameter range for the microstrip line model.*

The example models are generated by the commercial software package CST MICROWAVE STUDIO (MWS) [13]. In order to admit a large geometry variation range, a relatively coarse mesh is chosen. Of course, this is not an exact solution to the microstrip line problem, nevertheless in this chapter of main interest is the error directly stemming from the order reduction techniques. The original and the reduced system are compared by means of their S-parameters which are calculated from the impedance as described in chapter 3. At this point it has to be mentioned that for all methods described in this chapter the projection matrices are kept in

storage. This means that also the field solutions of the original and the reduced system could be compared. Nevertheless, only a comparison of the S-parameters has been analyzed. The calculations are performed in MATLAB® by making use of the system matrices produced by MWS. This gives a direct comparison of two MATLAB® solutions obtained from the same system matrices and reveals the error made by the respective order reduction technique. The reference S-parameter curves for the microstrip line are given in figure 6.2.

The error between the complex reduced-system S-parameters $\hat{S}$ and the reference S-parameters S is defined by:

$$\mathbf{f}_{\text{mult-arnoldi}}(s_i, \boldsymbol{\xi}_j) = \mathbf{f}(s_i, \boldsymbol{\xi}_j) = |\Delta S| = |\hat{S}(s_i, \boldsymbol{\xi}_j) - S(s_i, \boldsymbol{\xi}_j)|. \tag{6.3.19}$$

The S-parameters are discrete functions of the frequency samples $s_i, i = 1 \dots n_s$, where $n_s = 1000$ throughout this work, and $\boldsymbol{\xi}_j, j = 1 \dots n_{\boldsymbol{\xi}}$. In this example it is $\xi = (l, w, d)$. In order to have a measure for the error over the parameter space, the maximum and the mean error values have to be determined. The maximum over the frequency samples is:

$$f_{\max_{s_i}}(\boldsymbol{\xi}_j) = \max_{s_i} \{\mathbf{f}(s_i, \boldsymbol{\xi}_j)\}. \tag{6.3.20}$$

The mean value over the frequency sample range is given by:

$$\bar{f}_{s_i}(\boldsymbol{\xi}_j) = \frac{\sum_{i=1}^{n_s} \mathbf{f}(s_i, \boldsymbol{\xi}_j)}{n_s}. \tag{6.3.21}$$

The maximum and the mean value of $\bar{f}_{s_i}(\boldsymbol{\xi}_j)$ over the samples $\boldsymbol{\xi}_j$ are respectively:

$$f_{\max_{s_i, \boldsymbol{\xi}_j}} = \max_{\boldsymbol{\xi}_j} \{\bar{f}_{s_i}(\boldsymbol{\xi}_j)\}, \tag{6.3.22a}$$

$$\bar{f}_{s_i, \boldsymbol{\xi}_j} = \frac{\sum_{j=1}^{n_{\boldsymbol{\xi}}} \bar{f}_{s_i}(\boldsymbol{\xi}_j)}{n_{\boldsymbol{\xi}}}. \tag{6.3.22b}$$

Let t_{red} denote the time needed to set up the reduced order model and t_S the time to calculate the S-parameters of the original model from its system matrices, while $t_{\hat{S}}$ the time to calculate the reduced S-parameters from the reduced system matrices. Table 6.1 shows for each order $q = 0 \dots 6$ the errors $\bar{f}_{s_i, \boldsymbol{\xi}_j}$ and $f_{\max_{s_i, \boldsymbol{\xi}_j}}$, the times t_{red} and $t_{\hat{S}}$, as well as the size m of the respective projection matrix $\mathbf{V}$. As a comparison, the time t_S to calculate the original S-parameters from the imported system matrices in MATLAB® is approximately 6 min.

The multi-parameter Arnoldi method will not be used further in this work, as the contraction method presented next exhibits better numerical properties.

q	m	$\bar{f}_{s_i,\boldsymbol{\xi}_j}$	$f_{\max_{s_i,\boldsymbol{\xi}_j}}$	t_{red}/s	$t_{\hat{S}}/\text{s}$
0	2	0.76	0.77	0.86	0.1
1	10	0.68	0.7	163.5	0.4
2	30	0.84	0.85	580.6	0.54
3	70	0.37	0.3796	$1.55 \cdot 10^3$	4.6
4	126	0.1	0.15	$3.5 \cdot 10^3$	11
5	194	0.08	0.0882	$7.58 \cdot 10^3$	31.3
6	240	0.07	0.0726	$1.143 \cdot 10^4$	57

Table 6.1: *The table summarizes the mean and maximum error values as defined in equations (6.3.22) for increasing orders q. The value m gives the size of the projection matrix. t_{red} is the time needed to set up the reduced order model and $t_{\hat{S}}$ is the time to calculate the S-parameters from the reduced system matrices.*

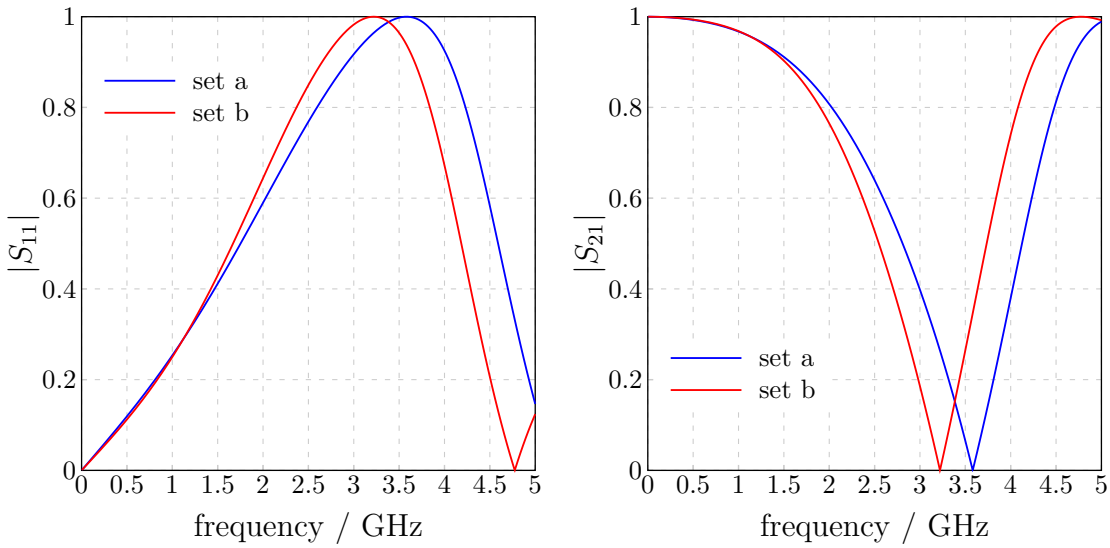


Figure 6.2: *The* MATLAB® *reference S-parameters S_{11} and S_{22} calculated for the bounds of the parameter range, set a $(l, w, d) = (8, 10, 1)$ and set b $(9, 11, 1.1)$.*

6.4 Contraction Method

Starting from (4.3.54), the vector $\mathbf{x}$ can be expressed as follows:

$$\begin{aligned}\mathbf{x}(\mathbf{s}) &= \mathbf{A}_{\mathbf{0}}^{-1}\left(\mathbf{I}_n + \sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)^{-1}\left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s}) \\ &= \sum_{k=0}^{\infty}\mathbf{A}_{\mathbf{0}}^{-1}\left(\sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)^{k}\left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}} \mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s}). \qquad (6.4.23)\end{aligned}$$

By equating with the multi-parameter Taylor-series expansion of (6.2.1), and having used the matrix-equivalent of the geometric series expansion (4.3.42), it is obtained:

$$\sum_{k=0}^{\infty}\sum_{|\boldsymbol{\beta}|=k}\mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}=\sum_{k=0}^{\infty}\mathbf{A}_{\mathbf{0}}^{-1}\left(\sum_{|\boldsymbol{\alpha}|=1}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)^{k}\left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s}), \tag{6.4.24}$$

from which follows:

$$\sum_{|\boldsymbol{\beta}|=k}\mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)(\mathbf{s}-\mathbf{s}_0)^{\boldsymbol{\beta}}=\mathbf{A}_{\mathbf{0}}^{-1}\left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)^{k}\left(\sum_{|\boldsymbol{\alpha}|=0}^{|\boldsymbol{\alpha}|\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s}). \tag{6.4.25}$$

Let $\mathbf{p}=\mathbf{s}-\mathbf{s}_0$, and notice that the left side of (6.4.25) describes a homogenous polynomial in r variables with degree k. Sets $\mathcal{S}_0\subset\ldots\mathcal{S}_k\subset\ldots\mathcal{S}_q\subset\mathbb{R}^r$ are considered, in which $\mathcal{S}_k, k=0\ldots q$ is unisolvent in homogenous polynomials of degree k in r variables, i. e. any polynomial with r variables and degree k is completely determined by its values at the points of $\mathcal{S}_k$ [21]. For efficiency reasons, $\mathcal{S}_k$ should be hierarchical. For example, $\mathcal{S}_k$ can be given by:

$$\mathcal{S}_k=\{(\boldsymbol{\gamma},q-|\boldsymbol{\gamma}|,\ |\boldsymbol{\gamma}|\le k)\},\ \ k=0\ldots q, \tag{6.4.26}$$

in which $\boldsymbol{\gamma}$ is a multi-index of dimension $r-1$. Then, for each $\mathcal{S}\supseteq\mathcal{S}_k, 0\le k\le q$, holds:

$$\bigcup_{|\boldsymbol{\beta}|=k}\text{colsp}\ \{\mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)\}=\bigcup_{\mathbf{p}\in\mathcal{S}}\text{colsp}\left\{\sum_{|\boldsymbol{\beta}|=k}\mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)\mathbf{p}^{\boldsymbol{\beta}}\right\}. \tag{6.4.27}$$

For $|\boldsymbol{\beta}|\le q$ and by using (6.4.25), it follows:

$$\begin{aligned}\bigcup_{|\boldsymbol{\beta}|\le q}\text{colsp}\{\mathbf{x}_{\boldsymbol{\beta}}(\mathbf{s}_0)\}&=\bigcup_{k=0}^{q}\bigcup_{\mathbf{p}\in\mathcal{S}}\text{colsp}\left\{\mathbf{A}_{\mathbf{0}}^{-1}\left(\sum_{|\boldsymbol{\alpha}|=0}^{\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)^{k}\left(\sum_{|\boldsymbol{\alpha}|=0}^{\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s})\right\}\\&=\bigcup_{\mathbf{p}\in\mathcal{S}}\bigcup_{k=0}^{q}\text{colsp}\left\{\mathbf{A}_{\mathbf{0}}^{-1}\left(\sum_{|\boldsymbol{\alpha}|=0}^{\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right)^{k}\left(\sum_{|\boldsymbol{\alpha}|=0}^{\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s})\right\}\\&=\bigcup_{\mathbf{p}\in\mathcal{S}}\mathcal{K}_{q+1}\left\{\mathbf{A}_{\mathbf{0}}^{-1}\left(\sum_{|\boldsymbol{\alpha}|=0}^{\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{A}_{\boldsymbol{\alpha}}\right),\left(\sum_{|\boldsymbol{\alpha}|=0}^{\le|\boldsymbol{\alpha}|_{\max}}\mathbf{s}^{\boldsymbol{\alpha}}\mathbf{b}_{\boldsymbol{\alpha}}\right)\mathbf{i}(\mathbf{s})\right\}.\end{aligned} \tag{6.4.28}$$

Equation (6.4.28) states that the determination of the multi-parameter moments can be reduced to the calculation of Krylov subspaces of single-parameter models at the points $\mathbf{p}\in\mathcal{S}_k$. In [17], these systems are called contractions of the multi-parameter polynomial system in the point $\mathbf{p}$, and therefore the method is called contraction

method. Figure 6.3 visualizes the set of points $\mathcal{S}_0, \mathcal{S}_1$ and $\mathcal{S}_2$ in $\mathbb{R}^3$. $\mathcal{S}_0$ contains only point $\boldsymbol{\xi}_0$, $\mathcal{S}_1$ contains $\boldsymbol{\xi}_0$ and two further points (red dots) of the parameter set and $\mathcal{S}_2$ contains the points of $\mathcal{S}_1$ as well as the three points that are indicated by red dots. Formally, the kth set is defined by equation (6.4.26). Here, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2) \geq \mathbf{0}$ are multi-indices of dimension 2.

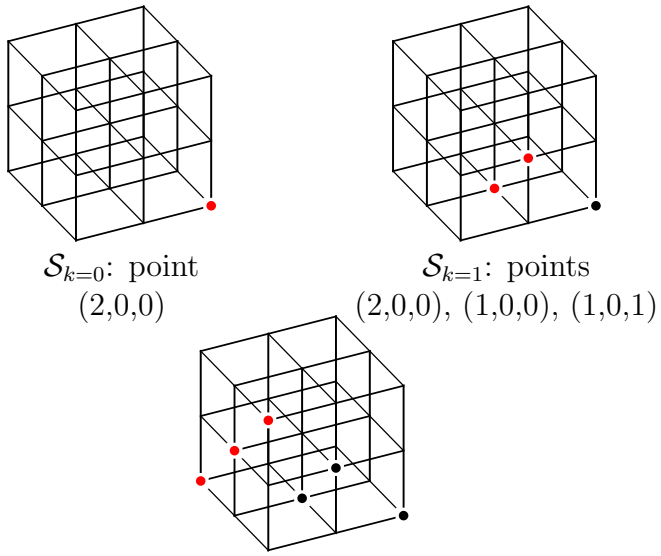


Figure 6.3: *Hierarchical bases of unisolvent sets in* $\mathbb{R}^3$ *for the contraction method for q=2.*

Algorithm 6.2 summarizes the contraction method.

Algorithm 6.2 Multi-parameter moment matching by contraction method

```
1: V
2: q
3: Calculate hierarchical bases S_0 ... S_q
4: for all p ∈ S_q do
5:    U = WCAWE ( Σ(s)|_p , q )
6:    for k = 0 ... q do
7:       if p ∈ S_k then
8:          V(:, v_i + (1 : blocksize)) = U(:, k · blocksize + (1 : blocksize)) %block-
            size:number of columns of B
9:       end if
10:   end for
11: end for
12: V = qr(V) % Orthogonalization by QR-decomposition
```

q	m	$\bar{f}_{s_i,\boldsymbol{\xi}_j}$	$f_{\max_{s_i,\boldsymbol{\xi}_j}}$	t_{red}/s	$t_{\hat{S}}$/s
0	2	0.76	0.78	8.75	0.06
1	10	0.69	0.7	35.45	0.15
2	30	0.79	0.83	98.34	0.9
3	70	0.34	0.34	214.18	3.55
4	140	0.0098	0.001	386.95	21.15
5	252	$3.196 \cdot 10^{-4}$	0.0012	621.31	60.04
6	420	$1.208 \cdot 10^{-4}$	0.00011	991.95	205
7	660	$3.515 \cdot 10^{-5}$	$3.514 \cdot 10^{-5}$	$1.6 \cdot 10^{3}$	601

Table 6.2: *The table summarizes the mean and maximum error values as well as the reduction and evaluation times t_{red} and $t_{\hat{S}}$, respectively, for different orders q for the contraction method.*

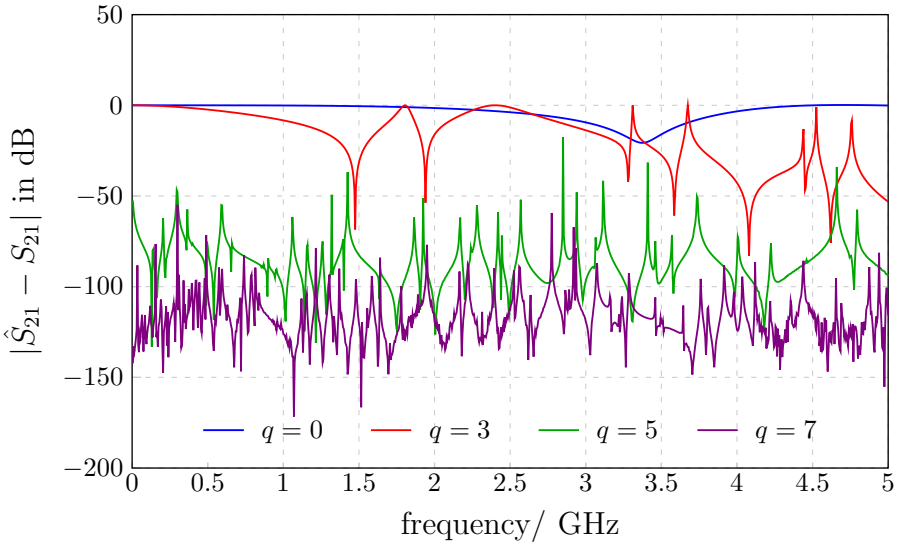


Figure 6.4: *Logarithmic error $\mathbf{f}_{\text{contraction}} = |\Delta S_{21}| = |\hat{S}_{21} - S_{21}|$, obtained with different reduction orders q for an arbitrary sample in the parameter range.*

The same test model as in the previous section is used. Table 6.2 gives for each order $q = 0 \ldots 7$ the errors $\bar{f}_{s_i,\boldsymbol{\xi}_j}$ and $f_{\max_{s_i,\boldsymbol{\xi}_j}}$, the times t_{red} and $t_{\hat{S}}$, as defined in the previous section, as well as the size m of the respective projection matrix $\mathbf{V}$. Figure 6.5 shows how the reduced S_{21}-parameter curve gradually approaches the MATLAB® reference curve, while figure 6.4 shows the error $\mathbf{f}_{\text{contraction}} = |\hat{S} - S|$, obtained with reductions of orders $q = 0, 3, 5, 7$, for an arbitrary sample in the parameter range.

Both methods introduced above require an explicit dependence on the parameters, as in (4.3.54), which is naturally not given for FIT systems in (4.2.28). The curl-curl equation (4.2.28) can only be linearized as described in section 4.3.1 when the mesh topology is fixed for all parameter changes. Obviously, this is a strong limiting factor for geometrical variations, as the systems result from FIT-models with automatically created meshes, which are not necessarily the same.

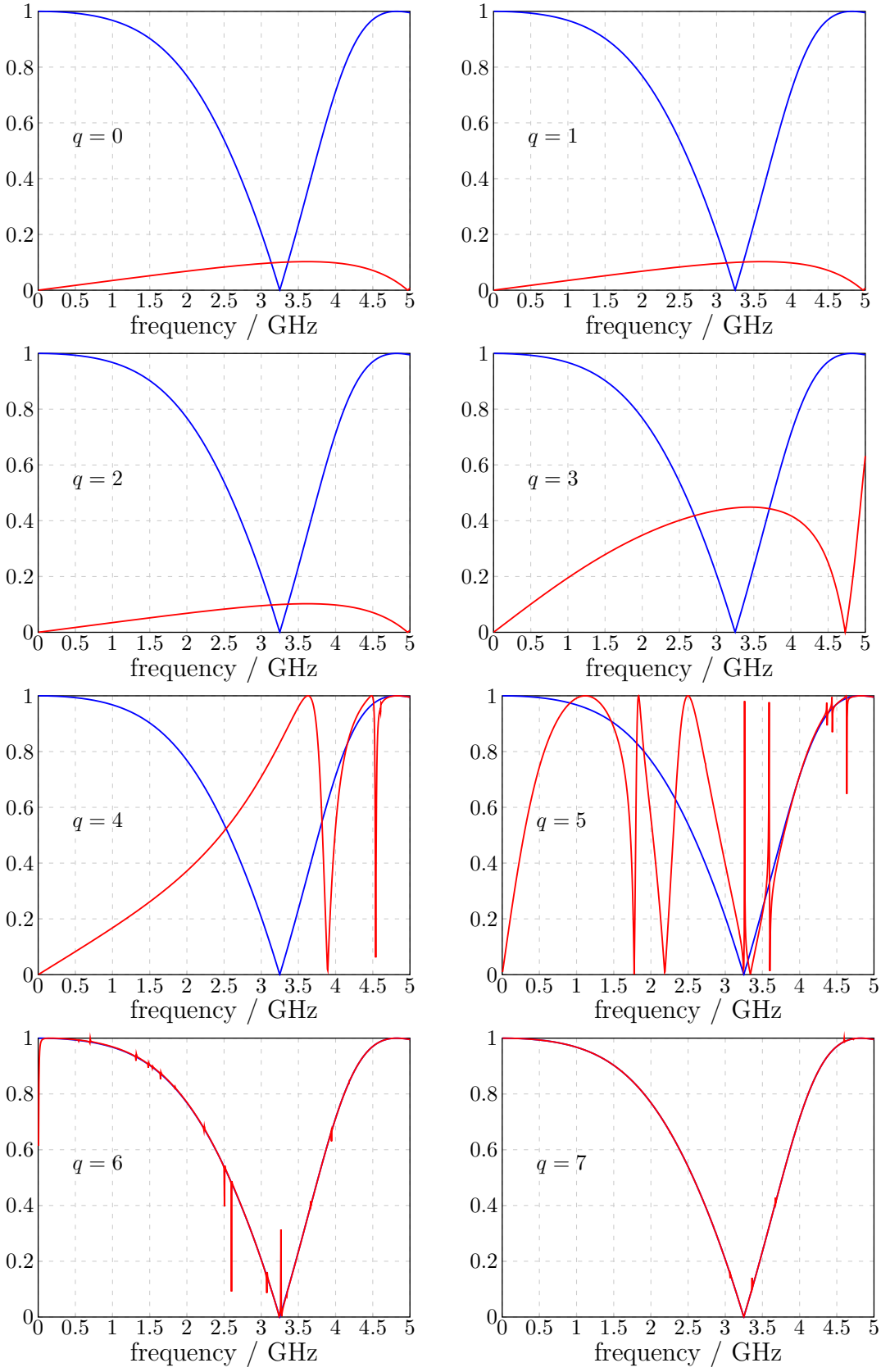


Figure 6.5: *The reduced S-parameter $\hat{S}_{21}$ curve for $q = 0 \ldots 7$ gradually approximates the original curve S_{21}. This approaching way is typical for the contraction method.*

6.5 Neighboring-Subspace Method

The necessary condition of topology preservation for parametric MOR with geometry parameters is relaxed in the approach presented in the following. For this approach, the parametric loss-free shifted curl-curl system (4.3.64) which is a special form of (1.2.5) is used. The system is considered at the discrete parameter values $\boldsymbol{\xi}$, where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_{r-1})$ denotes the vector containing parameters, excluding the frequency parameter s

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\xi},\text{curl,shift}} : \Big\{ \; & (\mathbf{M}_\varepsilon(\boldsymbol{\xi})(s-s_0)^2 \\ & +2\mathbf{M}_\varepsilon(\boldsymbol{\xi})s_0(s-s_0) + \mathbf{M}_\varepsilon(\boldsymbol{\xi})s_0^2 + \mathbf{A}_{\text{CC}}(\boldsymbol{\xi}))\,\mathbf{x} \;\; &&= s\mathbf{B}\mathbf{i}, \\ & \mathbf{u} &&= \mathbf{C}\mathbf{x}. \end{aligned} \tag{6.5.29}$$

6.5.1 Projection on a "Neighboring" Subspace

The observation that the S-parameter curves change only slightly for small variations around an expansion point $\boldsymbol{\xi}_0$, leads to the assumption that the matrices $\mathbf{V}_\xi$ related to each $\boldsymbol{\Sigma}_\xi$ in the neighborhood of $\boldsymbol{\xi}_0$ differ also only slightly. Therefore, it makes sense to investigate how large the error is, if the systems $\boldsymbol{\Sigma}_\xi$ in the neighborhood of $\boldsymbol{\xi}_0$ are reduced uniformly with $\mathbf{V}_0$, instead of with their own projection matrices $\mathbf{V}_\xi$, whereby all projection matrices are obtained with the single-parameter Arnoldi method. This question is pursued in the following.

The microstrip line of figure 6.1 which has already been used in the previous sections, will illustrate the considerations. The length l is chosen as a parameter, that is the parameter vector is $\boldsymbol{\xi} = l$. Within the variation range of figure 6.1, the mesh remains constant, therefore it is possible to use the projection matrix corresponding to $\Delta l = 0$, let it be denoted by $\mathbf{V}_0$, in order to project all systems corresponding to $\Delta l = 0 \ldots 1$. The error $\mathbf{f}_{\text{nsm}}(s,l) = |\Delta\hat{S}_{21}| = |\hat{S}_{21} - S_{21}|$ for $\Delta l = 0, 0.1, 0.3, 0.5, 0.8, 1$, is shown in figure 6.6. The resulting error justifies the name "neighboring subspaces".

6.5.2 Projection on Merged Local Projection Matrices

In continuation of this though, several interpolation points $\boldsymbol{\xi}_i, i = 1 \ldots N$, can be selected to form the matrix

$$\mathbf{V}_{\text{all}} = [\mathbf{V}_1\, \mathbf{V}_2\, \ldots \mathbf{V}_N]. \tag{6.5.30}$$

Let m_i be the number of columns of each $\mathbf{V}_i$. In most subspaces corresponding to $\mathbf{V}_i$, common directions appear. In order to set up $\mathbf{V}$, the directions are sorted by relative importance with the help of the singular value decomposition (SVD) [1], i. e. :

$$\mathbf{V}_{\text{all}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{N}^{\text{T}}, \tag{6.5.31}$$

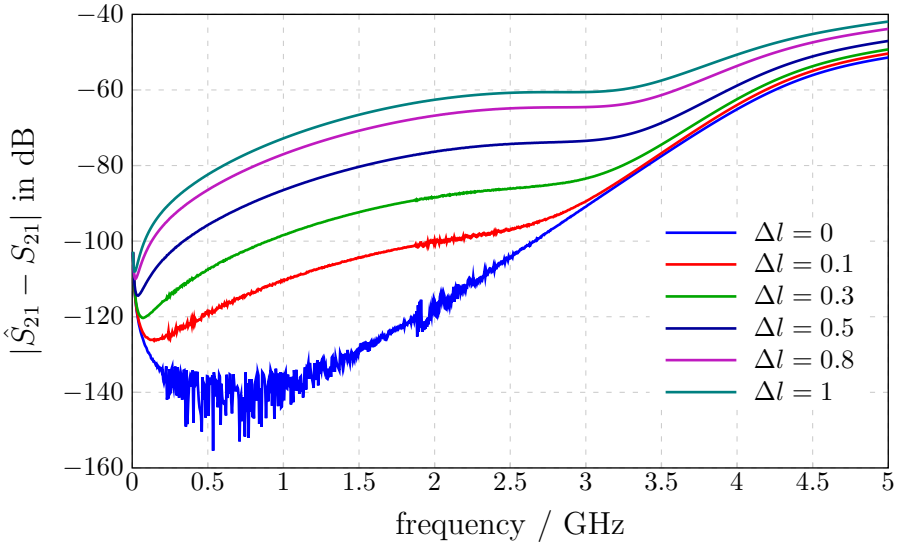


Figure 6.6: *The figure shows the error* $\mathbf{f}_{\text{nsm}} = |\hat{S}_{21} - S_{21}|$ *between the reduced models* $\Sigma_{\boldsymbol{\xi}}$ *corresponding to the indicated lengths and the* MATLAB® *reference solution, i. e. the original model solution. The projection was performed with a common matrix* $\mathbf{V}$ *corresponding to* $l = 8mm$.

with $\mathbf{U}$ an $n \times n$ unitary matrix and $\mathbf{N}$ an $\sum_{i=1}^{N} m_i \times \sum_{i=1}^{N} m_i$ unitary matrix. The matrix $\mathbf{\Sigma}$ contains the singular values of $\mathbf{V}_{\text{all}}$. Choosing for $\mathbf{V}$ the first m columns of $\mathbf{U}$, i. e. :

$$\mathbf{V} = \mathbf{U}(:, 1 : m), \tag{6.5.32}$$

guarantees to capture the m most important directions of $\mathbf{V}_{\text{all}}$. Apparently, besides the number of expansion points, also the number of columns of each $\mathbf{V}_i$, as well as the number of directions m kept in $\mathbf{V}$ can be chosen freely.

This technique has already been used in [31] and [32].

6.5.2.1 Choice of interpolation points

The decisive questions arising at this point are how many interpolation points have to be taken and how they have to be placed in the parameter space. While no established theory exists for answering these questions, for the latter, it is obvious that the points should rather be well distributed in the parameter space than lying close to each other. Pursuing the former question leads to experimenting with parameter sets $\mathcal{S}'_1, \mathcal{S}'_2 \ldots \mathcal{S}'_M$ with different numbers of elements. For efficiency reasons and in order to obtain an estimate of the current approximation error that allows an automated abortion when a desired accuracy is reached, the sets should be hierarchical, for example $\mathcal{S}'_1 \subset \mathcal{S}'_2 \subset \ldots \mathcal{S}'_M$. Obviously, the hierarchical bases introduced in section 6.4 and visualized in figure 6.3 do not fulfill both conditions, as the points are not well distributed in the parameter space.

Sets of points $\mathcal{S}'_i, i = 0 \ldots$ which are both hierarchical, i. e. $\mathcal{S}'_0 \subset \mathcal{S}'_1 \subset \mathcal{S}'_2 \subset \ldots$, and well distributed are the points comprising the multidimensional sparse grids [9].

Sparse grids are a numerical technique, initially investigated by the Russian mathematician Smolyak, which are used to represent, integrate or interpolate high dimensional functions. The points are chosen such that a smooth multi-parameter function can be approximated with a suitable interpolation formula. With the Smolyak method, which still forms the basis of all sparse grid methods, single-parameter interpolation formulas are extended to the multi-parameter case by using tensor products. With increasing problem dimension, this powerful interpolation method requires several orders of magnitude less nodes than conventional interpolation on a full grid. Furthermore, the property of full grid interpolation that the error decays asymptotically with increasing grid resolution is preserved up to a logarithmic factor in the sparse grid approach. The indices $i, i = 0 \ldots$ are called levels of the sparse grids. In [27], the link is given to the Sparse Grid Interpolation Toolbox, which is a MATLAB®toolbox for approximating expensive, possibly high-dimensional multi-parameter functions. Figure 6.7, which visualizes the sparse grids for dimension two and levels $1, 2$ and 3 has been produced with the Sparse Grid Interpolation Toolbox. An important drawback is that even for low levels the number of interpolation points is high. This leads to a large projection matrix $\mathbf{V}$. Therefore, for the purposes of the neighboring-subspace method, the principal point distribution of the sparse grids will be maintained, while occasionally points will be dropped.

6.5.2.2 Variation of the length

In a first instance, the method is analyzed with respect to l. Later on, variations of the width w and depth d will also be considered. The effect of using more than one interpolation points has been analyzed by starting with the variation range center point, and gradually adding points.

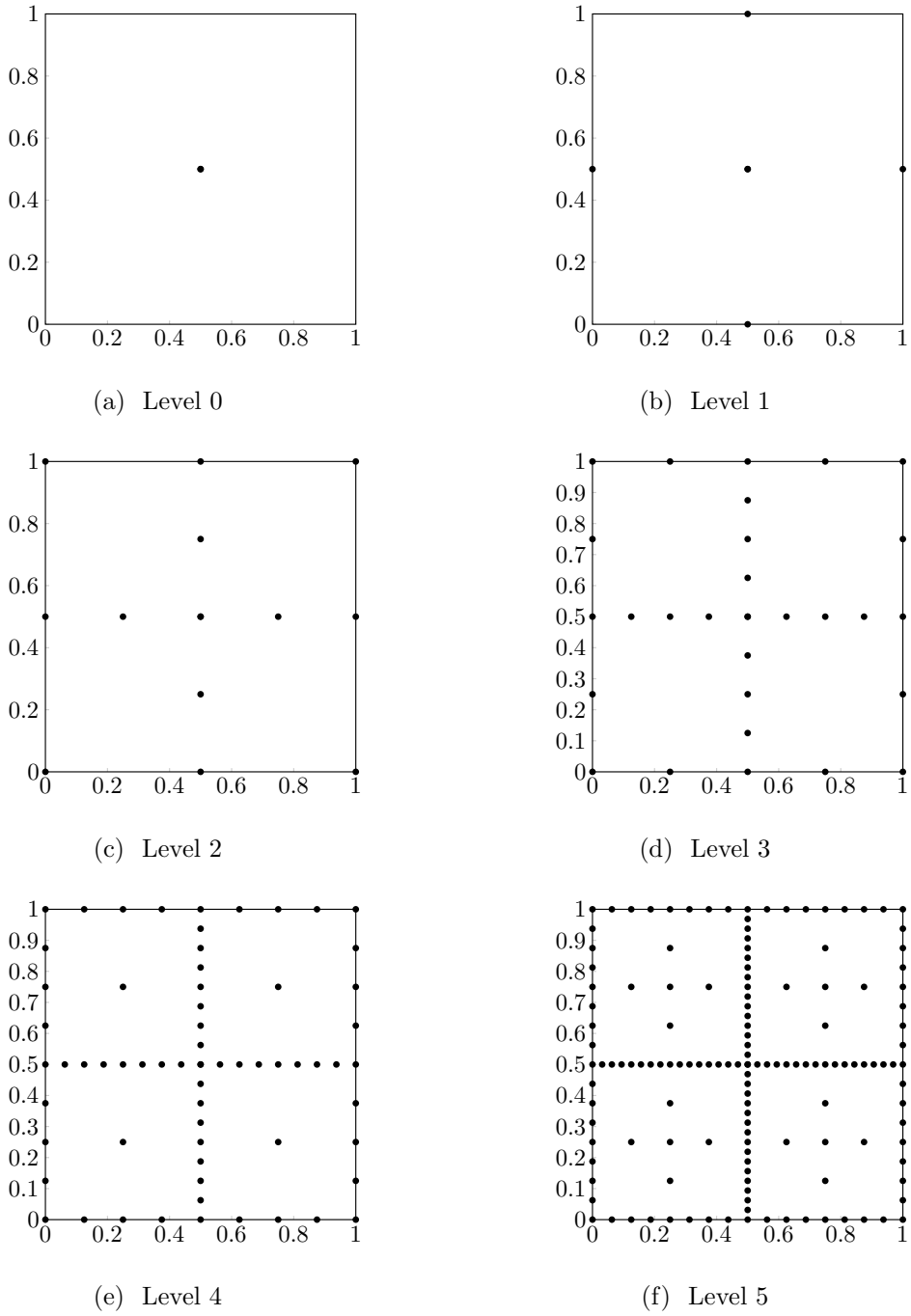
Let q_{A} be the order of the single-parameter Arnoldi method underlying the neighboring-subspace method. Due to the simplicity of the microstrip structure, according to [75, p.80] a choice of q_{A} between $6, 8$ and 10 is sufficient.

Preferably, the singular values of $\mathbf{V}_{\mathrm{all}}$ that are used to set-up the projection matrix $\mathbf{V}$ are given in percent p of the size of $\mathbf{V}_{\mathrm{all}}$. Then, the value m in equation (6.5.32), i. e. the size of the matrix $\mathbf{V}$, is given by: $m = \mathrm{floor}(p \cdot \mathrm{size}(\mathbf{V}_{\mathrm{all}}))$.

Table 6.3 shows the errors $\bar{f}_{s_i,\boldsymbol{\xi}_j}$ and $f_{\max_{s_i,\boldsymbol{\xi}_j}}$, the times t_{red} and $t_{\hat{S}}$, as defined in the previous section, as well as the size m of the respective projection matrix $\mathbf{V}$ for the various settings.

The results summarized in table 6.3 show that $\bar{f}_{s_i,\boldsymbol{\xi}_j}$ and $f_{\max_{s_i,\boldsymbol{\xi}_j}}$ are decreased for increasing numbers of interpolation points. Furthermore, if $p = 0.6$, the resulting $\mathbf{V}$ leads to larger deviations then other settings. On the contrary, the setting of five interpolations points, $p = 0.8$, leads to no accuracy loss. Moreover, no further improvement is denoted by taking more than five interpolation points or $p \geq 0.8$.

The question whether the set of interpolation points affects the error or not has also been pursued. The same settings have been taken as in the previous example. Investigations revealed that up to an almost constant factor the errors were the same. Concluding, it can be stated, that similar results are obtained independently



(a) Level 0

(b) Level 1

(c) Level 2

(d) Level 3

(e) Level 4

(f) Level 5

Figure 6.7: *Visualization of the sparse grids for dimension two and levels* $1, 2 \ldots 5$.

from the exact positions of the interpolation points, as long as more than three interpolation points are used and as long as they are well distributed in the parameter variation range. In the remainder of the work the interpolation points are chosen by using the sparse grids distribution, as described above.

Interpolation	q_A	p	m	$\bar{f}_{s_i,\boldsymbol{\xi}_j}$	$f_{\max_{s_i,\boldsymbol{\xi}_j}}$	t_{red}/s	
one point	6	1	6	0.0491	0.0604	16	
	8	1	8	$5.8{\cdot}10^{-4}$	$9.41{\cdot}10^{-4}$	16	
	10	1	10	$9.8{\cdot}10^{-5}$	$3.25{\cdot}10^{-4}$	16	
three points	6	0.6	10	0.02	0.02	48	
		0.8	14	0.02	0.02	48	
		1	18	0.01	0.05	48	
	8	0.6	14	$2.3{\cdot}10^{-4}$	$4{\cdot}10^{-4}$	48	
		0.8	19	$1.51{\cdot}10^{-4}$	$1.82{\cdot}10^{-4}$	48	
		1	24	$1.7{\cdot}10^{-4}$	$2{\cdot}10^{-4}$	48	
	10	0.6	18	$1.9{\cdot}10^{-4}$	$6.23{\cdot}10^{-4}$	48	
		0.8	24	$9.2{\cdot}10^{-6}$	$6.8{\cdot}10^{-5}$	48	
		1	30	$6.7{\cdot}10^{-6}$	$7{\cdot}10^{-5}$	48	
five points	6	0.6	18	0.05	0.05	80	
		0.8	24	$9.2{\cdot}10^{-3}$	0.01	80	
		1	30	0.012	0.0262	80	
	8	0.6	24	$1.7{\cdot}10^{-4}$	$2{\cdot}10^{-4}$	80	
		0.8	32	$4.68{\cdot}10^{-4}$	$9.28{\cdot}10^{-4}$	80	
		1	40	$4.6{\cdot}10^{-5}$	$8.47{\cdot}10^{-5}$	80	
	10	0.6	30	$6.7{\cdot}10^{-6}$	$7{\cdot}10^{-5}$	80	
		0.8	40	$7.64{\cdot}10^{-6}$	$8.1{\cdot}10^{-5}$	80	
		1	50	$9.39{\cdot}10^{-6}$	$1{\cdot}10^{-5}$	80	

Table 6.3: *Comparison of the effects of different method settings for the length-variable microstrip example.*

6.5.2.3 Variation of the length, the width and the depth

Adding the width w and the depth d, leads to the results shown in table 6.4. In this case, up to a number of 25 interpolation points, as shown in the third row of table 6.4 are used. A large number of interpolation points leads to a large total column size m of the projection matrix $\mathbf{V}$, as shown in table 6.4. This leads to both a larger time t_{red} to calculate $\mathbf{V}$, as well as to a larger time $t_{\hat{S}}$ to reduce the system by projection of $\mathbf{V}$. In general, the following considerations have to be taken into account: The time t_{red} as well as the time $t_{\hat{S}}$ depend on the size of the projection matrix, which in turn is influenced by two significant factors: firstly, if the resonance behavior is simple or complicated, i. e. how many steps q_A are needed in the single-parameter Arnoldi procedure, and secondly, the number of parameters. The size of $\mathbf{V}$ might be reduced by dropping some of the least important singular values. As shown, this alone could save up to 20 % of the matrix size.

The reductions at each point by the single-parameter Arnoldi algorithm can be performed in parallel. Assuming that the time to perform the SVD is negligible

compared to t_{red}, if it is a priori determined how many interpolation points are used, the overall time to calculate $\mathbf{V}$ is approximately the time to reduce one local matrix.

All above mentioned factors lead to different constellations that have to be considered for each problem.

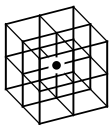
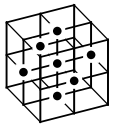
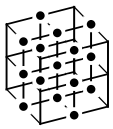
Interp. points	q_{A}	p	m	$\bar{f}_{s_i,\xi_j}$	$f_{\max_{s_i,\xi_j}}$	t_{red}/s	
	6	1	6	$5.8{\cdot}10^{-2}$	$7{\cdot}10^{-2}$	17	
	8	1	8	$7.8{\cdot}10^{-4}$	$1.2{\cdot}10^{-3}$	17	
	10	1	10	$2.4{\cdot}10^{-4}$	$6.6{\cdot}10^{-4}$	17	
 7 points	6	0.8		$5{\cdot}10^{-2}$	$5{\cdot}10^{-2}$	83	
		1	42	$1.6{\cdot}10^{-2}$	$1.7{\cdot}10^{-2}$	83	
	8	0.8		$2.1{\cdot}10^{-4}$	$7.6{\cdot}10^{-4}$	83	
		1	56	$5.6{\cdot}10^{-4}$	$1.4{\cdot}10^{-3}$	83	
	10	0.8		$7.5{\cdot}10^{-7}$	$3.4{\cdot}10^{-6}$	83	
		1	70	$5.3{\cdot}10^{-7}$	$1.6{\cdot}10^{-6}$	83	
 25 points	6	0.8		$1{\cdot}10^{-2}$	$2{\cdot}10^{-2}$	245	
		1	150	$1{\cdot}10^{-2}$	$1{\cdot}10^{-2}$	245	
	8	0.8		$3.3{\cdot}10^{-5}$	$8.6{\cdot}10^{-5}$	245	
		1	200	$2.2{\cdot}10^{-5}$	$4.7{\cdot}10^{-5}$	245	
	10	0.8		$1.9{\cdot}10^{-7}$	$6.3{\cdot}10^{-7}$	245	
		1	250	$1{\cdot}10^{-7}$	$1{\cdot}10^{-7}$	245	

Table 6.4: *Comparison of the effects of different method settings for the microstrip example with variable length, width and depth.*

Figure 6.10 shows the singular values $\sigma_j, j = 1 \dots n_V$ of the merged matrix $\mathbf{V}_{\text{all}}$ for the levels $0, 1$ and 2, corresponding to 1, 3 and 5 interpolation points respectively. With $q_A = 10$, the respective sizes of $\mathbf{V}_{\text{all}}$ are $10, 30$ and 50. For level 0, corresponding to one interpolation point, the singular vectors are of same significance as the corresponding singular values are all 1. For levels 1 or 2, the singular values rapidly decrease. As there is no obvious sudden decay, the truncation is based on empirical values. Omitting the 20% least important singular values has proven to be a good threshold. For the examples shown in the figures for the projection matrices corresponding to level 1 and level 2, this threshold means omitting the values with $j > 24$ and $j > 40$.

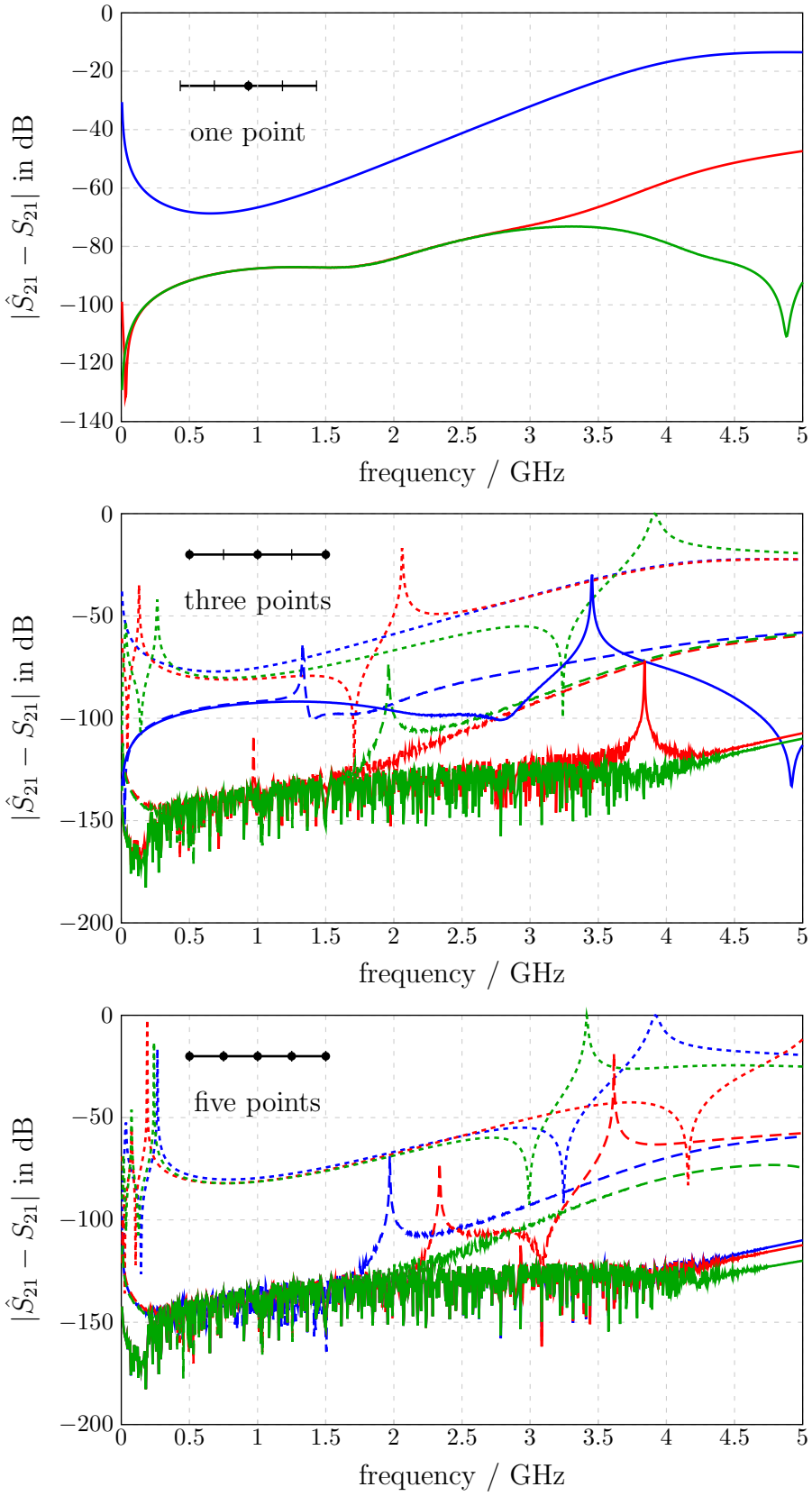


Figure 6.8: *The errors $|\hat{S}_{21} - S_{21}|$ corresponding to the settings of table 6.3 for an arbitrary point in the parameter range.*

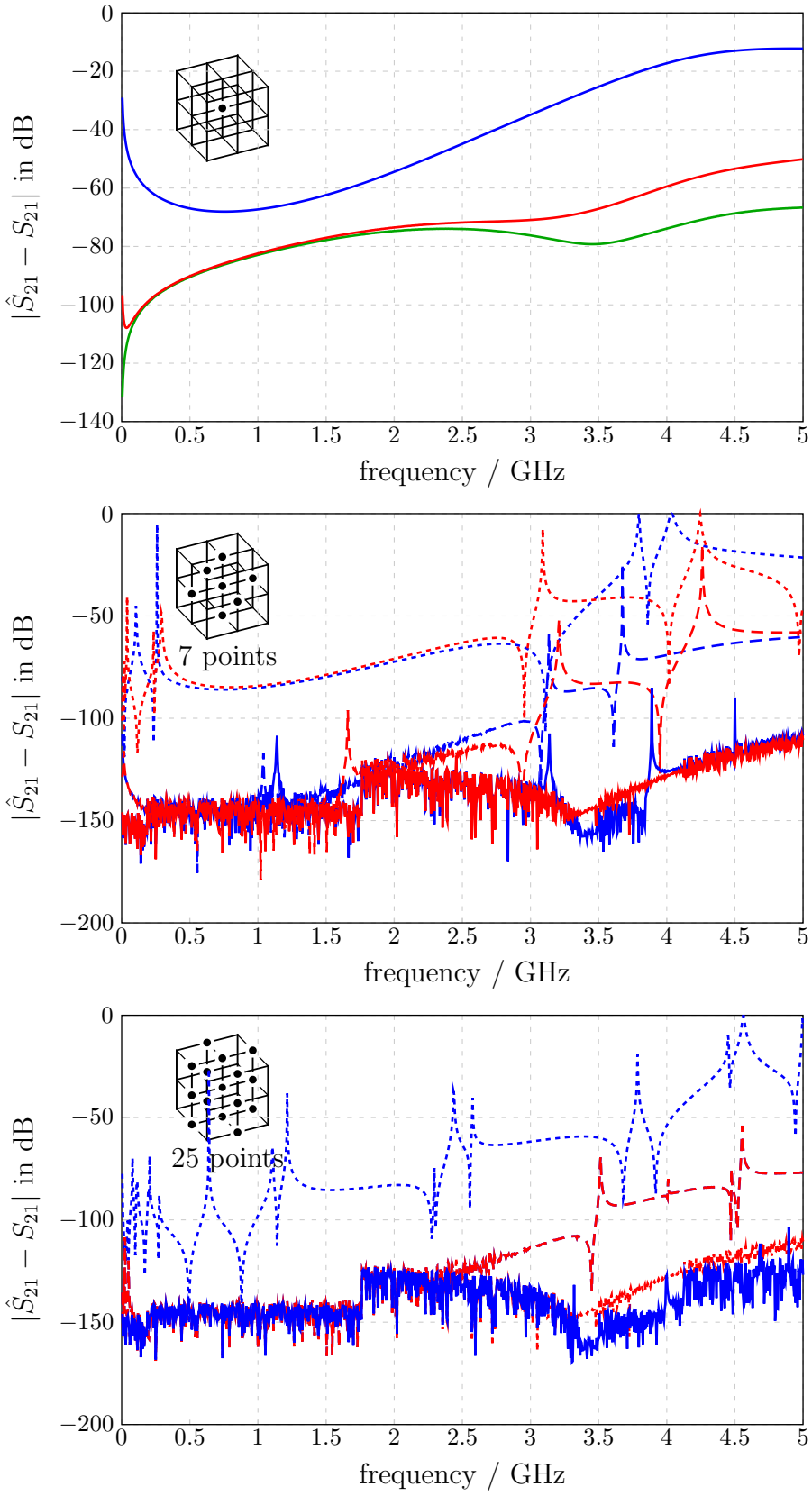


Figure 6.9: *The errors corresponding to the settings of table 6.4 for an arbitrary point in the parameter range.*

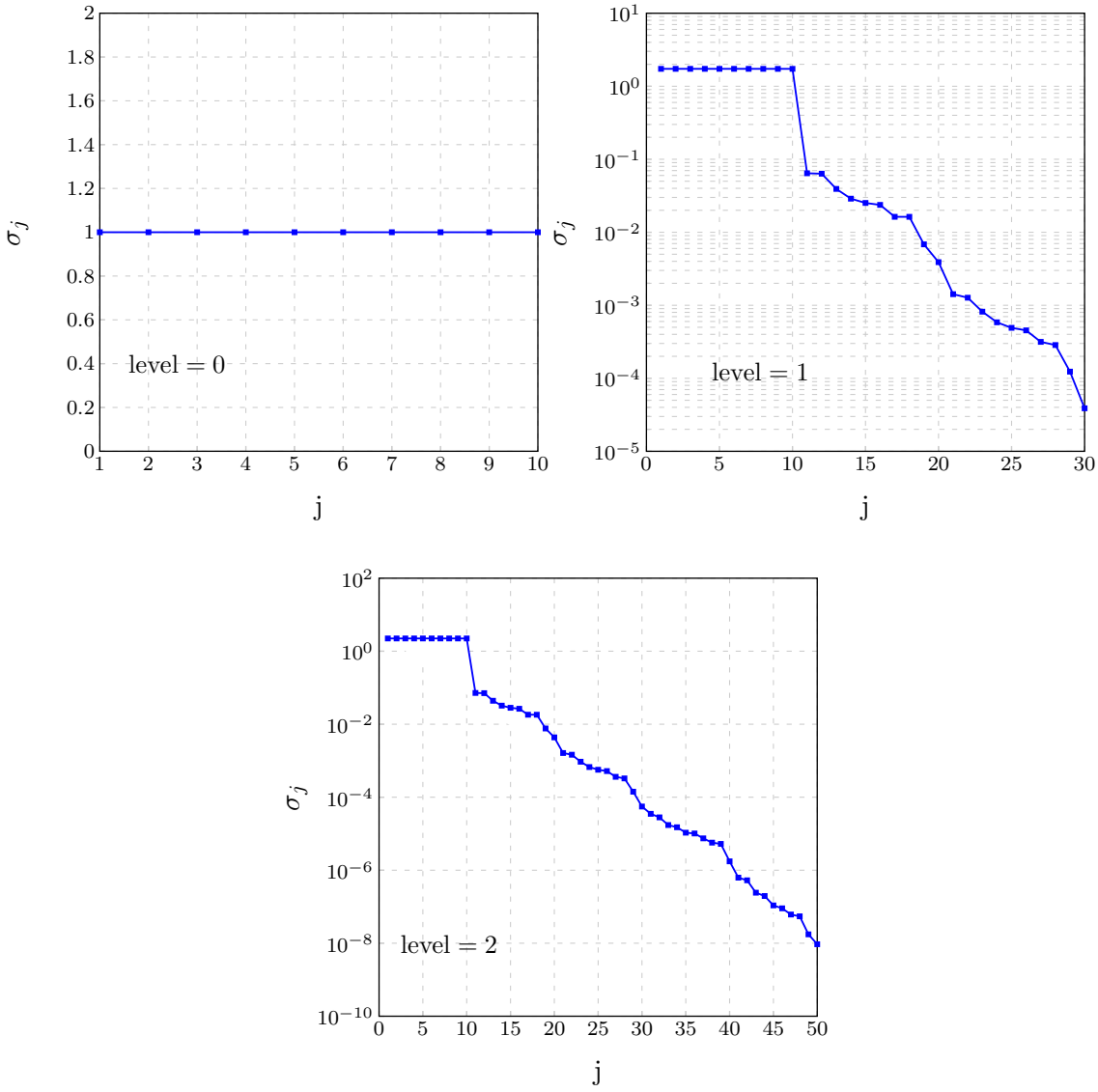


Figure 6.10: *The singular values* $\sigma_j, j = 1 \dots n_V$ *of the merged matrix* $\mathbf{V}_{\text{all}}$ *for the levels* $0, 1$ *and* 2*, corresponding to 1, 3 and 5 interpolation points respectively. With* $q_A = 10$*, the respective sizes of* $\mathbf{V}_{\text{all}}$ *are* $10, 30$ *and* 50*. While for one point all singular values are* 1*, meaning that the respective vectors are of the same importance, for levels 1 or 2, the singular values rapidly decrease.*

6.5.3 Geometry Variation Greater than One Mesh Cell

In the above methodology the row number in each $\mathbf{V}_i$ is constant. This is a necessary requirement for the neighboring-subspace method. Otherwise the merged matrix could not be set-up. Therefore, for the neighboring-subspace method the limitation is not the constant mesh topology, but the size of the system matrices. In general, during a parameter variation the mesh does not remain the same. There are basically two reasons that are responsible for the variation of the system-matrix size.

Most modern simulation software is based on automatic mesh generation, which in turn is based on several assumptions, for instance the geometry or the wave length. Therefore, not only geometrical variations affect the automatic mesh generation, but also material variations. For simple models, as the microstrip example, this can be achieved by manually intervening in the mesh creation, so that the parameter variation occurs on a constant mesh. For complex models this is more difficult.

Nevertheless, even if the mesh is kept fixed, the size of the material matrices can change. In order to be able to invert the diagonal material matrices, all zero diagonal entries are omitted. Entries corresponding to PEC material are also omitted. Let n_1, n_2 be the sizes of the material matrices $\mathbf{D}_{\varepsilon,1}, \mathbf{D}_{\varepsilon,2}$ corresponding to the geometry parameters l_1 and l_2, respectively, depicted in figure 6.11. In case that a dielectric material is varied, the omitted zero elements are the same for all values of the geometry parameter $l_1 \dots l_2$, thus $n_1 = n_2$. In case of a PEC material, as soon as the geometry parameter is larger than the mesh cell size, the material matrix entries corresponding to these cells, in the bottom picture of figure 6.11 they are indicated by the dark grey area, are set to zero. Therefore, in this case it is $n_1 \neq n_2$.

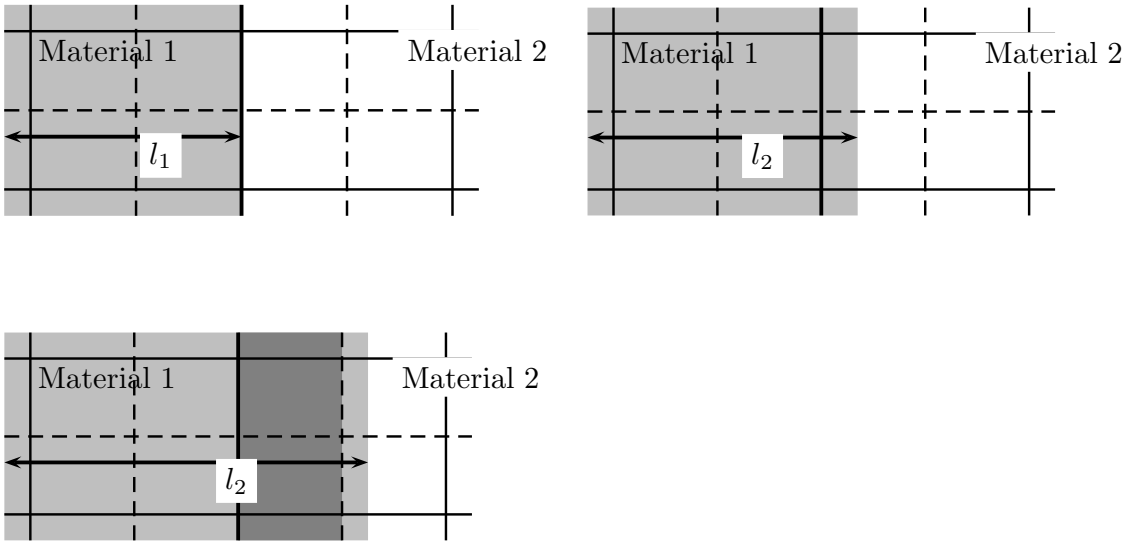


Figure 6.11: *Material 1 with the length l_1 is varied to the length l_2. The top right and bottom pictures show the case where the variation is within the mesh cell and the case where the variation is greater than the mesh cell size, respectively.*

In case of PEC material, the mesh is therefore "frozen" for one parameter. The structure is then discretized on this fixed mesh, for all parameter values e. g. for

the nominal value of the parameter. The variation within the mesh cell size, as indicated in the top right picture of figure 6.11, is possible due the feature of PFC in FIT as described in section 2.2.5. Locally, it acts as a modified mesh cell with the important difference that only one mesh cell is affected with no impact on the remainder of the mesh. This fact can be used for handling geometry variations in multi-parameter MOR. A linearization of the curl-curl equation as in section 4.3.1 is no longer necessary, and, furthermore, even radial geometries can be treated.

In continuation of this thought, in the case of PEC material, an interesting investigation is to determine how large the error is, if for a variation greater than a mesh cell, a common $\mathbf{D}_\varepsilon$ is used for all systems. That is, a fixed mesh and fixed material parameters would be used for all local systems. As long as the variation of the PEC material (other material do not produce zeros in the material matrices), affect only a few mesh cells, it is expected that the resulting projection matrices are close to the original ones. With a small error the restriction that the variation has to be within a mesh cell is relaxed. This issue is investigated in [60] and is an important advantage of the neighboring-subspace method compared to the other approaches.

Algorithm 6.3 summarizes the neighboring-subspace method introduced in this chapter.

Algorithm 6.3 Neighboring-subspace method

$\mathbf{V} = [\,]$
for level $= 0 \ldots \text{level}_{\max}$ **do**
  $\mathcal{S}' = \text{spgrid(level,r)}$ % *calculates the set of points corresponding to the sparse grids*
  for $ind = 0 \ldots \text{size}(\mathcal{S}', 1)$ **do**
    $\mathbf{p_v} = \mathcal{S}'(ind)$
    update CST MWS models, calculate material matrices and results
    import CST MWS matrices
    calculate system matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$
    $\mathbf{V}_p = \text{block arnoldi}(\mathbf{A}, \mathbf{B}, \mathbf{C}, q_A)$ % *consider only one* q_A
    $\mathbf{V}_{\text{all}} = [\mathbf{V}_{\text{all}} \quad \mathbf{V}_p]$
  end for
  $\mathbf{U} = \text{SVD}(\mathbf{V}_{\text{all}})$
  $\mathbf{V} = \mathbf{U}(:, 1:m)$ % *keep* m *largest singular vectors*
end for

6.6 Concluding Remarks

Based on the discrete equations resulting from the FIT, in this chapter three methods were presented to approximate the transfer function of the respective systems by lower dimensional systems, while paying attention on the parameter dependence. The parameter dependence should be maintained in the reduced model.

The various multi-parameter system representations derived in chapter 4 form the basis for the order reduction by means of a general projection framework. For the reduced model generation basically two principally different approaches are possible. Either the respective systems are reduced by means of higher-order Krylov subspaces which match the moments of the original and the reduced model, or the parametric reduced system is obtained from information resulting from the reduction at specified samples in the parameter range.

The iterative formulation to determine the moments of the multi-parameter system is a multi-level recursion. In each step not only the last moment is used, but a number of previously calculated moments. The numerical cost of this procedure is therefore increased compared to the single-parameter case.

The direct iterative calculation is numerically instable, similar to the classical case. Therefore, efficient algorithms are required in order to determine stable bases spanning the higher-order Krylov subspaces. This not trivial procedure is fulfilled by both the multi-parameter Arnoldi algorithm as well as the contraction method presented in this chapter. While the multi-parameter Arnoldi is more straight-forward than the contraction method, it requires the storage of large matrix parts.

Based on a tricky reformulation of the multi-parameter moment iteration, the contraction method reduces the determination of the higher-order Krylov subspace on the calculation of Krylov subspaces corresponding to local systems, more precisely to polynomial systems obtained for one parameter sample. The determination of the respective Krylov subspaces is performed by means of the classical WCAWE algorithm, which was presented in the previous chapter. Although being a reliable method that can reach high orders of accuracy, the contraction method has two main drawbacks. Being based on multiple evaluations of the WCAWE algorithm its time-cost is high even for small systems. Furthermore, it requires a constant mesh topology throughout the variation. Material parameters are not affected by this restriction, geometry parameters though are difficult to handle in practical applications with complex structures. In these cases, it is often not possible to guarantee a constant mesh topology.

The trend is therefore to methods that are mesh-independent or at least that do not pose the high restriction of a constant mesh topology. The neighboring-subspace method presented in this chapter for instance requires only a constant mesh. The projection matrix of the parametric system is defined as the composition of local projection matrices. Moment matching is in general not guaranteed for these methods. Its main advantage is that it can rapidly calculate the projection matrix for the parametric system, with the cost of a single-parameter system. The obtained accuracy can reach the accuracy of single parameter systems. In case of only one additional parameter besides the frequency even one or two local projection matrices are sufficient. A larger number of local projection matrices is required in case of more parameters, nevertheless the Krylov subspace determination of all local models can be performed in parallel, so that no additional time cost arises.

Chapter 7

Numerical Examples

So far, the microstrip example in the previous chapter has served to illustrate the presented PMOR methods. They have been designed to be accurate over a specified parameter space of interest, but no considerations have been performed concerning computational efficiency and practical applicability. In order to allow for such considerations the methods have to be applied to large and complex problems. On the basis of their performance practical questions arise. To this purpose, two filter examples of higher complexity than the microstrip line are examined namely a narrow-band filter and the Langer filter. Furthermore, a comparison of the PMOR methods is given.

7.1 General Considerations

7.1.1 Reference Solution

In order to make statements about the quality of the PMOR methods, for instance how well the S-parameters S of $\boldsymbol{\Sigma}$ are approximated by the S-parameters $\hat{S}$ of $\hat{\boldsymbol{\Sigma}}$, a reference solution is required. The reference solution is the analytical solution in case that it exists. As systems having an analytic solution are the exception, alternatively one reverts on taking a sufficient good approximation which is calculated by means of software simulation packages. In this work the CST MWS has been used.

Mainly, two sources of numerical inaccuracies affect the solution modules of simulation packages as the CST MWS. These are numerical errors caused by the finite digit number representation and inaccuracies due to the finite mesh resolution. Calculations in the frequency domain are additionally affected by the accuracy of the iterative linear equation system solvers. This value has been set to 10^{-9} throughout this chapter. Errors stemming from the finite mesh resolution are reduced by increasing the mesh resolution. Thus, one would calculate a solution to the structure by using the highest possible mesh $\mathcal{G}_{\text{ref}}$. The S-parameters corresponding to this mesh are defined as the reference S^{ref}. Obviously, the resulting systems can be very large. In practical applications, the approach is other way round: starting from a

coarse mesh, the mesh resolution is successively refined and the S-parameters are recalculated. When there is no longer a significant change in the results for increasing mesh density, convergence has been reached. Usually, the resulting mesh $\mathcal{G}$ is much coarser then $\mathcal{G}_{\text{ref}}$. In general, accurate S-parameter results for filter structures can only be obtained by mesh convergence studies, which can be performed either by manually changing the mesh settings or by making use of the automatic mesh adaptation tool. Obviously, the latter is the easiest way for this purpose. The final S-parameters generally differ significantly from the results obtained with the initial mesh settings. In this work, the following approach is chosen: a very fine mesh $\mathcal{G}_{\text{ref}}$ is set as the reference, and by means of error definitions given in the following, a coarser mesh $\mathcal{G}$ for the calculations is defined.

The FIT matrices obtained from the MWS calculation, which contain the mesh data, the structure information and the port signals can be imported in mathematical calculation packages. Depending on the system size the relevant system matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ can be determined. In this work the MATLAB® package has been chosen, nevertheless many other options exist. Given the system matrices, the original transfer function $\mathbf{H}$ and the S-parameters S can be then directly calculated and be compared with the reduced values. The microstrip example of the previous chapter has been treated in this way.

In case of large and complex structures where no analytical solution can be obtained, the systems resulting from the mesh $\mathcal{G}$ are too large for their S-parameters to be computed directly in MATLAB®. One option consists in calculating the S-parameters in MWS, S_{MWS}, and importing them in MATLAB®. This procedure has two drawbacks. Firstly, the representation accuracy of values imported in MATLAB® from CST MWS is 10^{-6}, therefore the quality of the PMOR methods can be traced back up to this accuracy. Additionally, $\hat{S}$ and S_{MWS} are calculated by different tools, so that a comparison of their computing times is questionable.

In the description given above, as well as in this work so far, the reduced S-parameters $\hat{S}$ are compared with the S-parameters S of the system from which they resulted. This is the usual policy in the MOR community and goes along with the problem set-up related to system-approximation described in section 1.3.1. Alternatively, one can detach from the idea that $\hat{S}$ should only be compared with S. In fact, both $\hat{S}$ and S can be compared with S^{ref}. The description of this approach is given in the subsequent section.

7.1.2 Fast Parameter Sweeps

At this point it is required to take one step back and think of what is aimed to achieve by using PMOR methods. One important practical application of PMOR methods are simulations with varying parameter values. In this case, the parametric reduced order model needs to be generated only once and can be used for instance for fast parameter-sweeps. One typical application is filter tuning, where a model needs to be evaluated many times in order to obtain the desired setting. Obviously, the time required to set-up the parametric reduced order model should be smaller

than a parameter-sweep of the original model. Time-saving is more important than high solution-accuracies.

Let $\mathcal{G}$ be a computational mesh and S be the corresponding S-parameters. In the following, S will denote the values S_{MWS} imported from MWS. Furthermore, let $F_{S,\text{ref}}$ denote the error of S with respect to S^{ref}. In practical applications usually $F_{S,\text{ref}} \approx 10^{-2}$. From a practical viewpoint it is therefore aimed to achieve reduced S-parameters with approximately the same accuracy with respect to S^{ref} in less computation time. As the error of the reduced values $\hat{S}$ obtained from the same system $\mathcal{G}$ with respect to S^{ref} is larger than the error between S and S^{ref}, eventually, a mesh $\mathcal{G}'$ that is finer than $\mathcal{G}$ should be used to obtain the reduced S-parameters. The S-parameters obtained from the mesh $\mathcal{G}'$ will be denoted by $\hat{S}'$. Let $F_{\hat{S}',\text{ref}}$ denote the error of $\hat{S}'$ with respect to S^{ref}. In practice it is therefore aimed that $F_{\hat{S}',\text{ref}} \approx F_{S,\text{ref}}$. Obviously, high MOR accuracies require unnecessarily larger computation times without practical improvement of $F_{\hat{S}',\text{ref}}$.

Assuming that a sweep with m values is performed, it is desired that

$$m \cdot t_S > m \cdot t_{S'} + t_{\text{red}}, \tag{7.1.1}$$

where t_S is the time needed to calculate the S-parameters in MWS, t_{red} is the time needed to reduce the system $\mathbf{\Sigma}'$ to $\hat{\mathbf{\Sigma}}'$ and $t_{S'}$ is the time to calculate the S-parameters from the respective system matrices. Again, it should be kept in mind that S and $\hat{S}'$ are calculated by different tools, namely by MWS and MATLAB®, respectively. The comparison of the computation times is therefore not completely fair.

While $t_{S'} \ll t_S$, usually t_{red} is quite large, and therefore for small m the inequality in (7.1.1) does not hold. This is in agreement with the fact, that PMOR methods do not make sense for one or two system evaluations, but for a parameter sweep, i. e. $m \geq 2$. Basically, a value m_0 exists, such that $m \cdot t_S > m \cdot t_{S'} + t_{\text{red}}$, $m \geq m_0$.

Concluding, in practical applications there is a redefinition of the problem set-up from section 1.3.1 as following: Consider a structure described by the system $\mathbf{\Sigma}_{\text{ref}}$, and $\mathbf{\Sigma}$, which result from the Maxwell grid equations for the meshes $\mathcal{G}_{\text{ref}}$ and $\mathcal{G}$, as described above, where $\mathbf{x} \in \mathbb{R}^n$. Determine a system $\hat{\mathbf{\Sigma}}'$ resulting from a mesh $\mathcal{G}'$ in the same form as $\mathbf{\Sigma}$ with $\hat{\mathbf{x}} \in \mathbb{R}^m$ with $m \ll n$, such that the approximation error between the $\hat{\mathbf{\Sigma}}'$ and $\mathbf{\Sigma}_{\text{ref}}$ is in the same order as the error between $\mathbf{\Sigma}$ and $\mathbf{\Sigma}_{\text{ref}}$, while m_0 in $m \cdot t_S > m \cdot t_{S'} + t_{\text{red}}$, $m \geq m_0$ is as small as possible.

7.2 The Narrow-Band Filter

The narrow-band filter shown in figure 7.1 consists of two perfectly conducting cylindrical resonators in a rectangular cavity. The resonators are coupled by a rectangular iris. Two coaxial ports are coupled to the device by extending the inner conductor of the coaxial cables into the cavity. The detailed description of the device including geometry specifications can be found in [14].

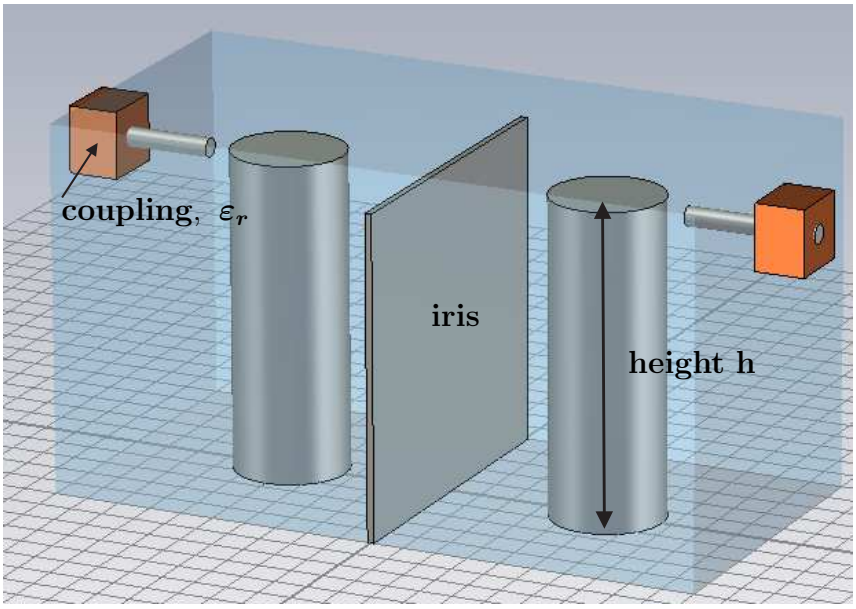


Figure 7.1: *The variable parameters for the narrow-band filter are the frequency, the permittivity of the coupling ε_r and the cylinder height h.*

Narrow-band filter parameter range	
frequency s	$0.58 \ldots 0.63$ GHz
permittivity ε_r	$1.2 \ldots 2.2$
permittivity ε_r	$2 \ldots 8$
height h	$96 \ldots 98$ mm

The frequency varies in the range of $0.58 \ldots 0.63$ GHz. The filter will be analyzed with respect to variations of the permittivity of the coupling ε_r, both for a small and a large range, and the height of the cylinders h.

7.2.1 Permittivity Variation

The reference S-parameters S^{ref} for the permittivity variation have been calculated with a mesh of 10.801.560 mesh cells with the "General Purpose Frequency Solver" in the CST Studio Suite and their magnitudes are shown in figure 7.2. In the complex plane S_{21} is shown in figure 7.3.

In order to make a choice for $\mathcal{G}$ and $\mathcal{G}'$, some further considerations concerning the error with respect to S^{ref} have to be taken into account. Besides the definitions given in the previous chapter, additional error definitions are necessary. In particular, let $s_i, i = 1 \ldots n_s$ denote the frequency samples. Throughout this work $n_s = 1000$ samples have been evaluated. Furthermore, let $\boldsymbol{\xi}_j, j = 1 \ldots n_{\boldsymbol{\xi}}$ be the parameter samples (excluding the frequency). The following error-norm is defined:

$$\max_{s_i}\{\min_{s_i}\{|S(s_i, \boldsymbol{\xi}_j) - S^{\text{ref}}(s_i, \boldsymbol{\xi}_j)|\}\}. \tag{7.2.2}$$

For convenience the notation $|S - S^{\text{ref}}|_{\text{maxmin}}$ meaning the above error norm will be used.

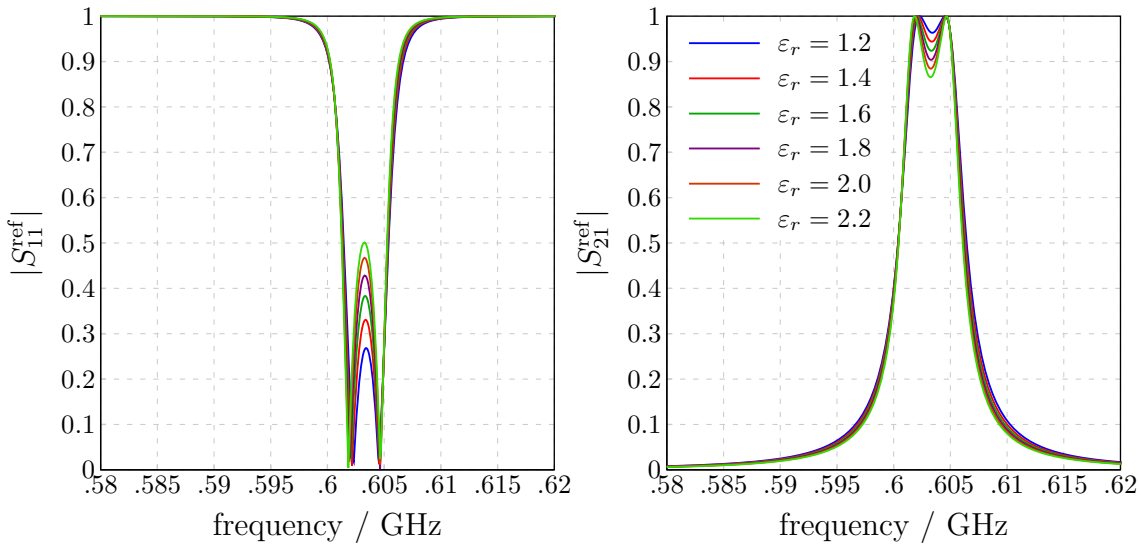


Figure 7.2: *The figure shows the magnitudes of the reference S-parameters for the permittivity variation of the capacitive coupling in the range* $\varepsilon_r = 1.2 \ldots 2.2$.

The same definitions hold for the comparison of $\hat{S}$ and S^{ref}, as well as of $\hat{S}$ and S:

$$\max_{s_i}\{\min_{s_i}\{|\hat{S}(s_i, \boldsymbol{\xi}_j) - S^{\text{ref}}(s_i, \boldsymbol{\xi}_j)|\}\}, \tag{7.2.3a}$$

$$\max_{s_i}\{\min_{s_i}\{|\hat{S}(s_i, \boldsymbol{\xi}_j) - \mathrm{S}(s_i, \boldsymbol{\xi}_j)|\}\}, \tag{7.2.3b}$$

with the abbreviating notations $|\hat{S}\text{–}S^{\text{ref}}|_{\text{maxmin}}$ and $|\hat{S}\text{–}S|_{\text{maxmin}}$, respectively.

The reason to consider $\min_{s_i}\{|S(s_i, \boldsymbol{\xi}_j) - S^{\text{ref}}(s_i, \boldsymbol{\xi}_j)|\}$ additionally to $|S - S^{\text{ref}}|$ is that in most cases the frequency plots S, $\hat{S}$ and S^{ref} in the complex plane have basically the same characteristics, featuring just a small shift.

A mesh convergence analysis gives the error $|S\text{–}S^{\text{ref}}|_{\text{maxmin}}$ for different meshes that successively get finer. Table 7.1 summarizes the errors $|S\text{–}S^{\text{ref}}|_{\text{maxmin}}$ for different meshes at an arbitrary point in the parameter space. Here, the system corresponding to $\varepsilon_r = 1.2$ is representative for the other parameters.

meshcells	$\lvert S_{11}\text{–}S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{21}\text{–}S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	t_S
19.296	0.067215	0.045213	174 s=2 m 54 s
86.961	0.050214	0.031972	186 s=3 m 6 s
212.976	0.027157	0.03354	701 s=11 m 41 s
409.360	0.028050	0.03158	900 s=15 m

Table 7.1: *Mesh convergence analysis for the narrow-band filter.*

Next, the error stemming from the order reduction is determined. To this purpose, the mesh with 212.976 cells is chosen, as it is a good tradeoff between system size and achieved error norm.

As a permittivity variation does not cause any mesh changes, the requirements for the contraction method are fulfilled. An initiated analysis though showed a large time-cost for one evaluation of the WCAWE algorithm. Considering that a high

number of WCAWE evaluations is required (at least order 6 is required), the overall calculation time is prohibitively large. Thus the contraction method proves to be not suitable for large practical applications. The reduction is therefore performed with the neighboring-subspace method throughout this chapter.

Investigations have shown that the MOR error is equally distributed in the parameter range of ε_r. It is therefore sufficient to observe the error at few parameter samples in the parameter range. Table 7.2 shows the errors $\bar{f}_{s,\varepsilon}$, $f_{\max_{s,\varepsilon}}$ and $|\hat{S}{-}S|_{\text{maxmin}}$, for different levels of the neighboring-subspace method. The table also gives the time t_{red} required to set-up the projection matrices $\mathbf{V}$ and the time $t_{\hat{S}}$ to calculate the reduced $\hat{S}$ from the reduced system matrices. In each field of the table two columns are given, the first refers to S_{11} and the second to the S_{21}. Obviously, in practice, the errors $f_{\max_{s,\varepsilon}}$ and $|\hat{S}{-}S|_{\text{maxmin}}$ are the same.

q	m	$\bar{f}_{s,\varepsilon}^{S_{11},\hat{S}_{11}}$	$\bar{f}_{s,\varepsilon}^{S_{21},\hat{S}_{21}}$	$f_{\max_{s,\varepsilon}}^{S_{11},\hat{S}_{11}}$	$f_{\max_{s,\varepsilon}}^{S_{21},\hat{S}_{21}}$	$\lvert\hat{S}_{11}{-}S_{11}\rvert_{\text{min}}^{\text{max}}$	$\lvert\hat{S}_{21}{-}S_{21}\rvert_{\text{min}}^{\text{max}}$	$t_{\hat{S}}/\text{s}$
0	10	$1.9{\cdot}10^{-4}$	$2.0{\cdot}10^{-4}$	$4.0{\cdot}10^{-3}$	$4.2{\cdot}10^{-3}$	$4.0{\cdot}10^{-3}$	$4.2{\cdot}10^{-3}$	0.5
1	24	$9.4{\cdot}10^{-5}$	$8.9{\cdot}10^{-5}$	$1.7{\cdot}10^{-3}$	$1.7{\cdot}10^{-3}$	$1.7{\cdot}10^{-3}$	$1.7{\cdot}10^{-3}$	1.2
2	40	$3.8{\cdot}10^{-5}$	$3.2{\cdot}10^{-5}$	$5.0{\cdot}10^{-4}$	$6.0{\cdot}10^{-4}$	$5.0{\cdot}10^{-4}$	$6.0{\cdot}10^{-4}$	3.1
3	72	$5.8{\cdot}10^{-5}$	$5.1{\cdot}10^{-5}$	$3.0{\cdot}10^{-3}$	$1.0{\cdot}10^{-3}$	$3.0{\cdot}10^{-3}$	$1.0{\cdot}10^{-3}$	4.1

Table 7.2: *Error stemming from the order reduction for different levels of the neighboring-subspace method. The number of points corresponding to the levels* $q = 0, 1, 2, 3$ *are* $1, 3, 5$ *and* 9*, respectively. The size of the reduction matrix* $\mathbf{V}$ *is given by* m*.*

Concluding from table 7.2 and by comparing with the results of table 7.1, as both $f_{\max_{s,\varepsilon}}$ and $|\hat{S}{-}S^{\text{ref}}|_{\text{maxmin}} < 10^{-3}$, it is sufficient to choose level 0 or at most 1. This is in agreement with the experience gained from the microstrip in the previous chapter. For variations with one parameter (additional to the frequency) it proved to be sufficient to use one or three interpolation points. Using more than 5 interpolation points did not further improve the results. Furthermore, the matrix obtained from the SVD of $\mathbf{V}_{\text{all}}$ is truncated for m such that $m = 0.8 \cdot \text{size}(\mathbf{V}_{\text{all}})$.

As the reductions can be performed in parallel, the times $t_{\text{red}} = 3865\,\text{s}$ are approximately the same for all levels. Nevertheless, even for $q = 0$, i. e. one interpolation point, the resulting error is still in the same range as the error $|S{-}S^{\text{ref}}|_{\text{maxmin}}$ of the original solution. In addition, it is not necessary to take different meshes $\mathcal{G}, \mathcal{G}'$, as the error convergence apparent from the values in table 7.1 is slow and the MOR error is smaller than the change between two subsequently chosen meshes. Table 7.3 shows the errors $|\hat{S}{-}S^{\text{ref}}|_{\text{maxmin}}$ and $|S{-}S^{\text{ref}}|_{\text{maxmin}}$ for six samples.

Finally, figure 7.3 depicts the S-parameter S corresponding to the chosen mesh and the reference S-parameter S^{ref} in the complex plane. Figure 7.4 depicts the MOR error compared to the errors of the reduced and the original model with respect to the reference solution.

With the time $t_{\text{red}} = 3865\,\text{s}$ and considering that the time to calculate the S-parameters in MWS is $t_S = 701\,\text{s}$, a parameter sweep by using the PMOR method makes sense for $m_0 = 3865\,\text{s}/701\,\text{s} = 5.5 \approx 6$ parameter values.

ε_r	$\lvert\hat{S}_{11}-S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert\hat{S}_{21}-S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{11}-S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{21}-S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$
1.2	0.02784	0.02376	0.02772	0.02427
1.4	0.03052	0.02613	0.03064	0.02665
1.6	0.03181	0.02833	0.03196	0.02885
1.8	0.03398	0.03039	0.03378	0.03092
2.0	0.03622	0.03233	0.03611	0.03286
2.2	0.03743	0.03417	0.03763	0.03471

Table 7.3: *Error comparison for six permittivity samples in the variation range between* $\varepsilon_r = 1.2 \ldots 2.2$.

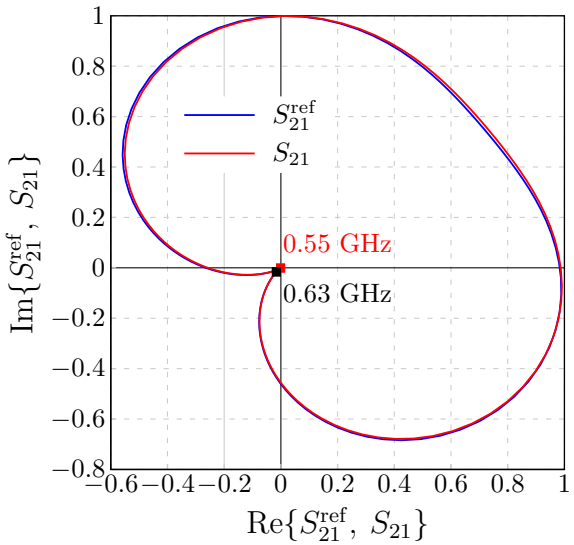


Figure 7.3: *The S-parameter* S_{21} *corresponding to the mesh* G *and the reference S-parameter* S_{21}^{ref} *in the complex plane. The reduced* $\hat{S}_{21}$ *has not been drawn as in this view its curve is indistinguishable from* S_{21}.

7.2.1.1 Wide-band permittivity variation

In this section the observed variation range will be enlarged to $\varepsilon_r = 2 \ldots 8$. Again, reference S-parameters have been calculated which are shown in figure 7.5. The mesh $\mathcal{G}$ used for this purpose was the same as in the previous setting for the small variation range.

Nevertheless, due to the higher permittivity values for the choice of the working mesh $\mathcal{G}$ a finer mesh is required. More precisely, using the mesh of the previous setting leads to the error values in the second column of table 7.4 implying that a higher grid resolution is required. It is important to notice, that the errors for the reduced system (first column of table 7.4) are approximately the same as the errors of the original system. Obviously, it is not the MOR which causes the largest error, and a higher value of q would not decrease the error compared to the reference solution. The results of table 7.4 clearly demonstrate that the MOR is as good as the system it stems from, always with respect to the reference solution. Observing

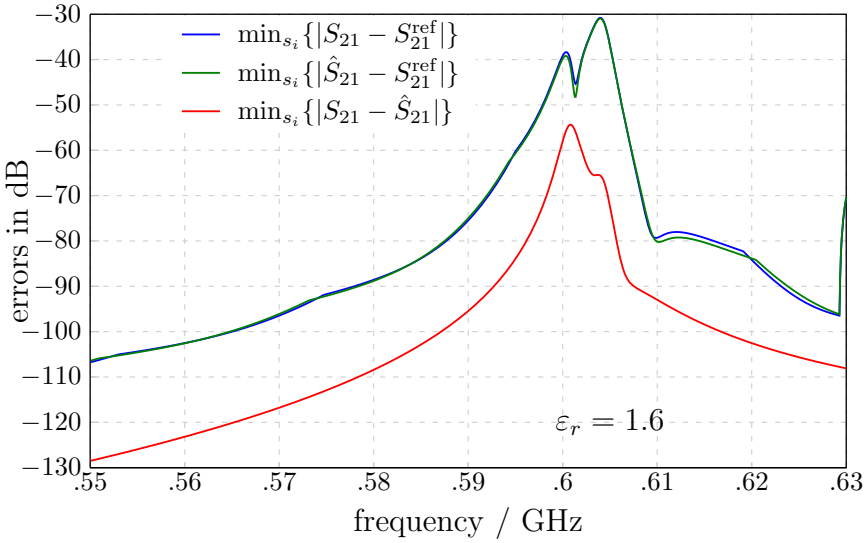


Figure 7.4: *As the error curves are similar for the whole parameter range, it is sufficient to observe it for one sample, here* $\varepsilon_r = 1.6$. *Obviously, the MOR error* $\min_{s_i}\{|S_{21}(s,\boldsymbol{\xi}) - \hat{S}_{21}(s,\boldsymbol{\xi})|\}$ *is not significant for the errors compared to the reference solution.*

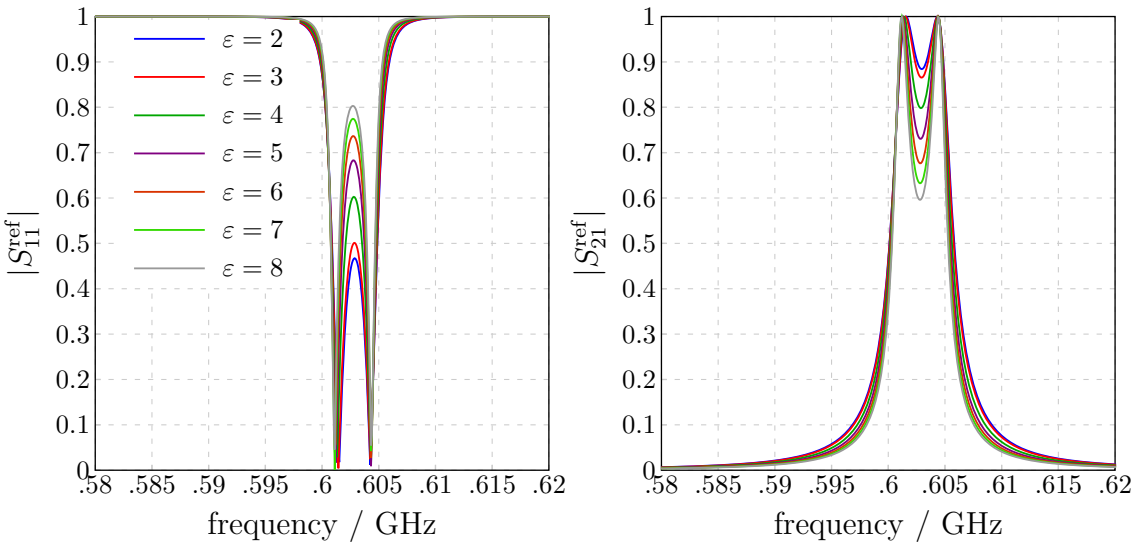


Figure 7.5: *Magnitude of the reference S-parameters for the permittivity variation between* $\varepsilon_r = 2$ *and* 8.

the errors of a finer mesh with 409.360 mesh cells, leads to results in the order 10^{-2}, similar to those for the small variation range. Concluding, for a wider range, as in this example, it is rather advised to refine the mesh, in order to capture the parameter variation requirements, than to increase the level q. Increasing q does not improve the results.

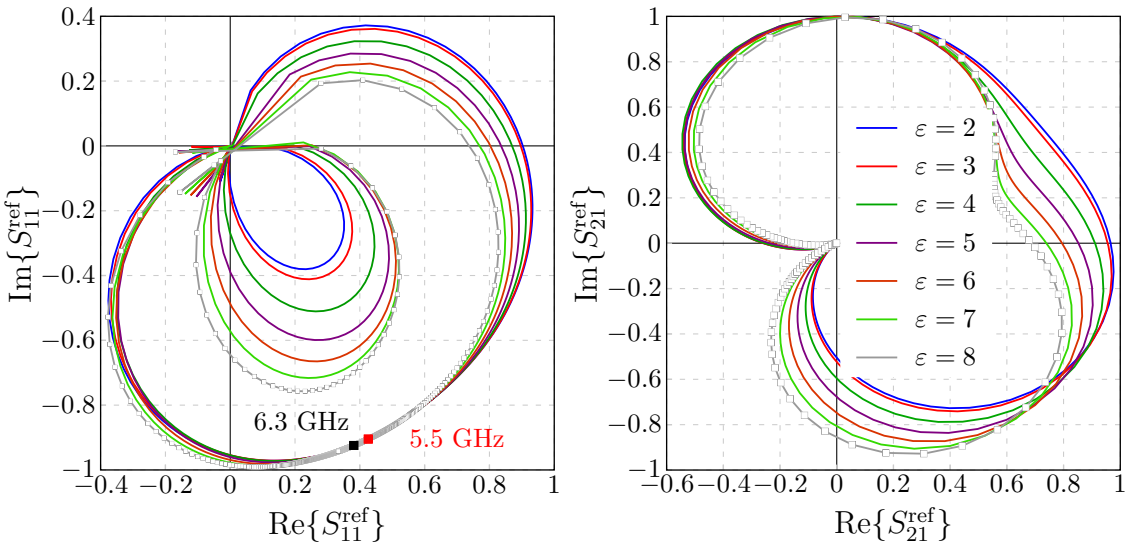


Figure 7.6: *The reference S-parameters for the permittivity variation between* $\varepsilon_r = 2$ *and* 8 *in the complex plane.*

ε_r	$\lvert\hat{S}_{11}-S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert\hat{S}_{21}-S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{11}-S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{21}-S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$
2	0.08076	0.04465	0.07773	0.04321
3	0.10856	0.04050	0.10820	0.04154
4	0.12193	0.04131	0.12060	0.04254
5	0.12704	0.04614	0.12479	0.04052
6	0.12952	0.05809	0.13158	0.05681
7	0.13370	0.06313	0.13370	0.06577
8	0.14067	0.06456	0.13904	0.06821

Table 7.4: *Comparison of* $|\hat{S} - S^{\text{ref}}|_{\text{maxmin}}$ *and* $|S - S^{\text{ref}}|_{\text{maxmin}}$ *for seven samples in the variation range. As in the previous setting with the small variation range one interpolation point is sufficient for the reduction. More interpolation points unnecessarily increase the size of the projection matrix.*

7.2.2 Variation of the Cylinder Height

In order to investigate variations of the cylinder height a mesh consisting of 216.920 cells has been chosen. The choice resulted again from a mesh convergence study. The permittivity has been set to $\varepsilon_r = 2.2$ throughout the variation of the height. The time required for the calculation with the "General Purpose Frequency Solver" with a solver accuracy of 10^{-9} in MWS is $t_S = 565\,\text{s} \approx 9\,\text{m}\,25\,\text{s}$. The reference curves shown in figure 7.7 have been obtained with a mesh consisting of 11.793.540 cells. They have also been calculated with the "General Purpose Frequency Solver" with a solver accuracy of 10^{-9}.

A problem arises in the area of the iris. The specified value of its height lies within the variation of the cylinder height. In this case, a common $\mathbf{D}_\varepsilon$ was used for all systems, i. e. for some parameter values, to few mesh cells a wrong permittivity

value was assigned in order to keep the mesh fixed, so that the neighboring-subspace method can be applied. Obviously, it did not affect the quality of the MOR.

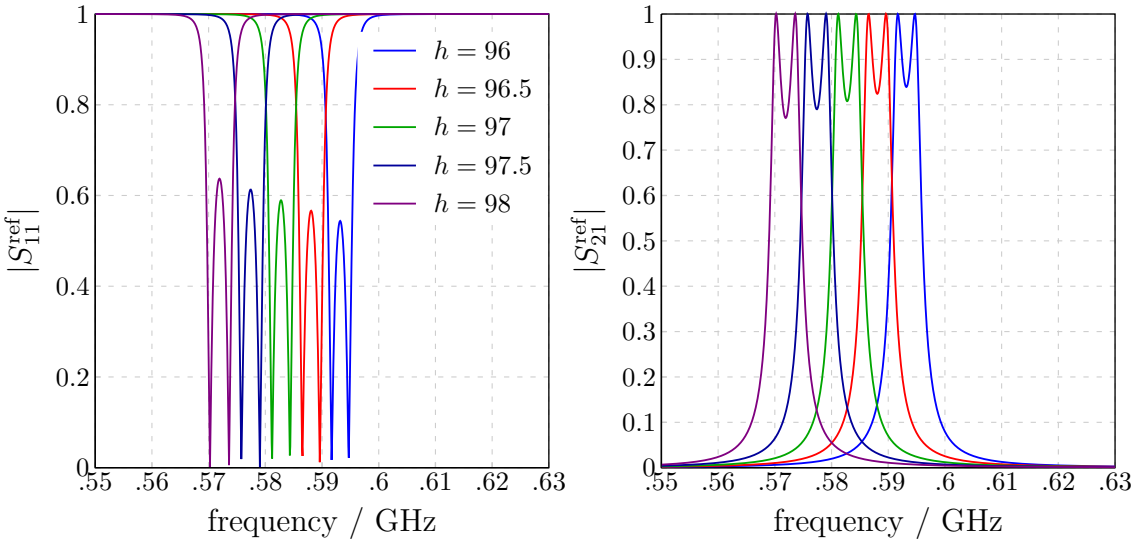


Figure 7.7: *Magnitude of the reference S-parameters for the height variation.*

Table 7.5 shows the error $|S–S^{\text{ref}}|_{\text{maxmin}}$, and the error $|\hat{S}–S^{\text{ref}}|_{\text{maxmin}}$, obtained by a reduction with a projection matrix corresponding to $q = 0$, i. e. to one interpolation point.

height h	$\lvert\hat{S}_{11}–S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert\hat{S}_{21}–S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{11}–S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{21}–S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$
96	0.02827	0.03323	0.03590	0.04373
96.5	0.02592	0.01773	0.02312	0.01238
97	0.12029	0.10336	0.03469	0.02919

Table 7.5: *Error comparison with respect to the reference solution for a reduction with one interpolation point (level $q = 0$). The error is not equally distributed in the parameter variation range. A higher level, i. e. more interpolation points, is recommended.*

At this point, a situation is apparent, that the error $|\hat{S}–S^{\text{ref}}|_{\text{maxmin}}$ of the reduced S-parameters for $q = 0$ is not equally distributed in the variation range, as at some parameter samples, for instance $h = 97$ mm in table 7.5, indeed larger than the error $|S–S^{\text{ref}}|_{\text{maxmin}}$ of the original S-parameters. Basically two options exist at this point. Either a higher level, i. e. $q = 1$, which corresponds to three interpolation points, can be taken, or the reduction should be based on a finer mesh $\mathcal{G}'$. Both options will be analyzed. Notice the errors $|\hat{S}–S^{\text{ref}}|_{\text{maxmin}}$ corresponding to the sample point $h = 96$ mm, which are lower than the respective $|S–S^{\text{ref}}|_{\text{maxmin}}$, due to the merging of the local projection matrices.

7.2.2.1 Mesh refinement

The error $|\hat{S}'-S^{\text{ref}}|_{\text{maxmin}}$ of the reduced system stemming from a mesh G′ that is finer than $\mathcal{G}$ can be compared with the error $|S'-S^{\text{ref}}|_{\text{maxmin}}$. In this case, the error again was not equally distributed for $q = 1$. This fact implies that in case that the error is not equally distributed, a finer mesh does not improve the results. It is rather recommended to increase the number of interpolation points as described above.

7.2.2.2 Increase of the level

Alternatively, increasing the reduction level leads to the results shown in table 7.6. Due to the merging of the local projection matrices, the error is distributed equally in the parameter range. Considering that the reductions at each point by the single-parameter Arnoldi algorithm can be performed in parallel, the overall time $t_{\text{red}}(q = 1)$ to calculate the projection matrix $\mathbf{V}(q = 1)$ is approximately the same as the time $t_{\text{red}}(q = 0)$ to calculate the projection matrix $\mathbf{V}(q = 0)$. More precisely, assuming that the time to perform the SVD is negligible compared to t_{red}, it is:

$$t_{\text{red}}(q = 1) = t_{\text{red}}(q = 0) + t_{\text{SVD}} \approx t_{\text{red}}(q = 0), \text{for parallel reductions.} \tag{7.2.4}$$

height h	$\lvert\hat{S}_{11}-S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert\hat{S}_{21}-S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{11}-S_{11}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$	$\lvert S_{21}-S_{21}^{\text{ref}}\rvert_{\text{min}}^{\text{max}}$
96	0.03746	0.04300	0.03590	0.04373
96.5	0.02672	0.01276	0.02312	0.01238
97	0.03389	0.02704	0.03469	0.02919

Table 7.6: *The table summarizes the errors with respect to the reference solution for a reduction with level $q = 1$. Three interpolation points correspond to this level. The results show an equal distribution in the variation range.*

The computation time in MWS is $t_{\text{S,MWS}} = 565\,\text{s}$, while the reduction time is $t = 7.300\,\text{s}$.

Finally, the pictures in figure 7.8 shows the magnitudes and the complex curves of the S-parameters S^{ref}, S and $\hat{S}$. Figure 7.8 depicts the errors $|S-S^{\text{ref}}|_{\text{maxmin}}$, $|\hat{S}-S^{\text{ref}}|_{\text{maxmin}}$ and $|\hat{S}-S|_{\text{maxmin}}$ for an arbitrary point in the parameter range. Obviously, the MOR error has the smallest value.

7.3 The Langer Filter

The filter shown in figure 7.10 is known as Langer filter. This filter has gained the status of a reference model, as it has already served as a test example for the single-parameter MOR methods in [75] as well as for the multi-parameter MOR methods for FE problems in [17]. But it has been also used in other works [16, 64] associated with electromagnetic field simulation problems with the FIT.

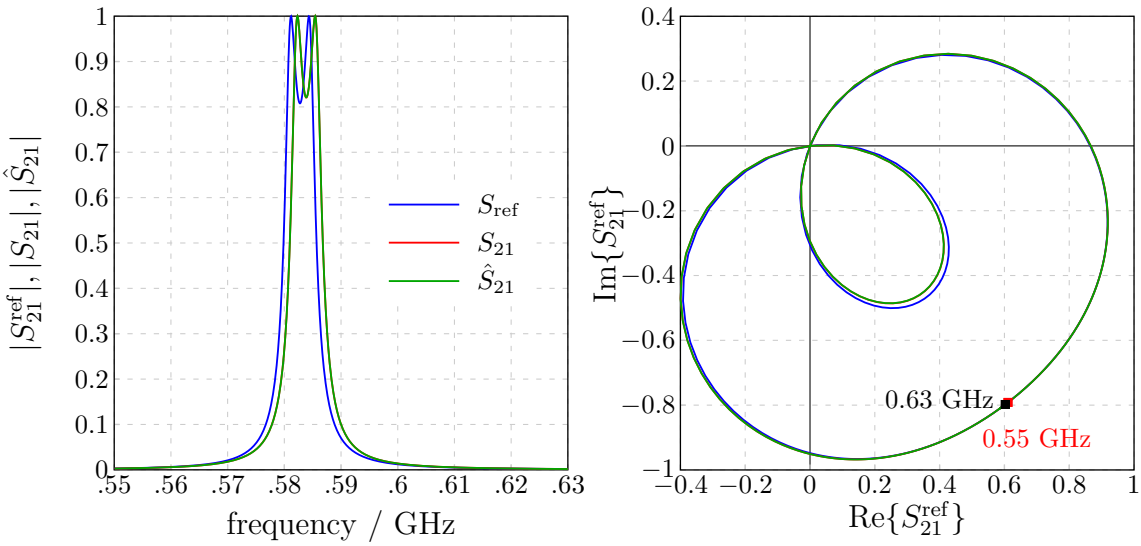


Figure 7.8: *Comparison of the magnitudes and the complex curves of the narrowband filter S-parameters S^{ref}, S and $\hat{S}$. The curves of the latter two are indistinguishable from each other, implying that the overall error is dominated by the error from the coarse mesh resolution of $\mathcal{G}$ compared to the mesh $\mathcal{G}_{\text{ref}}$.*

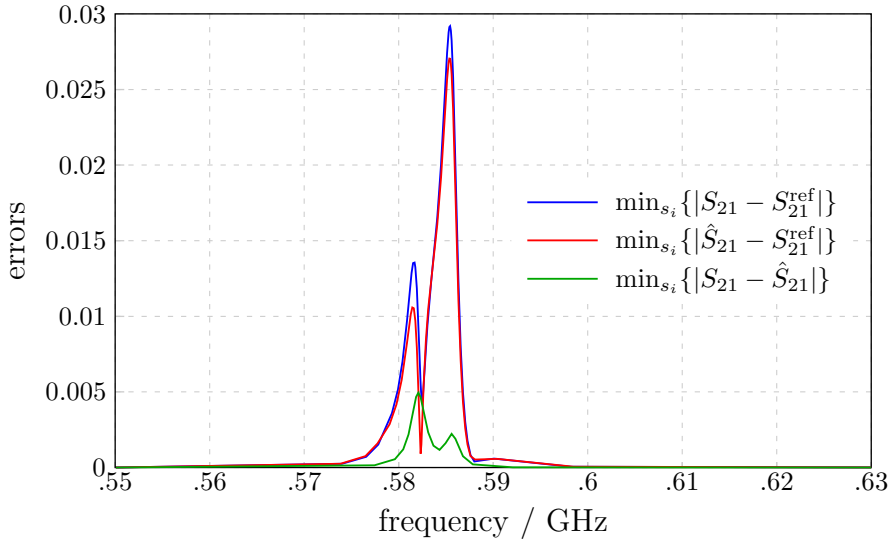


Figure 7.9: *The errors for an arbitrary point in the parameter range of the narrowband filter clearly show that the error stemming from the lower mesh resolution is significantly higher than the MOR error.*

The Langer filter has a symmetric structure consisting of a metallic box which includes two cylindrical, dielectric insertions. Therefore, it is also called dielectric filter. It is fed by two coaxial ports which are elongated inner conductors of two coaxial lines. The resonant behavior is controlled by the two dielectric rings, which act as coupled resonators and which have nominal value of $\varepsilon_r = 38$. For this value the filter serves as a sharp band pass at approximately 4.6 GHz. From a numerical viewpoint, it is interesting to consider also resonances above 6.5 GHz. The exact geometrical specifications can be found in [16].

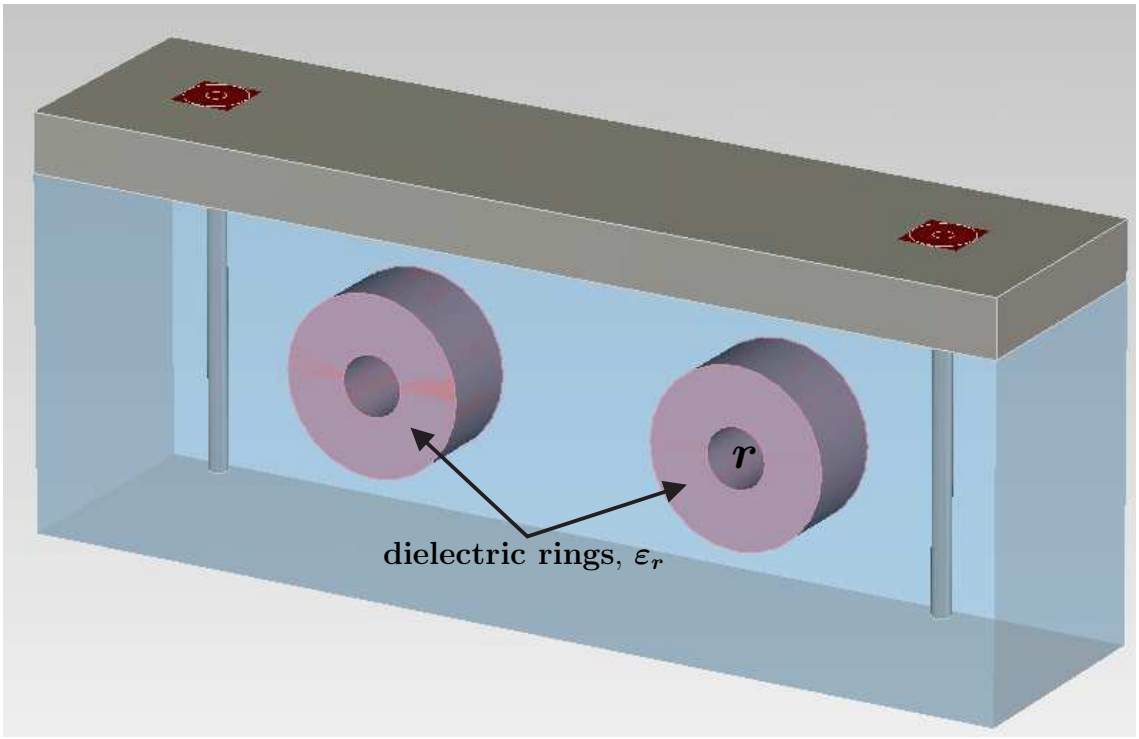


Figure 7.10: *The Langer filter modeled with the FIT method. The parameters considered are the permittivity ε_r of the cylindrical discs and their inner radius r.*

Due to the symmetry, a magnetic boundary condition can be applied, thus only half of the filter needs to be discretized to obtain the results of the entire structure. A hexahedral mesh is used. In the frequency range between 4.0 GHz and 7.5 GHz the dielectric filter has 8 eigenfrequencies, thus being a demanding problem type.

Langer filter parameter	Range
frequency s	$4.0 \ldots 7.5$ GHz
permittivity ε_r	$33 \ldots 41$
inner radius of disc	$2.0 \ldots 2.8$ mm

The Langer filter is a highly resonant filter which is very sensitive to the variations of the cylindrical inclusions. Both the relative permittivity ε_r of the rings as well as the values of the inner radii affect the frequency response.

7.3.1 Variation of the Inner Radius of the Disc

The variation of the radius is a characteristic example of a variation that is larger than one mesh cell. As the disc material is dielectric, the entries of the material matrices do not change to zero for parameter variations that are larger than the mesh cell size. As described in the previous chapter, this would be the case for PEC material. Therefore, during the variation of the parameters the mesh can be kept constant leading to constant-size material matrices.

The reference curves have been calculated with a mesh of 10.868.736 cells corresponding to 90 lines per wavelength with the "Time Domain Solver". Due to stored energy inside the Langer filter its solutions converge very slowly. The solver terminated for a remaining stored energy of -50 dB. As the computation is very

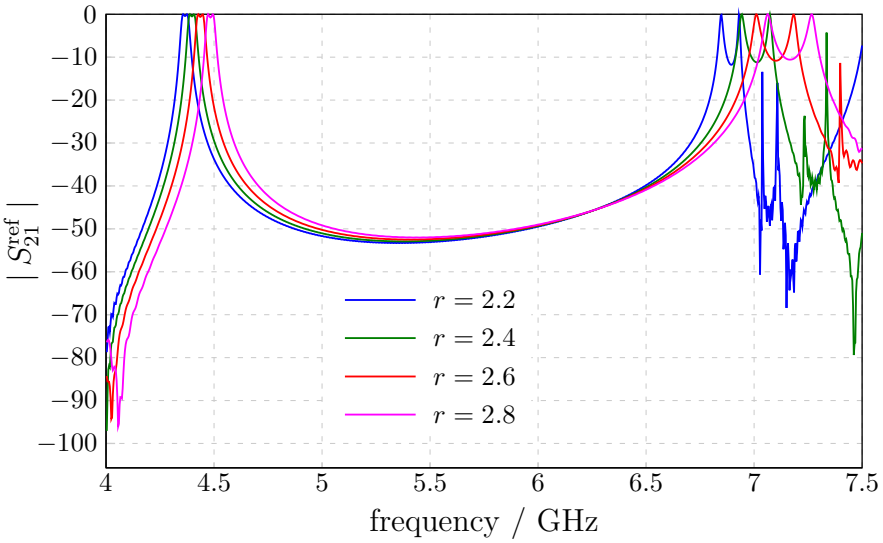


Figure 7.11: *Magnitude of the reference S-parameters* S_{21}^{ref} *for the variation of* r.

time-consuming, the computation was performed by means of computing on graphics processing units of a cluster computer.

A mesh convergence analysis showed that even for the mesh $\mathcal{G}$ a high grid size is required. Therefore, $\mathcal{G}$ has been chosen to consist of 560.000 elements, despite the fact, that there is still a deviation from the reference as indicated in table 7.7. Figure 7.11 shows the magnitudes of the reference S-parameters for different values of r.

A deviation from the original curves and the reference solution can be observed only on the right side of the spectrum. The results comply with the error investigation of the Langer filter in [75, p.114]. Therein the TSL algorithm was applied, which in the second step consists of the single-parameter Arnoldi algorithm. Calculating the projection matrix with 50 columns lead to results in the range of 10^{-9} for the spectrum lower than 6.5 GHz, and for errors in the range of 10^{-6} for the spectrum above 6.5 GHz. Nevertheless, in view of the lower calculation time and smaller reduced subspace, a projection matrix consisting of 22 columns was advised. In this case, while in the range between 4.0 and 6.5 GHz the errors are again in the order of 10^{-9}, in the higher spectrum errors reach orders of 10^{-2}. This is in agreement with the experience that the classical Arnoldi algorithm approximates the small eigenvalues first. Nevertheless, in practical settings, even this error is tolerable.

Based on the experience gained by the microstrip example in the previous chapter and table 7.2, for variations with one parameter (additional to the frequency) it is sufficient to use one or three interpolation points. Using more than 5 interpolation points may even deteriorate the result. Thus, in the following, first a reduction with one interpolation point will be analyzed and then with three points. Again, for the latter, the matrix $\mathbf{V}_{\text{all}}$ is truncated for m such that $m = 0.8 \cdot \text{size}(\mathbf{V}_{\text{all}})$.

Table 7.7 compares the errors $|\hat{S}-S^{\text{ref}}|_{\text{maxmin}}$ and $|S-S^{\text{ref}}|_{\text{maxmin}}$ from the reduction with $q = 0$. The higher values are explained by the comparably large error in the higher spectrum. Again, some parameter samples have larger errors than others, implying to choose a higher level for the reduction.

r	$\lvert\hat{S}_{11}-S_{11}^{\mathrm{ref}}\rvert_{\mathrm{min}}^{\mathrm{max}}$	$\lvert\hat{S}_{21}-S_{21}^{\mathrm{ref}}\rvert_{\mathrm{min}}^{\mathrm{max}}$	$\lvert S_{11}-S_{11}^{\mathrm{ref}}\rvert_{\mathrm{min}}^{\mathrm{max}}$	$\lvert S_{21}-S_{21}^{\mathrm{ref}}\rvert_{\mathrm{min}}^{\mathrm{max}}$
2.2	0.15211	0.32710	0.15921	0.26940
2.4	0.12480	0.21135	0.12184	0.14469
2.6	0.06520	0.0730	0.121851	0.14353
2.8	0.08421	0.0700	0.11870	0.14227

Table 7.7: *Comparison of* $|\hat{S}-S^{\mathrm{ref}}|_{\mathrm{maxmin}}$ *and* $|S-S^{\mathrm{ref}}|_{\mathrm{maxmin}}$ *for the samples in the variation range. One interpolation point has been used for the reduction.*

r	$\lvert\hat{S}_{11}-S_{11}^{\mathrm{ref}}\rvert_{\mathrm{min}}^{\mathrm{max}}$	$\lvert\hat{S}_{21}-S_{21}^{\mathrm{ref}}\rvert_{\mathrm{min}}^{\mathrm{max}}$	$\lvert\hat{S}_{11}-S_{11}\rvert_{\mathrm{min}}^{\mathrm{max}}$	$\lvert\hat{S}_{21}-S_{21}\rvert_{\mathrm{min}}^{\mathrm{max}}$
2.2	0.14731	0.17999	0.13543	0.16920
2.4	0.12175	0.20569	0.13753	0.12211
2.6	0.11772	0.13899	0.0856	0.0914
2.8	0.1835	0.14227	0.0521	0.08119

Table 7.8: *Comparison of* $|\hat{S}-S^{\mathrm{ref}}|_{\mathrm{maxmin}}$ *and* $|\hat{S}-S|_{\mathrm{maxmin}}$ *for the samples in the variation range. The reduction has been performed with three interpolation points, which leads to a projection matrix with* 52 *columns.*

The left pictures in figure 7.12 compare the S-parameters $S_{21}^{\mathrm{ref}}, S_{21}$ and $\hat{S}_{21}$ for 4 samples in the parameter range. The right pictures compare the corresponding errors $|S_{21}-S_{21}^{\mathrm{ref}}|_{\mathrm{maxmin}}, |\hat{S}_{21}-S_{21}^{\mathrm{ref}}|_{\mathrm{maxmin}}$ and $|\hat{S}_{21}-S_{21}^{\mathrm{ref}}|_{\mathrm{maxmin}}$. All figures show a deviation in the right frequency spectrum in the area of the closely placed poles, while in the last figure they have shifted out of the observed frequency range. This deviation does not stem from the neighboring-subspace method, but is due to the choice of $q_A = 22$ iterations in the single-parameter Arnoldi. As already observed by comparing table 7.7, the error is not equally distributed over the parameter range. The smallest values are obtained for the sample $r = 2.4$ which is close to the parameter range center, where the first interpolation has been set.

The time t_S varies for each value of r, approximately it is 1 h 50 min, while the time to calculate the reduction matrix is $t_{\mathrm{red}} \approx 23\mathrm{h}$. Nevertheless, an average $t_{\hat{S}} = 23\,\mathrm{s}$ is required to calculate $\hat{S}$ from the reduced system matrices.

Table 7.8 shows the errors $|\hat{S}-S^{\mathrm{ref}}|_{\mathrm{maxmin}}$ and $|\hat{S}-S|_{\mathrm{maxmin}}$ for the reduction with three interpolation points. Obviously, $|S-S^{\mathrm{ref}}|_{\mathrm{maxmin}}$ is the same as in the previous setting and has not been shown again. The error norms are now in the same range for all samples, furthermore their values lie below those in the previous setting. Figure 7.13 compares the S-parameters S_{21}^{ref}, S and $\hat{S}$ for the sample $r = 2.4$. Figure 7.14 shows the errors obtained for two samples.

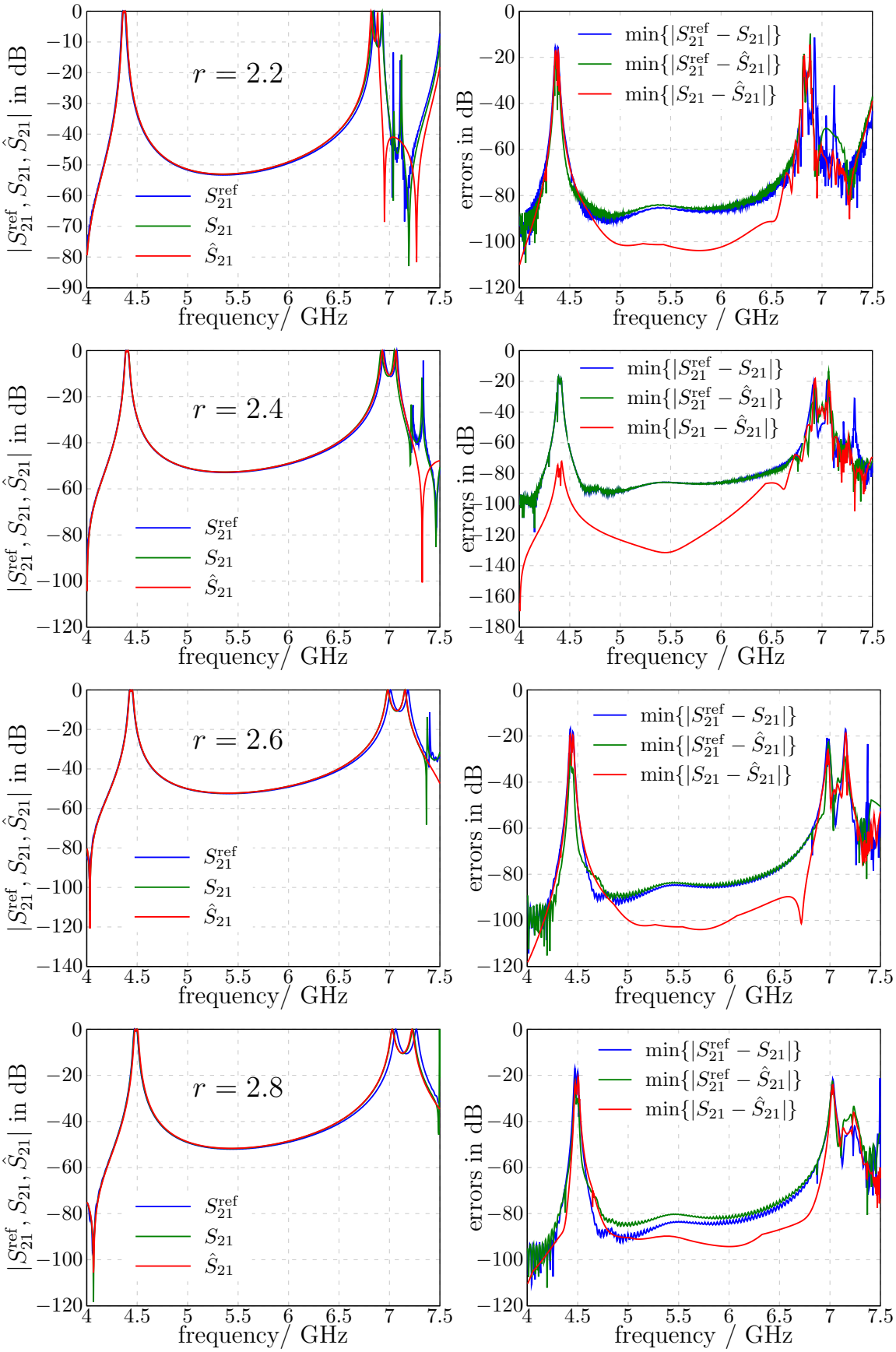


Figure 7.12: *The S-parameter $S_{21}^{\text{ref}}, S_{21}$ and $\hat{S}_{21}$ show a good agreement for all observed samples in the variation range.*

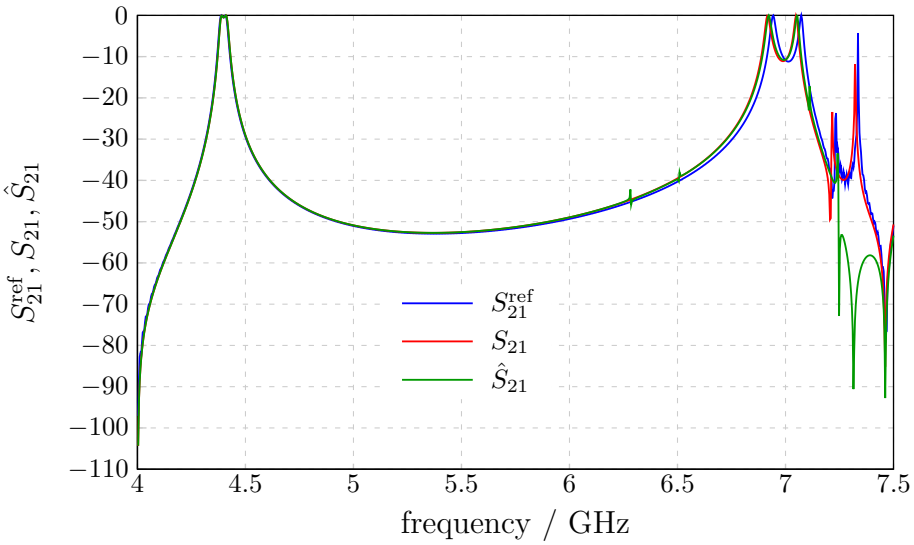


Figure 7.13: *Comparison of the S-parameters $S_{21}^{\mathrm{ref}}, S_{21}$ and $\hat{S}_{21}$. The reduced S-parameter deviates at the right side of the spectrum, nevertheless with a tolerable error, as discussed previously.*

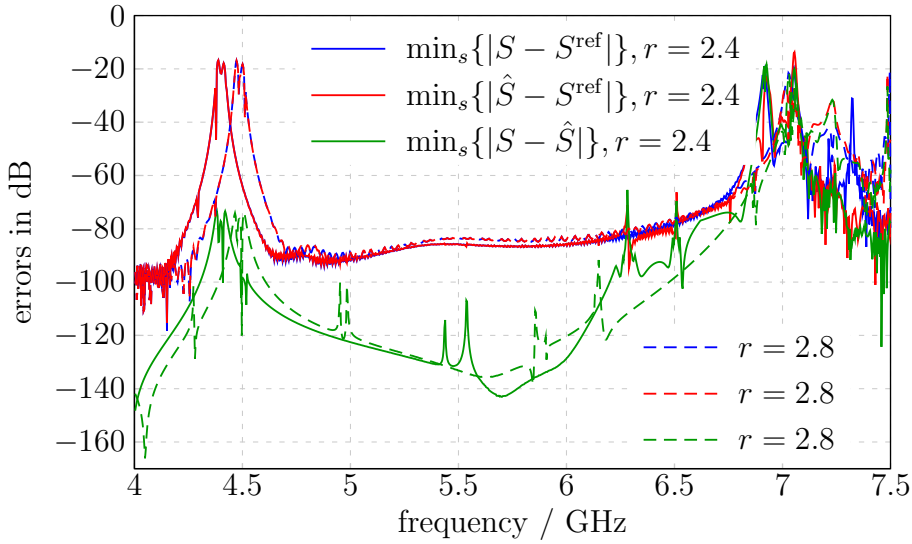


Figure 7.14: *The blue lines corresponding to the error of S compared to the reference indistinguishably lie on the red lines belonging to the error of $\hat{S}$ compared to the reference. The green lines which give the PMOR error now lie much below the blue and red lines. Furthermore they have similar values for both parameters. This is the case for the remaining parameters which are not shown in the picture.*

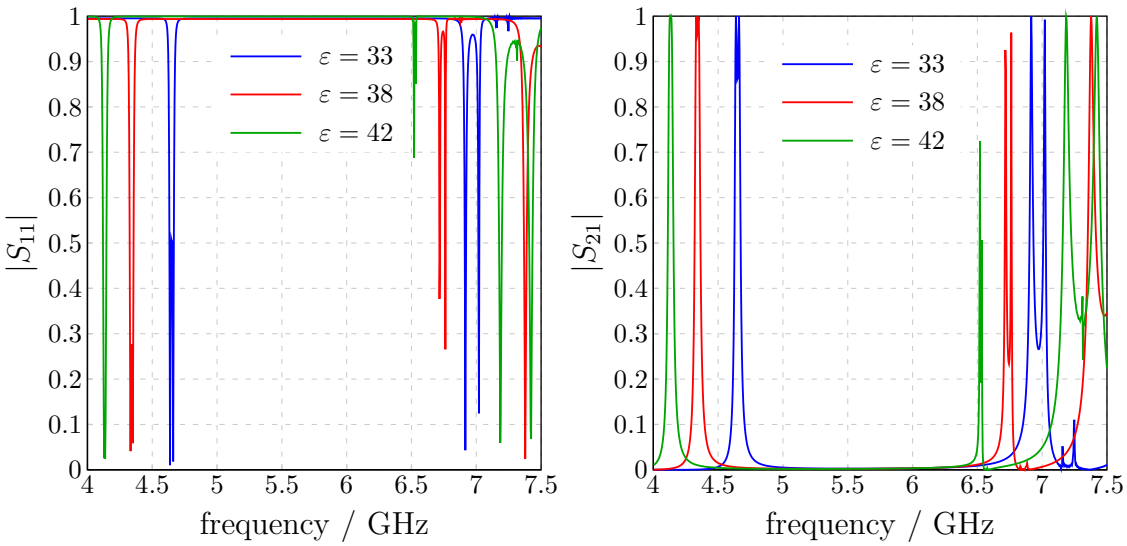


Figure 7.15: *Magnitude of the reference S-parameters S^{ref} for the ε_r variation.*

ε_r	$\lvert\hat{S}_{11}-S_{11}\rvert_{\text{min}}^{\text{max}}$	$\lvert\hat{S}_{21}-S_{21}\rvert_{\text{min}}^{\text{max}}$
33	0.1540	0.2048
34	0.1721	0.1260
35	0.1558	0.1453
36	0.1635	0.3495
37	0.1908	0.3037
38	0.1525	0.1440
39	0.1569	0.3005
40	0.3103	0.2565
41	0.1057	0.3167

Table 7.9: *Comparison of $|\hat{S}_{21}-S_{21}^{\text{ref}}|_{\text{maxmin}}$ and $|S_{21}-S_{21}^{\text{ref}}|_{\text{maxmin}}$ for the samples in the variation range. One interpolation point has been used for the reduction.*

7.3.2 Permittivity Variation

The reference S-parameter have been calculated in CST MWS with the "Time Domain Solver" on a mesh consisting of 13.553.760 cells. Again, the computation was performed on a cluster computer on graphics processing units. The solver terminated for a remaining stored energy of $-50\,\text{dB}$. Figure 7.15 shows the magnitude of the reference S-parameters S_{ref} for the ε_r variation.

Analogous considerations have been performed for the permittivity variation as previously for the radius variation. Again, both a reduction with one and three parameters will be compared.

Table 7.9 contains the errors $|\hat{S}_{11}-S_{11}|_{\text{maxmin}}$ and $|\hat{S}_{21}-S_{21}|_{\text{maxmin}}$, which are equally distributed despite the large variation range. Figure 7.16 depicts the S-parameters for one sample point and the respective errors.

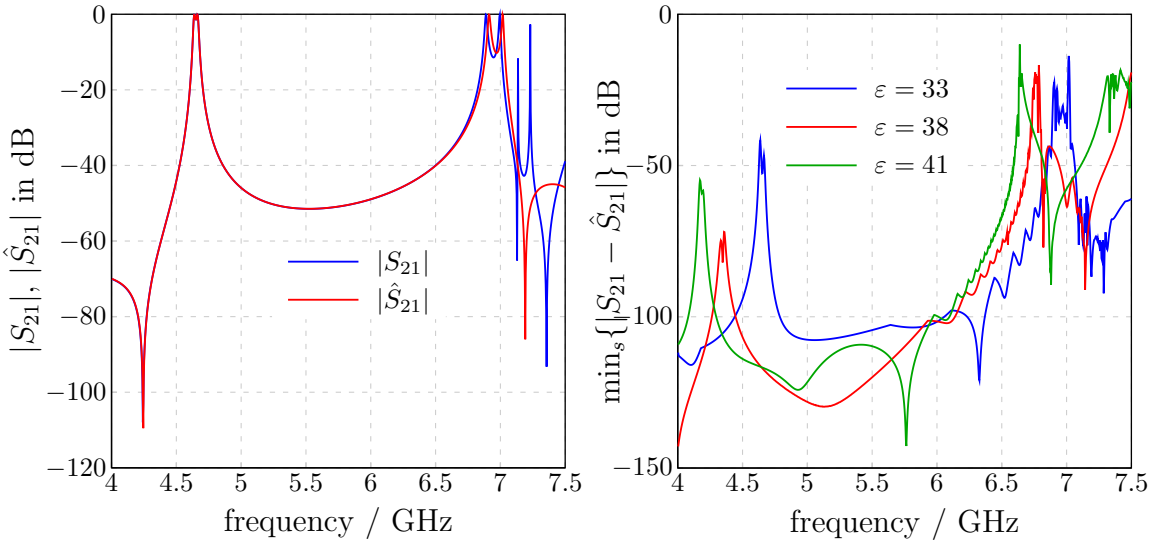


Figure 7.16: *The reduction with one interpolation point leads to a good agreement between original and reduced model. The error increases at the resonant points and would be decreased for a larger iteration step in the Arnoldi algorithm, in exchange with a larger runtime and a higher reduced system dimension. Nevertheless, the error is almost equally distributed in the large variation range of ε, as shown in the right picture for three samples.*

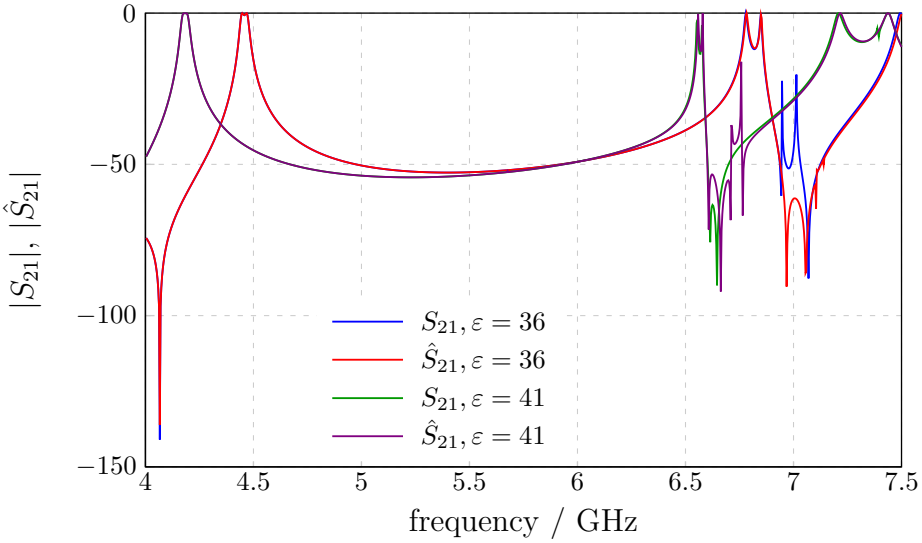


Figure 7.17: *Comparison of the S-parameters for the permittivity variation and a reduction with level 1.*

A comparison with the results obtained with the reduction using three interpolation points (figures 7.17 and 7.18), clearly shows that the latter are smaller over the whole frequency and parameter range, and that they are additionally equally distributed in the parameter range.

Both figure 7.11 and figure 7.15 show that the variation of the respective parameters leads to a nonlinear shift of the S-parameter curves. The PMOR methods being

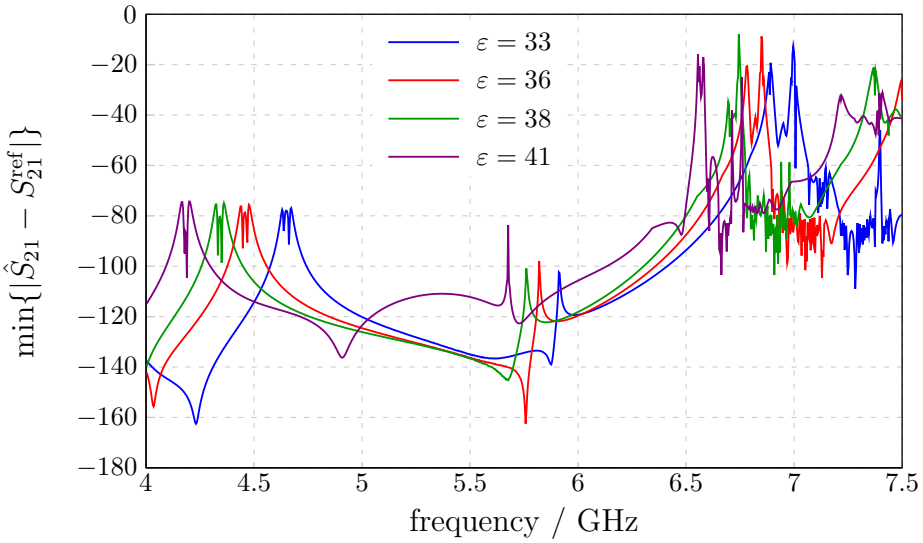


Figure 7.18: *Results obtained by the reduction with level 1 of the neighboring-subspace method for the permittivity variation. The error is smaller over the whole frequency and parameter range, and additionally it is equally distributed in the parameter range. The samples are representative for the whole parameter range.*

able to follow this variation and accurately approximate the curves even for a wide variation range gives them for this case a definitive advantage over methods based on the linearization around a working point, e. g. the sensitivity analysis.

7.4 Concluding Remarks

The discussion in this chapter clearly showed that real-world examples require a fast S-parameter approximation while featuring a tolerable accuracy. Moreover, they require a flexible variation of both material and geometry variations. PMOR methods based on the iterative construction of higher-order Krylov subspaces which require an explicit dependence on the parameters although they give accurate results over the whole variation domain, they have a double drawback. They are quasi restricted to material variations and furthermore, they are too slow for large-scale models.

The neighboring-subspace method leading to similar accuracies is superior with respect to both aspects. A parametric system with one parameter additional to the frequency parameter can be reduced with an MOR accuracy of $|\hat{S}-S^{\mathrm{ref}}|_{\mathrm{maxmin}} \approx 10^{-3}$ by using only three interpolation points. Approximately, the 20% least important singular values of the merged projection matrix can be truncated, leading to a small projection matrix. As the projections of the local systems can all be performed in parallel, the time-cost of the neighboring-subspace method is equivalent to the reduction of one system. Concluding, the neighboring-subspace approach is a very competitive method.

Chapter 8

Conclusion and Outlook

The goal of this dissertation was to develop model order reduction (MOR) methods for parametric systems in electromagnetic field simulations.

More precisely, large-scale and parameter-depending models should be reduced to smaller models which approximate the transfer function of the original model while retaining the parameter dependence in the reduced model. For an effective approach, parametric reduced systems should be both accurate over the entire parameter space of interest, as well as computationally efficient to solve. Then the parametric reduced order model needs to be generated only once for simulations with varying parameter values and can be used for instance for fast parameter sweeps. Obviously, the time required to set up the parametric reduced model should be smaller than a parameter-sweep of the original model. The systems stem from the Maxwell grid equations which result from the continuous Maxwell equations by using the Finite Integration Technique (FIT).

The FIT is an established method for the discretization of the continuous Maxwell equations. It makes use of a staggered pair of grids and its principal idea is to split the closed line integrals in the continuous equations into integrals along the grid edges and the closed surface integrals into integrals over the grid surfaces. In chapter 2 the discretization of the four Maxwell equations has been shown. Special emphasis has been given to the description of the material discretization and the resulting material matrices. The material matrices are diagonal, which is a very useful property in the context of MOR. The boundary conditions and the excitation of the computational domain have also received special attention, as the definition of the in- and output of the systems resulting from the FIT equations is based on it. A variety of system descriptions was derived from the Maxwell grid equations, in order to be able to describe the PMOR methods.

For the generation of the reduced order model there are basically two main challenges. The first consists in incorporating the parameter dependence in the original electromagnetic model such that a reduction can be generated. The reason for this is that most existing approaches to generate a parametric reduced order model are based on electromagnetic systems that require the same discretization grid for all parameters.

The second challenge consists in the reduction method itself. As mentioned earlier, the reduction method should efficiently produce a parametric reduced order model, i.e. a low-cost model with respect to the parameters. The efficiency is important, as if the parameter dependence is not handled carefully, the computational cost to generate the reduced model can easily exceed the time gained from using the reduced models, despite their fast evaluation.

The main focus was set to material and geometry parameters. These two parameter kinds require different approaches, as they exhibit an important difference: geometry variations do affect the underlying mesh, while material parameter variations do not. As the mesh is responsible for the sizes of the system matrices, geometry parameter variations may lead to different sized matrices. Already existing methods for parametric MOR (PMOR) cannot cope with different sized matrices. While the reduction step is performed by a general projection framework, based on these considerations two avenues for developing PMOR methods for FIT systems have been followed.

The multi-parameter Arnoldi and the contraction method belonging to the first category are based on moment matching between original and reduced system. They offer a stable calculation of the higher-order Krylov subspaces spanned by the multi-parameter moments, as their direct calculation features numerical instabilities similar to the single-parameter case. The example of the microstrip line shows that in particular the contraction method is a reliable method leading to high accuracies. Nevertheless, due to the multiple evaluation of the WCAWE algorithm on which it relies, its application to real-world problems lead to a very high time-cost. The Langer filter and the narrow-band filter examined in chapter 7 could not be efficiently handled with the contraction method.

Both methods require an explicit dependence on the parameters which is naturally not given in FIT systems. Therefore they are not applicable to systems in the form directly obtained by the Maxwell grid equations. Nevertheless, by applying the linearization step introduced in this work, both for material as well as for geometry parameters, the FIT systems were appropriately adapted to the desired form. The projection matrix is associated with higher-order Krylov subspaces, and is therefore the multi-parameter equivalent of single-parameter systems based on ordinary Krylov subspaces.

The linearization step requires a constant mesh topology. For material parameter variations the mesh topology does not change, therefore the methods can be directly applied. As geometry variations cause mesh modifications, using the same grid during the variation is challenging. There are attempts to overcome this difficulty, for instance by guarantying a constant mesh topology by manually intervening in the mesh creation. By means of the commercial software package CST Microwave Studio (MWS), which is used to generate the FIT models, although possible by using fixpoints this procedure is not practicable for complex problems. Furthermore, it has to be adjusted to each specific problem separately. The trend thus leads to methods that are independent from the underlying discretization grid. This forms one of the most actual research topics in the field of MOR. In this work an alternative approach was developed, the neighboring-subspace method.

This approach is based on using the system in the form directly obtained from the Maxwell grid equations, that is, no explicit dependence of the system matrices on the parameters is required. The projection matrix of the parametric system is defined as the composition of local projection matrices. No moment matching property can be guaranteed for this method, but it provides more flexibility in the geometry variation than the contraction method, i. e. it does not require a constant mesh-topology but only a constant mesh. Naturally, the variation range of geometrical parameters in this method is restricted to the mesh cell size, though considerations have been performed to loosen this restriction.

The neighboring-subspace method has been successfully applied to the microstrip example of chapter 6, where both geometry and material parameters were varied. As the example of the variation of the length, the width and the depth of the microstrip line shows, it is even efficient for more than one parameter. Compared to all three methods it is so far the only method that can cope with large-scale models. The systems stemming from the narrow-band filter and the Langer filter of chapter 7 were successfully reduced by using this method. Its efficiency lies in the fact that it is based on powerful single-parameter MOR methods. As both the contraction and the neighboring-subspace method make use of single-parameter reduction methods, in the first part of chapter 5 a brief overview of basic single-parameter reduction methods was given.

Filter tuning has already been addressed in this work as an important application of PMOR methods. Nevertheless, a variety of potential applications of PMOR methods exist indicating several directions of future research topics.

One potential application that is currently subject to research is the coupled simulation of nano-electric structures. Using simulations of the full nonlinear coupled model, equivalent models for a nano-electric structure are obtained. For these models suitable mathematical models are used in order to develop PMOR techniques for a fast simulation of new semiconductor structures for nano-technology and microsystems technology. Due to the small scale and high density of conductors, quantum effects have to be considered where appropriate.

In the field of network analysis the idea of replacing a complex circuit with a simpler circuit that features the same approximate behavior is very common. This idea is extended to replacing an electromagnetic device with a physically motivated equivalent setting, e. g. an electric circuit, featuring an approximate output-input behavior to the original device. That is, there is a correspondence between an electromagnetic device and an equivalent electric circuit.

The equivalent circuit cannot be obtained directly from the original electromagnetic system, as the model size and therefore the number of required circuit elements would be too high. Both the classical as well as the parametric reduced models can be used for this purpose. In industry, these equivalent circuits play an important role, as they are more versatile that the electromagnetic systems they model.

If the topology of the equivalent electrical circuit is known, i. e. the way the circuit components are arranged with respect to each other, with the help of MOR techniques the circuit component values can be determined. Nevertheless, the ultimate

goal is a full automated generation of the equivalent circuit. At the moment only semi-analytic methods have been developed, that is, the topology is known, but some further information such as specific geometry parameters have to be specified in order to generate the equivalent circuit. Important open questions are e. g. how parameter variations affect the equivalent circuit.

In view of the high research activities initiated to face environmental challenges such as global climate change or dwindling fossil energy, techniques as described above are truly substantial. Currently, alternative solutions based on regenerative electric power are developed. A major field is wind power, which is one of the most important current energy sources. The actual research is focused on developing powerful electric energy converters for electric motors, energy storage technology and power electronics. The trend described above is apparent also in automotive industry, where new drive solutions are developed, as conventional combustion-motors are progressively replaced by electric motors. The enormous size and actuality of this research area reveal the importance of the development of supporting sciences such as PMOR.

Symbols and Abbreviations

General Symbols

$\mathbb{R}$	real numbers

Continuous Electrodynamics

$\vec{H}$	magnetic field strength
$\vec{E}$	electric field strength
$\vec{D}$	dielectric flux density
$\vec{J}$	electric current density
ϱ	charge density
A	arbitrary surface area
V	arbitrary volume
$\vec{J}$	current density
$\vec{J}_e$	impressed current density
ε_0	permittivity of free space
$\vec{M}(\vec{H}, \vec{r}, t)$	magnetization
$\vec{P}(\vec{E}, \vec{r}, t)$	polarization
χ_e	electric susceptibility
χ_m	magnetic susceptibility
ε_r	relative permittivity
μ_r	relative permeability
σ_F	surface charge density
$\vec{J}_A$	surface current density
$\vec{n}$	unit normal vector
$\underline{k}$	complex wave vector
$\delta = \sqrt{\frac{2}{\omega\mu\kappa}}$	skin depth
λ	wave length

Discrete electrodynamics

$\Omega \subset (\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3)$	domain of interest in the continuous space
$\mathcal{G}_i$	finite set of discrete elements, grid cells
$\mathcal{G}$	computational grid, primary grid, mesh
$\widetilde{\mathcal{G}}$	dual grid
$P(i,j,k)$	primary elementary points
$L(i,j,k)$	primary elementary edges
$A(i,j,k)$	primary elementary surfaces
$\widetilde{P}(i,j,k)$	dual elementary points
$\widetilde{L}(i,j,k)$	dual elementary edges
$\widetilde{A}(i,j,k)$	dual elementary surfaces
I, J, K	number of points in each coordinate direction of the primary grid
n_P	number of mesh points
i, j, k	corresponding indices
$\widehat{\mathbf{e}}$	discrete electric voltage
$\widehat{\mathbf{h}}$	magnetic voltage
$\widehat{\widehat{\mathbf{b}}}$	discrete magnetic flux
$\widehat{\widehat{\mathbf{d}}}$	discrete dielectric flux
$\mathbf{C}_{\mathrm{FIT}}, \widetilde{\mathbf{C}}_{\mathrm{FIT}}$	discrete curl operator for primary and dual grid, respectively
$\mathbf{S}_{\mathrm{FIT}}, \widetilde{\mathbf{S}}_{\mathrm{FIT}}$	discrete divergence operator for primary and dual grid, respectively
$\widetilde{\mathbf{D}}_S$,$\widetilde{\mathbf{D}}_A$	diagonal geometry matrices
$\overline{\mu^{-1}}$	averaged permeability
$\mathbf{M}_\mu, \mathbf{M}_\varepsilon, \mathbf{M}_\kappa$	material matrices
f_{A_m}	proportion of mesh cell surface A_m filled with PEC material
f_{L_i}	proportion of the mesh edges L_i which are inside the PEC material
$\widehat{\mathbf{e}}_t$,$\widehat{\mathbf{h}}_t$	two-dimensional transversal components of $\widehat{\mathbf{e}}$ and $\widehat{\mathbf{h}}$, respectively
a_m	$m-$th mode traveling in positive coordinate direction
b_m	$m-$th mode traveling traveling in negative direction
$\widehat{\mathbf{e}}_{t,m}$ and $\widehat{\mathbf{h}}_{t,m}$	2D phasors containing only transversal elements $\widehat{\mathrm{e}}, \widehat{\mathrm{h}}$ of mode m
$\mathbf{N}_w$	discrete cross product operator
$Z = E^{\mathrm{T}}/H^{\mathrm{T}}$	port impedance
$\mathbf{Z}_L^{1/2}$	diagonal matrix containing $\sqrt{Z_{L,m}}$
$\mathbf{u}, \mathbf{i}$	generalized voltage and current at the ports
$k_{w,3D}$	constant for calculating the excitation vector
$F_{\mathrm{TE}}, F_{\mathrm{TM}}$	correcting factors for TE and TM modes at the waveguide ports

Systems

$\mathbb{U}$	set in which the outputs of a system are possible
$\mathfrak{B}$	system behavior, according to the behavioral approach
$\mathbf{s} = (s_1, \ldots, s_r)$	parameter-vector including s
$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{r-1})$	parameter-vector excluding s
$\mathbf{i}, \mathbf{u}$	in- and output
$\mathbf{x}$	state for systems in state-space form, auxiliary variable for systems in other forms
$\mathbf{X}$	matrix containing $\mathbf{x}$ (MIMO)
$\boldsymbol{\Sigma}$	system in state-space form
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$	matrices corresponding to the system in state-space form
$\boldsymbol{\Sigma}_{\mathrm{higher-order}}$	higher-order system
$\mathbf{A}_k, \mathbf{B}_k, \mathbf{C}_k, \mathbf{D}_k$	matrices corresponding to the higher-order-system
$\boldsymbol{\Sigma}_\mathbf{s}$	$\mathbf{s}$-parameterized system
$\mathbf{A}_\mathbf{s}, \mathbf{B}_\mathbf{s}, \mathbf{C}_\mathbf{s}, \mathbf{D}_\mathbf{s}$	matrices corresponding to the $\mathbf{s}$-parameterized system
$\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$	$\boldsymbol{\xi}$-parameterized system
$\mathbf{A}_{\boldsymbol{\xi}}, \mathbf{B}_{\boldsymbol{\xi}}, \mathbf{C}_{\boldsymbol{\xi}}, \mathbf{D}_{\boldsymbol{\xi}}$	matrices corresponding to the $\boldsymbol{\xi}$-parameterized system
$\mathcal{P}, \mathcal{Q}$	reachability and observability gramians
$\mathcal{P}_t, \mathcal{Q}_t$	transformed reachability and observability gramians
$\mathbf{h}$	impulse response
$\mathbf{H}$	transfer impedance
$\boldsymbol{\eta}_k(\infty), \boldsymbol{\eta}_k(0), \boldsymbol{\eta}_k(\sigma)$	kth Markov parameter, moment or shifted moment, respectively
$\mathcal{K}_q(\mathbf{A}, \mathbf{b})$	Krylov subspace of order q of matrix $\mathbf{A}$ with respect to vector $\mathbf{b}$
$\mathbf{H}$	upper Hessenberg matrix arising in the Arnoldi process
$\mathbf{r}$	vector perpendicular to the Arnoldi vectors
$\mathbf{T}_k$	triband matrix in the Lanczos algorithm
q	number of moments matched
$\mathbf{z}, \mathbf{Z}$	test vector (SISO), matrix containing test vectors (MIMO)

Order Reduction

$\hat{\mathbf{x}}$	reduced state for systems in state-space form, reduced auxiliary variable for systems in other forms
$\hat{\boldsymbol{\Sigma}}$	reduced system in state-space form
$\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}}$	reduced matrices for system in state-space form
$\hat{\boldsymbol{\Sigma}}_{\mathrm{higher-order}}$	reduced higher-order system

$\hat{\mathbf{A}}_k, \hat{\mathbf{B}}_k, \hat{\mathbf{C}}_k, \hat{\mathbf{D}}_k$	matrices corresponding to the reduced higher-order system
σ	Hankel singular value
$\mathbf{K}$	matrix with eigenvectors
$\mathbf{U}$	Cholesky factor of the reachability gramian
Σ	matrix with eigenvalues
$\mathbf{T}$	balancing transformation matrix
$\mathbf{\Gamma}$	coefficient matrix in POD

Parametric Model Order Reduction

$\hat{\mathbf{\Sigma}}_\mathbf{s}$	reduced $\mathbf{s}$-parameterized system
$\hat{\mathbf{A}}_\mathbf{s}, \hat{\mathbf{B}}_\mathbf{s}, \hat{\mathbf{C}}_\mathbf{s}, \hat{\mathbf{D}}_\mathbf{s}$	matrices corresponding to the reduced $\mathbf{s}$-parameterized system
$\hat{\mathbf{\Sigma}}_{\boldsymbol{\xi}}$	reduced $\boldsymbol{\xi}$-parameterized system
$\hat{\mathbf{A}}_{\boldsymbol{\xi}}, \hat{\mathbf{B}}_{\boldsymbol{\xi}}, \hat{\mathbf{C}}_{\boldsymbol{\xi}}, \hat{\mathbf{D}}_{\boldsymbol{\xi}}$	matrices corresponding to the reduced $\boldsymbol{\xi}$-parameterized system
f_{method}	norm of difference between reduced and reference S-parameters
$\bar{f}_{s,j}$	average error over frequency for one parameter-set
$\bar{f}_\xi$	average error over frequency and parameter samples
$\bar{f}_{\max,\xi}$	$\max(\bar{f}_\xi)$ over parameter samples
t_{red}	time to calculate projection matrix
t_S	time to evaluate S-parameters from the original system matrices
$t_{\hat{S}}$	time to evaluate S-parameters from the reduced system matrices
$\mathcal{S}$	hierarchical bases for contraction method
$\mathbf{p}$	contraction point
q_A	order of Arnoldi method underlying neighboring-subspace method

General

CST AG	Computer Simulation Technology AG
MWS	Microwave Studio
ODE	ordinary differential equation
PDE	partial differential equation
ENIAC	Electronic Numerical Integrator and Computer

System Theory

LTI	Linear Time Invariant System
SVD	Singular Value Decomposition

EVD	Eigenvalue Decomposition

Continuous electrodynamics

TE modes	Transverse Electric
TM modes	Transverse Magnetic
TEM modes	Transverse Electro-Magnetic

Discrete electrodynamics

FIT	Finite Integration Theory
PEC	Perfectly Electric Conducting
PFC	Partially Filled Cells
PML	Perfectly Matched Layer boundary

Model Order Reduction

MOR	Model Order Reduction
PMOR	Parametric Model Order Reduction
POD	Proper Othogonal Decomposition
AWE	Asymptotic Waveform Evaluation
CFH	Complex Frequency Hopping
IRAM	Implicitly Restarted Arnoldi Method
GMRES	Generalized Minimal Residuals
PVL	Padé Via Lanczos
PRIMA	Passive Reduced-Order Interconnect Macro-Modeling Algorithm
TSL	Two-Step Lanczos
MIMO	Multi Input Multi Output
SISO	Single Input Single Output
WCAWE	Well Conditioned Asymptotic Waveform Evaluation

Index

Bibliography

[1] A. Antoulas. *Approximation of Large Scale Dynamical Systems*. Advances in Design and Control. SIAM, 2005.

[2] W. E. Arnoldi. The principle of minimized iterations in the solution of the matrix eigenproblem. *Quarterly of Applied Mathematics*, 9:17–29, 1951.

[3] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12:273–288, 2000.

[4] U. Baur and P. Benner. Modellreduktion für parametrisierte Systeme durch balanciertes Abschneiden und Interpolation. *Automatisierungstechnik*, 57(8), 2009.

[5] T. Bechthold, B. Salimbahrami, and B. Lohmann. A two-sided Arnoldi algorithm with stopping criterion and MIMO selection procedure. *Mathematical and Computer Modeling of Dynamical Systems*, 11(1):79–93, 2005.

[6] P. Benner and L. Feng. A robust algorithm for parametric model order reduction. *Proceedings in Applied Mathematics and Mechanics*, 2007.

[7] P. Benner, P. Kürschner, and J. Saak. Improved second-order balanced truncation for symmetric systems. *Proceedings of the Vienna Mathmod Conference*, February 2012.

[8] A. Bultheel and M. Van. Barel. Padé techniques for model reduction in linear system theory. *Journal of Computational and Applied Mathematics*, 14:401–438, 1986.

[9] H.J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica, Cambridge University Press*, pages 1–123, 2004.

[10] E. Chiprout and M. S. Nakhla. Asymptotic waveform evaluation and moment matching for interconnect analysis. *Boston, MA:Academic publishers*, 1994.

[11] L. Codecasa. Boundary condition independent compact dynamic thermal networks of packages. *IEEE Transactions on Components and Packaging Technologies*, 28(4):593–604, 2005.

[12] R. E. Collin. *Foundations for Microwave Engineering.* The IEEE Press Series on Electromagnetic Wave Theory, 1992.

[13] CST AG, Bad Nauheimer Strasse 19, 64289 Darmstadt, Germany. *MICROWAVE STUDIO.*

[14] CST AG, Bad Nauheimer Strasse 19, 64289 Darmstadt, Germany. *Online Help.*

[15] L. Daniel, O. Siong, K. Lee, and J. White. A multiparameter moment matching model reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 5:678–693, 2004.

[16] M. Dohlus. *Ein Beitrag zur numerischen Berechnung elektromagnetischer Felder im Zeitbereich.* Ph. D. thesis, Technische Universität Darmstadt, 1992.

[17] O. Farle. *Ordnungsreduktionsverfahren für die Finite-Elemente-Simulation parameterabhängiger passiver Mikrowellenstrukturen.* Ph. D. thesis, Universität des Saarlandes, Online Catalogue of the "Saarländische Universitäts- und Landesbibliothek", 2007.

[18] P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *Proceedings of EURO-DAC 94, EURO-VHDL 94, IEEE Computer Society Press*, pages 179–175, 1994.

[19] O. Föllinger. *Regelungstechnik.* Hüthig, 1994.

[20] R. W. Freund and P. Feldmann. Reduced-order modeling of large linear passive multi-terminal circuits using matrix-Padé approximation. *Proceedings of the Design Automation and Test in Europe Conference*, pages 530–537, 1998.

[21] M. Gasca and T. Sauer. Polynomial interpolation in several variables. *Advances in Computational Mathematics*, 12:377–410, 2000.

[22] G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, 3rd edition, 1996.

[23] W. B. Gragg and A. Lindquist. On the partial realization problem. *Linear Algebra and its Applications*, 50, 1983.

[24] E. J. Grimme. *Krylov projection methods for model reduction.* Ph. D. thesis, University of Illinois at Urbana Champaign, 1997.

[25] E. J. Grimme, D. S. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. Technical report crpc-tr94458, Rice University Houston, Rice University Houston, TX 77251-1892, USA, 1994.

[26] G. A. Baker (jr) and P. Graves-Morris. *Padé approximants*, volume 59 of *Encyclopedia of mathematics and its applications.* second edition, 1996.

[27] Andreas Klimke. *Sparse Grid Interpolation Toolbox*. Universität Stuttgart.

[28] M. Krüger and A. Trächtler. Approximation of Pareto-optimal systems using parametric model-order reduction. *Proceedings of the Vienna Mathmod Conference*, February 2012.

[29] K. Krohne. *Verfahren zur Reduzierung der Modellordnung FIT diskretisierter elektromagnetischer Strukturen*. Diplomarbeit, TU Darmstadt, 2002.

[30] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45(4):255–282, October 1950.

[31] A. T. Leung and R. Khazaka. Parametric model order reduction technique for design optimization. In *Proceedings of International Symposium on Circuits Systems*, pages 1290–1293, 2005.

[32] P. Li, F. Liu, X. Li, L. T. Pileggi, and S. R. Nassif. Modelling interconnect variability using efficient parametric model order reduction. In *Proceedings of the Design, Automation and Test in Europe Conference and Exhibition*, pages 958–963, Munich, Germany, 2005.

[33] H. G. Liddell and R. Scott. *A Greek-English Lexicon*. Perseus Digital Library, Tufts University.

[34] J. Lienemann. *Complexity reduction techniques for advanced mems actuators simulation*. Ph. D. thesis, Albert-Ludwigs-Universität Freiburg, 2006.

[35] The MathWorks Inc. MATLAB®, 2011.

[36] B. C. Moore. Principal component analysis in linear systems: Controllability, observability and model reduction. *IEEE Transactions on Automatic Control*, 26:17–32, 1981.

[37] N. Nicolova, J. W. Bandler, and M. H. Bakr. Adjoint techniques for sensitivity analysis in high-frequency structure CAD. *IEEE Transactions on Microwave Theory and Techniques*, 52(1):403–419, January 2004.

[38] A. Odabasioglu, M. Celik, and L. T. Pileggi. PRIMA: Passive reduced-order interconnect macromodeling algorithm. *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, 17(8):645–654, 1998.

[39] H. Panzer, J. Mohring, R. Eid, and B. Lohmann. Parametric model order reduction by matrix interpolation. *at-Automatisierungstechnik*, 8:475–484, 2010.

[40] L. T. Pillage and R. A. Rohrer. Asymptotic waveform evaluation for timing analysis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 9(4):352–366, 1990.

[41] J. Poldeman and J. Willems. *Introduction to Mathematical Systems Theory*. Springer Verlag, 1998.

[42] T. Reis and T. Stykel. Balanced truncation model reduction of second-order systems. *Mathematical and Computer Modeling of Dynamical Systems*, 14(5):391–406, 2008.

[43] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, 2nd edition, 2003.

[44] J. Saak, P. Kürschner, and P. Benner. A goal-oriented dual LRCF-ADI for balanced truncation. *Proceedings of the Vienna Mathmod Conference*, February 2012.

[45] W. Schilders. *Model Order Reduction*, chapter Introduction to MOR, pages 3–32. Springer, 2008.

[46] P. Scholz. *Analysis and numerical modeling of inductively coupled antenna systems*. Ph. D. thesis, Technische Universität Darmstadt, 2010.

[47] R. Schumann. *Die nichtorthogonale Finite-Integrations-Methode zur Simulation elektromagnetischer Felder*. Ph. D. thesis, Technische Universität Darmstadt, 1999.

[48] Y. Shamash. Linear system reduction using Padé approximation to allow retention of dominant modes. *International Journal of Control*, 21:257–272, 1975.

[49] G. Mouil Sil. *Model order reduction for efficient EMC simulation*. Ph. D. thesis, Technische Universität Darmstadt, 2010.

[50] R. D. Slone, R. Lee, and J. F. Lee. Multipoint Galerkin asymptotic waveform evaluation for model order reduction of frequency domain FEM electromagnetic radiation problems. *IEEE transactions on Antennas and Propagation*, 49(10):1504–1513, 2001.

[51] R. D. Slone, R. Lee, and J. F. Lee. Broadband model order reduction of polynomial matrix equation using single-point well-conditioned asymptotic waveform evaluation: Derivation and theory. *International Journal for Numerical Methods in Engineering*, 58:2325–2342, December 2003.

[52] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Linearization of parametric FIT-discretized systems for model order reduction. *Proceedings of the 13th Biennial IEEE Conference on Electromagnetic Field Computation (CEFC), Athens*, 2008.

[53] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Order reduction of linear time-invariant FIT-discretized structures with length and frequency parameterization. *6th International Conference on Engineering Computational Technology (ECT), Athens*, 5, 2008.

[54] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Ordnungsreduktion Linearisierter Frequenz- und Längenvariabler FIT-Modelle. *GMA-Fachausschüsse 1.30 und 1.40, Salzburg (GMA)*, 2008.

[55] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Linearization of parametric FIT-discretized systems for model order reduction. *IEEE Transactions on Magnetics*, 45(3):1380–1383, 2009.

[56] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Parameterized model order reduction using neighbouring subspaces. *Proceedings of the 8th Conference on Scientific Computing in Electrical Engineering (SCEE)*, pages 117–118, 2010.

[57] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Threedimensional geometry variations of FIT systems for model order reduction. *International Symposium on Electromagnetic Theory, Berlin (EMTS)*, 2010.

[58] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Model order reduction methods for multivariate parameterized dynamical systems. *Proceedings of the XXX URSI General Assembly and Scientific Computing of the International Union of Radio Science (URSI GASS 2011)*, pages 1–25, 2011.

[59] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Model order reduction methods for multivariate parameterized dynamical systems obtained by the finite integration theory. *Proceedings of the XXX URSI General Assembly and Scientific Computing of the International Union of Radio Science (URSI GASS 2011)*, pages 1–4, 2011.

[60] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. Model order reduction methods for systems with geometrical parameters. *accepted for the International Workshop on Finite Elements for Microwave Engineering*, 2012.

[61] K. Stavrakakis, T. Wittig, W. Ackermann, and T. Weiland. *Scientific Computing in Electrical Engineering*, chapter Parametric Model Order Reduction by Neighboring Subspaces. Springer, 2012.

[62] J. A. Stratton. *Electromagnetic Theory.* Mc Graw Hill Book company, 1941.

[63] P. Thoma. *Zur numerischen Lösung der Maxwellschen Gleichungen im Zeitbereich.* Ph. D. thesis, Technische Universität Darmstadt, 1997.

[64] B. Trapp. *Zur numerischen Berechnung hochfrequenter elektromagnetischer Felder auf der Basis von Eigenlösungen.* Ph. D. thesis, Technische Universität Darmstadt, 2002.

[65] L. N. Trefethen and D. Bau. *Numerical Linear Algebra.* SIAM, 1997.

[66] J. A. Tsemo, W. F. O. Müller, K. K. Stavrakakis, and T. Weiland. Pickup design with beta matching. *Proceedings IPAC, San Sebastián*, pages 1189–1191, September 2011.

[67] R. Unbehauen. *Systemtheorie.* Oldenbourg, 1983.

[68] C. D. Villemagne and R. E. Skelton. Model reduction using a projection formulation. *International Journal of Control*, 46:2141–2169, 1987.

[69] T. Weiland. Eine Methode zur Lösung der Maxwellschen Gleichungen für sechskomponentige Felder auf diskreter Basis. *Electronics and Communication (AEÜ)*, 31(3):116–120, March 1977.

[70] T. Weiland. A numerical method for the solution of the eigenwave problem of longitudinally homogenous waveguides. *Archiv für Elektronik und Übertragungstechnik (AEÜ)*, 31:308–314, 1977.

[71] T. Weiland. Time domain electromagnetic field computation with finite difference methods. *International Journal of Numerical Modeling*, 9:295–319, 1996.

[72] T. Weiland. *Softwarepraktikum zu Verfahren und Anwendungen der Feldsimulation*. Institut für Theorie elektromagnetischer Felder (TEMF), 2002.

[73] T. Weiland. Verfahren und Anwendungen der Feldsimulation. Skriptum zur Vorlesung, Technische Universität Darmstadt, 2010.

[74] D. S. Weile, E. Michielsen, E. Grimme, and K. Gallivan. A method for generating rational interpolant reduced order models of two-parameter linear systems. *Applied Mathematics Letters*, 12:93–102, 1999.

[75] T. Wittig. *Zur Reduzierung der Modellordnung in elektromagnetischen Feldsimulationen*. Ph. D. thesis, Technische Universität Darmstadt, 2003.

[76] T. Wittig, I. Munteanu, R. Schuhmann, and T. Weiland. Model order reduction with a two-step Lanczos algorithm. *IEEE Transactions on Magnetics*, 38(2):673–676, 2002.

[77] T. Wittig, R. Schuhmann, and T. Weiland. Model order reduction for large systems in computational electromagnetics. *Linear Algebra and its Applications*, 415(2):499–530, May 2006.

[78] T. Wolf, H. Panzer, and B. Lohmann. Sylvester equations and the factorization of the error system in Krylov subspace methods. *Proceedings of the Vienna Mathmod Conference*, February 2012.

[79] K.S. Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, AP-14:302–307, May 1966.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die in den letzten Jahren und speziell in der Phase des Niederschreibens zum Gelingen dieser Arbeit beigetragen haben. Insbesondere gilt mein Dank:

- Herrn Prof. Dr.-Ing. Thomas Weiland für die wissenschaftliche Betreuung und die ausgezeichneten Arbeitsbedingungen am Institut,
- Herrn Prof. Dr. Romanus Dyczij-Edlinger für die freundliche Übernahme des Korreferats,
- Herrn Dr.-Ing. Tilmann Wittig für die vielen wertvollen Diskussionen und Anmerkungen, die freundschaftliche Betreuung sowie die inhaltlichen Korrekturen der Arbeit,
- Herrn Dr.-Ing. Wolfgang Ackermann für die kompetente fachliche Betreuung sowie die Durchsicht und Korrektur des Manuskripts,
- Frau Marcia Bennett für sprachliche Korrekturen,
- allen meinen ehemaligen und jetzigen Kolleginnen und Kollegen am Institut für Theorie Elektromagnetischer Felder der TU-Darmstadt für die gute Zusammenarbeit,
- meinen Eltern und meiner Schwester für die bedingungslose Unterstützung in jeder Hinsicht in meinem gesamten Werdegang, sowie
- meiner Schwiegermutter Maria Tokatlidou für die Betreuung meiner Tochter während des Niederschreibens.

Ganz besonders möchte ich auch meinem Mann Giorgo Tokatlidi und unserem Sonnenschein Sophia danken für die Unterstützung und das Verständnis in den vergangenen Jahren. Danke!

Curriculum Vitae

Kynthia Katherina Stavrakakis
Geboren am 15.03.1982

2006-2012	Promotion, Institut für Theorie Elektromagnetischer Felder, Fachbereich Elektrotechnik und Informationstechnik, Technische Universität Darmstadt
2011	Fünfter Preis bei der zweiten internationalen URSI Students Paper Competition 2011
2008	Elternzeit
2002-2006	Diplom in Elektrotechnik und Informationstechnik, Vertiefungsrichtung Theoretische Elektrotechnik, Technische Universität Darmstadt
2002-2004	Grundstudium Mathematik, Technische Universität Darmstadt
1999-2002	Vordiplom in Elektrotechnik und Informationstechnik, Technische Universität Darmstadt
1996-1999	Apoletyrion (Hochschulreife), Lyzeum, Athen, Griechenland
1993-1996	Gymnasium, Athen, Griechenland
1987-1993	Grundschule, Athen, Griechenland