

## Parameterized model order reduction

***Citation for published version (APA):***

Ciuprina, G., Fernández Villena, J., Ioan, D., Ilievski, Z., Kula, S., Maten, ter, E. J. W., Mohaghegh, K., Pulch, R., Schilders, W. H. A., Silveira, L. M., Stefanescu, A., & Striebel, M. (2015). *Parameterized model order reduction*. (CASA-report; Vol. 1508). Technische Universiteit Eindhoven.

***Document status and date:***

Published: 01/01/2015

***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

***Take down policy***

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

**EINDHOVEN UNIVERSITY OF TECHNOLOGY**  
Department of Mathematics and Computer Science

CASA-Report 15-08  
February 2015

Parameterized model order reduction

by

G. Ciuprina, J. Fernández Villena, D. Ioan, Z. Ilievski, S. Kula, E.J.W. ter Maten,  
K. Mohaghegh, R. Pulch, W.H.A. Schilders, L. Miguel Silveira,  
A. Ștefănescu, M. Striebel



Centre for Analysis, Scientific computing and Applications  
Department of Mathematics and Computer Science  
Eindhoven University of Technology  
P.O. Box 513  
5600 MB Eindhoven, The Netherlands  
ISSN: 0926-4507

## Chapter 4

# Parameterized Model Order Reduction

Gabriela Ciuprina, Jorge Fernández Villena, Daniel Ioan, Zoran Ilievski, Sebastian Kula, E. Jan W. ter Maten, Kasra Mohaghegh, Roland Pulch, Wil H.A. Schilders, L. Miguel Silveira, Alexandra Ștefănescu and Michael Striebel

---

Gabriela Ciuprina and Daniel Ioan

Politehnica University of Bucharest, Spl.Independentei 313, 060042, Bucharest, Romania, e-mail: {Gabriela, Daniel}@lmi.pub.ro.

Jorge Fernández Villena and L. Miguel Silveira

INESC ID / IST - TU Lisbon, Rua Alves Redol 9, 1000-029 Lisbon, Portugal, e-mail: {Jorge.Fernandez, LMS}@inesc-id.pt

Zoran Ilievski

European Space & Technology Centre, Keplerlaan 1, P.O. Box 299, 2200 AG Noordwijk, the Netherlands, e-mail: ZoranI@gmail.com.

Sebastian Kula

Kazimierz Wielki University, Institute of Mechanics and Applied Computer Science, ul. Kopernika 1, 85-074 Bydgoszcz, Poland, e-mail: SKula@ukw.edu.pl.

E. Jan W. ter Maten

Chair of Applied Mathematics/Numerical Analysis, Bergische Universität Wuppertal, Gaußstraße 20, D-42119 Wuppertal, Germany, e-mail: Jan.ter.Maten@math.uni-wuppertal.de, and

Eindhoven University of Technology, Department of Mathematics and Computer Science, CASA, P.O.Box 513, 5600 MB Eindhoven, the Netherlands, e-mail: E.J.W.ter.Maten@tue.nl.

Kasra Mohaghegh

Multiscale in Mechanical and Biological Engineering (M2BE), Aragón Institute of Engineering Research (I3A), University of Zaragoza, María de Luna, 3, E-50018, Zaragoza, Spain, e-mail: Kasra@unizar.es.

Roland Pulch

Institut für Mathematik und Informatik, Ernst Moritz Arndt Universität Greifswald, Walther-Rathenau-Straße 47, D-17487 Greifswald, Germany, e-mail: PulchR@uni-greifswald.de.

Alexandra Ștefănescu

Research Centre of Excellence "Micro- and nanosystems for radiofrequency and photonics", IMT Bucharest – National Institute for Research and Development in Microtechnologies, 126A (32B), Erou Iancu Nicolae str., 077190 Bucharest, Romania, e-mail: Alexandra.Stefanescu@imt.ro.

Michael Striebel

ZF Lenksysteme GmbH, Richard-Bullinger-Straße 77, D-73527 Schwäbisch Gmünd, Germany, e-mail: Michael.Striebel@zf-lenksysteme.com.

**Abstract** This Chapter introduces parameterized, or parametric, Model Order Reduction (pMOR). The Sections are offered in a preferred order for reading, but can be read independently. Section 4.1, written by Jorge Fernández Villena, L. Miguel Silveira, Wil H.A. Schilders, Gabriela Ciuprina, Daniel Ioan and Sebastian Kula, overviews the basic principles for pMOR. Due to higher integration and increasing frequency-based effects, large, full Electromagnetic Models (EM) are needed for accurate prediction of the real behavior of integrated passives and interconnects. Furthermore, these structures are subject to parametric effects due to small variations of the geometric and physical properties of the inherent materials and manufacturing process. Accuracy requirements lead to huge models, which are expensive to simulate and this cost is increased when parameters and their effects are taken into account. This Section introduces the framework of pMOR, which aims at generating reduced models for systems depending on a set of parameters.

Section 4.2, written by Gabriela Ciuprina, Alexandra Ștefănescu, Sebastian Kula and Daniel Ioan, provides robust procedures for pMOR. This Section proposes a robust specialized technique to extract reduced parametric compact models, described as parametric SPICE-like netlists, for long interconnects modeled as transmission lines with several field effects such as skin effect and substrate losses. The technique uses an EM formulation based on partial differential equations (PDE), which is discretized to obtain a finite state space model. Next, a variability analysis of the geometrical data is carried out. Finally, a method to extract an equivalent parametric circuit is proposed.

Section 4.3, written by Michael Striebel, Roland Pulch, E. Jan W. ter Maten, Zoran Ilievski, and Wil H.A. Schilders, covers ways to efficiently determine sensitivity of output with respect to parameters. First direct and adjoint techniques are considered with forward and backward time integration, respectively. Here also the use of MOR via POD (Proper Orthogonal Decomposition) is discussed. Next, techniques in Uncertainty Quantification are described. Here pMOR techniques can be used efficiently.

Section 4.4, written by Kasra Mohaghegh, Roland Pulch and E. Jan W. ter Maten, provides a novel way in extending MOR to Differential-Algebraic Systems. Here new MOR techniques for reducing semi-explicit system of DAEs are introduced. These techniques are extendable to all linear DAEs. Especially pMOR techniques are exploited for singularly perturbed systems.

## 4.1 Parametric Model Order Reduction

<sup>1</sup>Model Order Reduction (MOR) techniques are a set of procedures which aim at replacing a large-scale model of a physical system by a lower dimensional model which exhibits similar behavior, typically measured in terms of its input-output re-

---

<sup>1</sup> Section 4.1 has been written by: Jorge Fernández Villena, L. Miguel Silveira, Wil H.A. Schilders, Gabriela Ciuprina, Daniel Ioan and Sebastian Kula. For additional topics and applications see also the Ph.D.-Thesis of the last author [19].

sponse. Reducing the order or dimension of these models, while guaranteeing that the input-output response is accurately captured, is crucial to enable the simulation and verification of large systems [1–3, 33]. Since the first attempts in this area [31], the methods for linear model reduction have greatly evolved and can be broadly characterized into two types: those that are based on subspace generation and projection methods [12, 27], and those based on balancing techniques [25, 28] (sometimes also referred to as Singular Value Decomposition (SVD)-based [2]). Hybrid techniques that try to combine some of the features of each family have also been presented [17, 18, 20, 30].

Although previously ignored when analyzing or simulating systems, parameter variability can no longer be disregarded as it directly impacts system behavior and performance. Accounting for the effects of manufacturing or operating variability, such as geometric parameters, temperature, etc., leads to parametric models whose complexity must be tackled both during the design and verification phases. For this purpose, *Parametric MOR* (pMOR, also known as *Parameterized MOR*) techniques that can handle parameterized descriptions are being considered as essential in the determination of correct system behavior. The systems generated by pMOR procedures must retain the ability to model the effects of both geometric and operating variability, in order to accurately predict behavior and optimize designs.

Several pMOR techniques have been developed for modeling large-scale parameterized systems. Although the first approaches were based on perturbation based techniques, such as [16, 24], the most common and effective ones appear to be extensions of the basic projection-based MOR algorithms [27, 30] to handle parameterized descriptions. An example of these are multiparameter moment-matching pMOR methods [8] which can generate accurate reduced models that capture both frequency and parameter dependence. The idea is to match, via different approaches, generalized moments of the parametric transfer function, and build an overall projector. Sample-based techniques have been proposed in order to contain the large growth in model order for multiparameter, high accuracy systems [29, 37]. They rely on sampling the joint multi-dimensional frequency and parameters space. This approach allows the inclusion of *a priori* knowledge of the parameter variation, and provides some error estimation. However, the issue of sample selection becomes particularly relevant when done in a potentially high-dimensional space.

### ***4.1.1 Representation of Parametric Systems***

In order to include parametric systems inside an efficient simulation flow, the parametric dependence should be explicit. This means that it must be possible to access the parameter values and modify them inside the same representation, while avoiding, if possible, re-computing the parametric systems, i.e. to perform another extraction.

Parameters usually affect the geometrical or electrical properties of the layout, and thus, most of these variations can be represented as modifications of the values

of the system matrices inside a state-space descriptor. For this reason, in most cases, the input and output ports are not affected by these variations (this of course depends on how the system is built), and in the case when they are in fact affected, these variations can be shifted to the inner states. The variability leads to a dependence of the extracted circuit elements on several parameters, of electrical or geometrical origin. This dependence results in a parametric state-space system representation, which in descriptor form can be written as

$$\begin{aligned} C(\lambda_1, \dots, \lambda_p) \dot{x}(\lambda_1, \dots, \lambda_p) + G(\lambda_1, \dots, \lambda_p) x(\lambda_1, \dots, \lambda_p) &= Bu, \\ y(\lambda_1, \dots, \lambda_p) &= Lx(\lambda_1, \dots, \lambda_p), \end{aligned} \quad (4.1)$$

where  $C, G \in \mathbb{R}^{n \times n}$  are again, respectively, the dynamic and static matrices,  $B \in \mathbb{R}^{n \times p}$  is the matrix that relates the input vector  $u \in \mathbb{R}^p$  to the inner states  $x \in \mathbb{R}^n$  and  $L \in \mathbb{R}^{q \times n}$  is the matrix that links those inner states to the outputs  $y \in \mathbb{R}^q$ . The elements of the matrices  $C$  and  $G$ , as well as the states of the system  $x$ , depend on a set of  $P$  parameters  $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_p]$  which model the effects of the mentioned uncertainty. This time-domain descriptor yields a parametric dependent frequency response modeled via the transfer function

$$H(s, \lambda_1, \dots, \lambda_p) = L (sC(\lambda_1, \dots, \lambda_p) + G(\lambda_1, \dots, \lambda_p))^{-1} B \quad (4.2)$$

for which we seek to generate a reduced order approximation, able to accurately capture the input-output behavior of the system for any point in the parameter space

$$\widehat{H}(s, \lambda_1, \dots, \lambda_p) = \widehat{L} (s\widehat{C}(\lambda_1, \dots, \lambda_p) + \widehat{G}(\lambda_1, \dots, \lambda_p))^{-1} \widehat{B}. \quad (4.3)$$

In general, one attempts to generate a reduced order model whose structure is, as much as possible, similar to the original, i.e. exhibiting a similar parametric dependence. The "de facto" standard used in most of the literature for representing a parametric system is based on a Taylor series expansion with respect to the parameters (shown here for first order in the frequency domain):

$$\begin{aligned} ((C_0 + C_1\lambda_1 + \dots + C_p\lambda_p)s + (G_0 + G_1\lambda_1 + \dots + G_p\lambda_p)) x(s, \lambda) &= Bu(s), \\ y(s, \lambda) &= Lx(s, \lambda), \end{aligned} \quad (4.4)$$

where  $G_0$  and  $C_0$  are the nominal values of the matrices, whereas  $G_i$  and  $C_i$  are the sensitivities with respect to the parameters. Novel extraction methodologies can efficiently generate such sensitivity information [5, 26].

A nice feature of this representation is that this explicit parameter dependence allows to obtain a reduced, yet similar representation when a projection scheme is applied

$$\begin{aligned} ((\widehat{C}_0 + \widehat{C}_1\lambda_1 + \dots + \widehat{C}_p\lambda_p)s + (\widehat{G}_0 + \widehat{G}_1\lambda_1 + \dots + \widehat{G}_p\lambda_p)) x(s, \lambda) &= \widehat{B}u(s), \\ y(s, \lambda) &= \widehat{L}x(s, \lambda), \end{aligned} \quad (4.5)$$

where  $\widehat{C}_i = V^T C_i V$ ,  $\widehat{G}_i = V^T G_i V$ ,  $\widehat{B} = V^T B$  and  $\widehat{L} = LV$ .

Some questions may be raised about the order necessary for an accurate representation of the parametric model. This depends on the range of variation and the effect of each parameter, and therefore is not trivial to ascertain.

However, some literature presents interesting results in this area [4, 6], with the conclusion that low order (first order in most cases) Taylor series are a good and useful approximation to the real parametric system. As it will be shown later, this statement has important consequences from the point of view of some parametric algorithms, especially those which rely on moment matching techniques.

### 4.1.2 Reduction of Parametric Systems

The most straight-forward approach for the reduction of such a parametric system is to apply nominal techniques. A first possibility is to apply nominal reduction methodologies on the perturbed system. This means that the model in (4.4) is evaluated for a set of parameter values. This model is no longer parametric, and thus standard reduction methodologies can be applied on it. However, once a "perturbed" system is evaluated and reduced, the parameter dependence is lost, and thus the result is a system which is no longer parametric, and therefore only accurate for a set of parameters.

A slightly different approach that overcomes this issue is to apply the projection on the Taylor series approximation. In this case, depending on the framework applied, we can distinguish two cases:

- First, in a projection methodology, the projector is computed from the nominal system, and later applied on the nominal and on the sensitivity matrices, obtaining a model as in (4.5).
- Second, in the case of Balanced Truncation realizations, the computation of the Gramians is done via the nominal system, but the balancing and the truncation is done both on the nominal matrices and on the sensitivities.

These methods, although not oriented to accurately capture the behavior of the system under variation of the parameters, can yield good approximations in cases of small variations or mild effect of the parameters. However, they are not reliable, and their performance heavily depends on the system.

#### 4.1.2.1 Perturbation Based Parametric Reduction

The first attempts to handle and reduce systems under variations were focused on perturbation techniques.

One of the earliest attempts to address this variational issue was to combine perturbation theory with moment matching MOR algorithms [24] into a **Perturbation-based Projector Fitting scheme**. To model the variational effects of the interconnects, an affine model was built for the capacitance and conductance matrices,

$$\begin{aligned} G(\lambda_1, \dots, \lambda_P) &= G_0 + \lambda_1 G_1 + \dots + \lambda_P G_P, \\ C(\lambda_1, \dots, \lambda_P) &= C_0 + \lambda_1 C_1 + \dots + \lambda_P C_P, \end{aligned} \quad (4.6)$$

where now  $G_0$  and  $C_0$  are the nominal matrix values, i.e. the value of the matrices under no parameter variation, and  $C_i$  and  $G_i$ ,  $i = 1, \dots, P$ , are their sensitivities with respect to those parameters. For small parameter variations, the projection matrix obtained via a moment-matching type algorithm such as PRIMA also may show small perturbations. To capture such effect, several samples in the parameter space were drawn  $G(\lambda_1, \dots, \lambda_P)$  and  $C(\lambda_1, \dots, \lambda_P)$ , and for each sample PRIMA was applied resulting a projector. A fitting methodology was later applied in order to determine the coefficients of a parameter dependent projection matrix

$$V(\lambda_1, \dots, \lambda_P) = V_0 + \lambda_1 V_1 + \dots + \lambda_P V_P. \quad (4.7)$$

To obtain a reduced model, both the parametric system and the projector are evaluated with the parameter set. Projection is applied and the reduced model obtained. However, this reduced model is only valid for the used parameter set. If a reduced model for a different parameter set is needed, the evaluation and projection must be applied again, what makes hard to include this method in a simulation environment.

Another method combined perturbation theory with the Truncated Balanced Realization (TBR) [25, 28] framework. A perturbation matrix was theoretically obtained starting from the affine models shown in (4.6) [16]. This matrix was applied via a congruence transformation over the Gramians to address the variability, obtaining a set of perturbed Gramians. These in turn were used inside a Balancing Truncation procedure. As with most TBR-inspired methods, this one is also expensive to compute and hard to implement. The above methods have obvious drawbacks, perhaps the most glaring of which is the heavy computation cost required for obtaining the reduced models and the limitation that comes from perturbation based approximations, possibly leading to inaccuracy in certain cases.

#### 4.1.2.2 Multi-Dimensional Moment Matching

Most of the techniques in the literature extend the moment matching paradigm [12, 27, 34] to the multi-dimensional case. They usually rely on the implicit or explicit matching of the moments of the parametric transfer function (4.2). These moments depend not only on the frequency, but on the set of parameters affecting the system, and thus are denoted as multi-dimensional or multi-parameter moments.

This family of algorithms assume that a model based on the Taylor Series expansion can be used for approximating the behavior of the conductance and capacitance,  $G(\lambda)$  and  $C(\lambda)$ , expressed as a function of the parameters

$$\begin{aligned} G(\lambda_1, \dots, \lambda_P) &= \sum_{i_1=0}^{\infty} \dots \sum_{i_P=0}^{\infty} G_{i_1, \dots, i_P} \lambda_1^{i_1} \dots \lambda_P^{i_P}, \\ C(\lambda_1, \dots, \lambda_P) &= \sum_{i_1=0}^{\infty} \dots \sum_{i_P=0}^{\infty} C_{i_1, \dots, i_P} \lambda_1^{i_1} \dots \lambda_P^{i_P}, \end{aligned} \quad (4.8)$$



where  $G_{i_1, \dots, i_p}$  and  $C_{i_1, \dots, i_p}$  are the multidimensional Taylor series coefficients. This Taylor series can be extended up to the desired (or required) order, including cross derivatives, for the sake of accuracy. If this formulation is used, the structure for parameter dependence may be maintained if the projection is not only applied to the nominal matrices, but to the sensitivities as well.

Multiple methodologies follow these basic premises, but they differ in how and which such moments are generated and used in the projection stage.

The **Multi-Parameter Moment Matching** method [8] relies on a single-point expansion of the transfer function (4.2) in the joint space of the frequency  $s$  and the parameters  $\lambda_1, \dots, \lambda_p$ , in order to obtain a power series in several variables,

$$x(s, \lambda_1, \dots, \lambda_p) = \sum_{k=0}^{\infty} \sum_{k_s=0}^k \sum_{k_1=0}^{k-k_s} \dots \sum_{k_p=0}^{k-k_s-k_1-\dots-k_{p-1}} M_{k, k_s, k_1, \dots, k_p} s^{k_s} \lambda_1^{k_1} \dots \lambda_p^{k_p}, \quad (4.9)$$

where  $M_{k, k_s, k_1, \dots, k_p}$  is a  $k$ -th ( $k = k_s + k_1 + \dots + k_p$ ) order multi-parameter moment corresponding to the coefficient term  $s^{k_s} \lambda_1^{k_1} \dots \lambda_p^{k_p}$ .

A basis for the subspace spanned from these moments can be built and the resulting orthonormal basis  $V$  can be used as a projection matrix for reducing the original system

$$\text{colspan}V = \text{colspan}\{M_{00\dots 0}, \dots, M_{k, k_s, k_1, \dots, k_p}\}. \quad (4.10)$$

This parametrized reduced model matches up to the  $k$ -th order multi-parameter moment of the original system.

However, the main inefficiencies of this method are twofold:

- On one hand, this method generates pure multi-dimensional moments (see equation (4.9)), which means that the number of moments grows dramatically (all the possible combinations for a given order must be done) when the number of parameters is increased, even for a modest number of moments for each parameter. For this reason, the reduced model size grows exponentially with the number of parameters and the moments to match.
- On the other hand, the process parameters fluctuate in a small range around their nominal value, whereas the frequency range is much larger, and a higher number of moments are necessary in order to capture the global response for the whole frequency range. This algorithm treats the frequency as one parameter more, which turns to be highly inefficient.

An improvement of the previous approach is to perform a **Low-Rank Approximation** of the multi-dimensional moments [21]. An SVD-based low-rank approximation of the generalized moments,  $G^{-1}G_i$  and  $G^{-1}C_i$  (which are related to the multidimensional moments), is applied. Then, separate subspaces are built from these low-rank approximations for every parameter. The global projector is obtained from the orthonormalization of the nominal moments (computed via Arnoldi for example), and the moments of the subspaces related to the parameters. The projector is applied on the Taylor Series approximation to obtain a reduced parametric model. This approach, although providing more flexibility and improving the matching, re-

quires the low-rank SVD of the generalized moments, which comes at a cost of  $O(n^3)$ , i.e., limiting its applicability to small-medium size problems.

A different multi-dimensional moment matching approach was also presented in [15], called **Passive Parameterized Time-Domain Macro Models**. It relies on the computation of several subspaces, built separately for each dimension, i.e. the frequency  $s$  (to which respect  $k_s$  block moments are obtained in a basis denoted as  $Q_s$ ) and the parameter set  $\lambda$  (generating the basis  $Q_i$  which match  $k_{\lambda_i}$  block moments with respect to parameter  $\lambda_i$ ). These independent subspaces can be efficiently computed using standard nominal approaches, e.g. PRIMA. Once all the subspaces have been computed, an orthonormal basis can be obtained so that its columns span the joint of all subspaces. For example, in the affine Taylor Series representation, using Krylov spaces  $Kr(A, B, k)$  (matrix  $A$ , multi-columns vector  $B$ , moments  $k$ ):

$$\begin{aligned} \text{colsp}\{Q_s\} &\equiv Kr\{A, R, k_s\} \quad \text{with} \quad \begin{cases} A = -G^{-1}C, \\ R = G^{-1}B \end{cases} \\ \text{colsp}\{Q_i\} &\equiv Kr\{A_i, R_i, k_i\} \quad \text{with} \quad \begin{cases} A_i = -(G + sC)^{-1}(G_i + sC_i), \\ R_i = -(G + sC)^{-1}B \end{cases}, \quad (4.11) \\ V &= QR [Q_s \ Q_1 \ \dots \ Q_i \ \dots \ Q_p], \end{aligned}$$

where subscript  $i$  refers to the  $i$ -th parameter  $\lambda_i$ , and the parameter related moments have been generalized to any shifted frequency  $s$ .  $QR$  stands for the  $QR$ -factorization based orthonormalization. Applying the resulting matrix  $V$  in a projection scheme ensures that the parametric Reduced Order Model matches  $k_s$  moments of the original system with respect to the frequency, and  $k_i$  moments with respect to the parameter  $\lambda_i$ . If the cross-term moments are needed for accuracy reasons, the subspace that spans these moments can also be included by following the same scheme.

A different approach is explored in **CORE** [22]. Here an explicit moment matching with respect to the parameters is first done, via Taylor-series expansion, followed by an implicit moment matching in frequency (via projection). The first step is done by expanding the state space vector  $x$  and the matrices  $G$  and  $C$  in its Taylor Series only with respect to the parameters,

$$x(s, \lambda) = \sum_{i_1=0}^{\infty} \dots \sum_{i_p=0}^{\infty} x_{i_1, \dots, i_p}(s) \lambda_1^{i_1} \dots \lambda_p^{i_p}, \quad (4.12)$$

$$\begin{aligned} G(\lambda) &= \sum_{i_1=0}^{\infty} \dots \sum_{i_p=0}^{\infty} G_{i_1, \dots, i_p} \lambda_1^{i_1} \dots \lambda_p^{i_p}, \\ C(\lambda) &= \sum_{i_1=0}^{\infty} \dots \sum_{i_p=0}^{\infty} C_{i_1, \dots, i_p} \lambda_1^{i_1} \dots \lambda_p^{i_p}, \end{aligned} \quad (4.13)$$

where  $G_{0, \dots, 0}, C_{0, \dots, 0}$  and  $x_{0, \dots, 0}(s)$  are the nominal values for the matrices and the states vector, respectively. The remaining  $G_{i_1, \dots, i_p}, C_{i_1, \dots, i_p}$  and  $x_{i_1, \dots, i_p}$  are the sensitivities with respect to the parameters. Explicitly matching the coefficients of the same powers leads to an augmented system, in which the parametric dependence is shifted to the output related matrix  $L_A$ :

$$\begin{aligned}
C_A &= \begin{bmatrix} C_0 & & & & \\ C_1 & C_0 & & & \\ \vdots & 0 & C_0 & & \\ C_i & 0 & 0 & C_0 & \\ \vdots & & & & \ddots \end{bmatrix}, & B_A &= \begin{bmatrix} B \\ 0 \\ \vdots \\ 0 \\ \vdots \end{bmatrix}, & x_A &= \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_i \\ \vdots \end{bmatrix}, \\
G_A &= \begin{bmatrix} G_0 & & & & \\ G_1 & G_0 & & & \\ \vdots & 0 & G_0 & & \\ G_i & 0 & 0 & G_0 & \\ \vdots & & & & \ddots \end{bmatrix}, & L_A &= [L \ \lambda_1 L \ \cdots \ \lambda_i L \ \cdots].
\end{aligned} \tag{4.14}$$

The second step applies a typical nominal moment matching procedure (e.g. PRIMA [27]) to reduce this augmented system. This is possible because the matrices  $G_A$ ,  $C_A$  and  $B_A$  used to build the projector do not depend on the parameters. The projector is latter applied on all the matrices of the augmented system in (4.14). Furthermore, the Block Lower Triangular structure of the system matrices  $G_A$  and  $C_A$  can be exploited in recursive algorithms to speed-up the reduction stage. This two-step approach allows to increase the number of the matched multi-parameter moments with respect to other techniques, for a similar reduced order. In principle, in spite of the larger size of the augmented model, the order of the reduced system can be much smaller than in the previous cases. On the other hand, the structure of the dependence with respect to the parameters is lost, since the parametric dependence is shifted to the later projected output related  $L$  matrix. The projection mixes all the parameters, losing the possibility of modifying them without need of recomputation. This method also has the disadvantage that the explicit computation of the moments with respect to the parameters can lead to numerical instabilities. The method, although stability-preserving, is unable to guarantee passivity preservation.

Some algorithms [23, 37] try to match the same moments as CORE, but in a more numerical stable and efficient fashion, using **Recursive and Stochastic Approaches**. They generalize the CORE paradigm up to an arbitrary expansion order with respect to the parameters, and apply an iterative procedure in order to compute the frequency moments related to the nominal matrices, and the ones obtained from the parametric part (this means, to obtain a basis for each block of states  $x_i$  in (4.14), but without building such system).

$$\begin{aligned}
\text{colspan}\{V_0\} &\equiv Kr\{A, R, q_0\} = [V_0^0 \ V_0^1 \ \cdots \ V_0^{q-1}], \\
\text{with } A &= -(G_0 + s_k C_0)^{-1} C_0, \quad R = (G_0 + s_k C_0)^{-1} B, \\
\text{colspan}\{V_i\} &= [V_i^0 \ V_i^1 \ \cdots \ V_i^j \ \cdots], \\
\text{with } V_i^j &= -(G_0 + s_k C_0)^{-1} (G_i V_0^j + s_k C_i V_0^{j-1} + C_0 V_i^{j-1}), \\
G_i &= G_0 \dots 010 \dots 0, \\
C_i &= C_0 \dots 010 \dots 0,
\end{aligned} \tag{4.15}$$

where  $s_k$  is the expansion point for the Krylov subspace generation, and  $V_i^j$  is the  $j$ -th moment with respect to the frequency for the  $i$ -th parameter. This general recursive scheme, here presented for first order with respect to the parameters, can be extended to any (independent) order with respect to each parameter.

The technique in [37] uses a tree graph scheme, in which each node is associated to a moment, and the branches represent recursive dependences among moments. Each tree level contains all the moments of the same multi-parameter order. On this tree, a random sampling approach is used to select and generate some representative moments, preventing the exponential growth.

On the other hand, the technique in [23] advocates for an exhaustive computation at each parameter order. This means that all the moments for zero-parameter order (i.e. nominal), are computed until no rank is added. The same procedure is repeated for first order with respect to all parameters. If the model is not accurate, more order with respect to the parameters can be added.

Notice that both schemes provide a large degree of flexibility, as different orders with respect to each parameter and with respect to the frequency can be applied. In both approaches, the set of all the moments generated is orthonormalized, so an overall projector is obtained. This is used inside a congruence transformation on the Taylor Series approximation (4.4), to generate a reduced model in the same representation. Another advantage of these methodologies is that the passivity is PRIMA-like preserved, and the basis is built in a numerical stable fashion.

#### 4.1.2.3 Multi-Dimensional Sampling

Another option present in the literature relies on sampling schemes for capturing the variational nature of the parametric model. They are applied for the building of a projector to later apply congruence transformation on the original model.

A simple generalization of the multi-point moment matching framework [11] to a multi-dimensional space can be done via **Variational Multi-Point Moment Matching**. Small research has been devoted to this family of approaches, but one algorithm can be found in [21]. The flexibility it provides is also one of its main drawbacks, as the methodology can be applied in a variety of schemes, from a single-frequency multi-parameter sampling to a pure multi-dimensional sampling. From these expansion points, several moments are computed following a typical moment matching scheme. The orthonormalization of the set of moments provides the overall projector which is applied in a congruence reduction scheme. However, it is hard to determine the number and placement of samples, and the number of moments to match with respect to the frequency and to the parameters.

Another scheme, which overcomes some of the issues of the previous approach is the **Variational Poor Man's TBR** [29]. This approach is based on the statistical interpretation of the algorithm (see [30] for details) and enhances its applicability to multiple dimensions. In this interpretation, the Gramian  $X_\lambda$  is seen as a covariance matrix for a Gaussian variable  $x_\lambda$ , obtained by exciting the (presumed stable) system with  $u$  involving white noise. Rewriting the Gramian as

$$X_\lambda = \int_{S_\lambda} \int_{-\infty}^{\infty} (j\omega C_\lambda + G_\lambda)^{-1} B B^T (j\omega C_\lambda + G_\lambda)^{-H} p(\lambda) d\omega d\lambda, \quad (4.16)$$

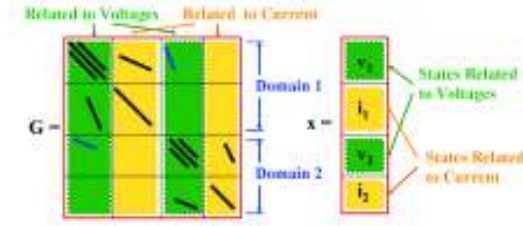
where  $p(\lambda)$  is the probability density of  $\lambda$  in the parameter space,  $S_\lambda$ . Just as in PMTBR, a quadrature rule can be applied in the overall parameter plus frequency space to approximate the Gramian via numerical computation. But in this case the weights are chosen taking into account the *Probability Density Function* (PDF) of  $\lambda_i$  and the frequency constraints. This can be generalized to a set of parameters, where a joint PDF of all the parameters can be applied on the overall parameter space, or the individual PDF of each parameter can be used. This possibility represents an interesting advantage, since *a-priori* knowledge of the parameters and the frequency can be included in order to constrain the sampling and yield a more accurate reduced model. The result of this approach is an algorithm which generates Reduced Order Models whose size is less dependent on the number of parameters. In the deterministic case, an error analysis and control can be included, via the eigenvalues of the SVD. However, in the variational case only an expected error bound can be given:

$$E\{\|\hat{x}_\lambda(0) - x_\lambda(0)\|_2^2\} \leq \sum_{i=r+1}^n \sigma_i^2, \quad (4.17)$$

where  $r$  is the reduced order and  $n$  the original number of states. On the other hand, in this method the issue of sample selection, already an important one in the deterministic version, becomes even more relevant, since the sampling must now be done in a potentially much higher-dimensional space.

### 4.1.3 Practical Consideration and Structure Preservation

Inside the pMOR realm, the moment matching algorithms based on single point expansion may not be able to capture the complete behavior along the large frequency range required for common RF systems, and may lead to excessively large models if many parameters are taken into account. Therefore the most suitable techniques for the reduction seem to be the multipoint ones. Among those techniques, Variational PMTBR [29] offers a reliable framework with some interesting features that can be exploited, such as the inclusion of probabilistic information and the trade off between size and error, which allows for some control of the error via analysis of the singular values related to the dropped vectors. On the other hand, it requires a higher computational effort than the multi-dimensional moment matching approaches, as it is based on multidimensional sampling schemes and Singular Value Decomposition (SVD), but the compression ratio and reliability that it offers compensates this drawback. The effort spent in the generation of such models can be amortized when the reduced order model generated is going to be used multiple times. This is usually the case for parametric models, as the designer may require several evaluations for different parameter sets (e.g. in the case of Monte Carlo simulations, or optimization steps). Furthermore, this technique offers some extra



**Fig. 4.1** Two-level hierarchy: Domain level (given by the numbers, 1 and 2) and Variable level (voltages  $v_k$  and currents  $i_k$ ).

advantages when combined with block structured systems [14], such as the block-wise error control with respect to the global input-output behaviour, which can be applied to improve the efficiency of the reduction. This means that each block can be reduced to a different order depending on its relevance in the global response.

An important point to recall here is that the block division may not reflect different sub-domains. Different sub-divisions can be done to address different hierarchical levels. For instance, it may be interesting to divide the complete set in sub-domains connected by hooks, which generates a block structured matricial representation. But inside each block corresponding to a sub-domain, another block division may be done, corresponding either to smaller sub-domains or to a division related to the different kind of variables used to model each domain (for example, in a simple case, currents and voltages). This variable related block structure preservation has already been advocated in the literature [13] and may help the synthesis of and equivalent SPICE-like circuit [35] in the case that is required. Figure 4.1 presents a more intuitive depiction of the previous statements, in which a two domain example is shown with its hierarchy, and each domain has also some inner hierarchy related to the different kind of variables (in this case, voltages and currents, but it can also be related to the electric and magnetic variables, depending on the formulation and method used for the generation of the system matrices).

The proposed flow starts from a parametric state-space descriptor, such as (4.1), which exhibits a multi-level hierarchy, and a block parametric dependence (as different parameters may affect different sub-domains). The matrices of size  $n$  have  $K$  domains, each with size  $n_i$ ,  $n = \sum_i n_i$ . For instance, for the static part,

$$G = \begin{bmatrix} G_{11}(\lambda_{\{11\}}) & \dots & G_{1K}(\lambda_{\{1K\}}) \\ \vdots & \ddots & \vdots \\ G_{K1}(\lambda_{\{K1\}}) & \dots & G_{KK}(\lambda_{\{KK\}}) \end{bmatrix}, \quad (4.18)$$

where  $\lambda_{\{ij\}}$  is the set of parameters affecting  $G_{ij} \in \mathbb{R}^{n_i \times n_j}$ . Then we perform the multi-dimensional sampling, both in the frequency and the parameter space. For each point we generate a matrix or vector  $z_j$  (a matrix in case  $B$  includes multiple inputs)

$$z_j = (C(\lambda_j)s_j + G(\lambda_j))^{-1} B, \quad (4.19)$$

where  $C(\lambda)$  and  $G(\lambda)$  are the global matrices of the complete domain, with  $n$  degrees of freedom (dofs). To generate the matrix  $z_j \in \mathbb{R}^{n \times m}$ , with  $m$  the number of global ports, we can apply a direct procedure, meaning a factorization (at cost  $O(n^\beta)$ , with  $1.1 \leq \beta \leq 1.5$  for sparse matrices) and a solve (at cost  $O(n^\alpha)$ , with  $1 \leq \alpha \leq 1.2$  for sparse matrices). Novel sparse factorization schemes can be applied to improve the efficiency [9, 10]. In cases when a direct method may be too expensive iterative procedures may be used [32].

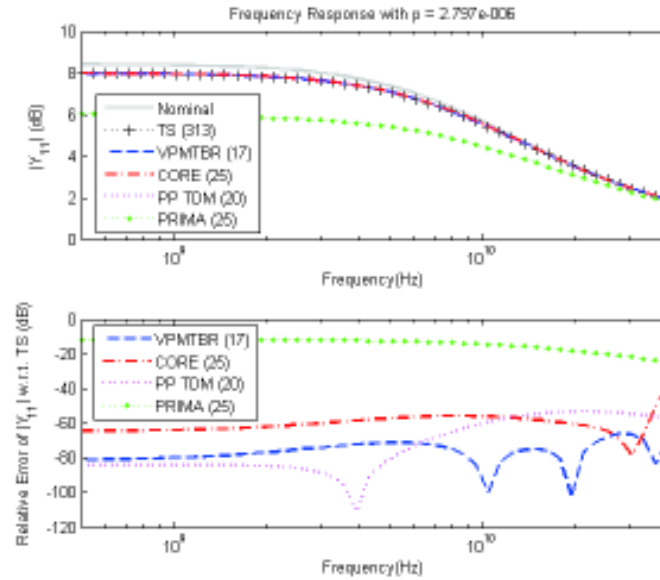
The choice of the sampling points may be an issue, as there is no clear scheme or procedure that is known to provide an optimal solution. However, as stated in [29], the accuracy of the method does not depend on the accuracy of the quadrature (and thus in the sampling scheme), but on the subspace generated. For this reason, a good sampling scheme is to perform samples in the frequency for the nominal system, and around these nominal samples, perform some parametric random sampling in order to capture the vectors that the perturbed system generates. The reasoning behind this scheme is that for small variations, such as the ones resulting from process parameters, the subspace generated along the frequency is generally more dominant than the one generated by the parameters. In addition, under small variations, the nominal sampling can be used as a good initial guess for an iterative solver to generate the parametric samples. For the direct solution scheme, to generate  $P$  samples (and thus  $Pm$  vectors) for the global system has a cost of  $O(Pn^\alpha + Pn^\beta)$ . Note that since  $m$  is the number of global (or external) ports, the number of vectors is smaller than if we take all the hooks into account.

The next step is the orthonormalization, via SVD, of the  $Pm$  vectors for generating a basis of the subspace in which to project the matrices. Here an **independent basis**  $V_i$ ,  $i \in \{1, \dots, K\}$ , can be generated for each  $i$ -th sub-domain. To this end the columns in  $z_j$  are split according to the block structure present in the system matrices (i.e., the  $n_i$  rows for each block), and an SVD is performed on each of these set of vectors, at a cost of  $O(n_i(Pm)^2)$ , where  $n_i$  is the size of the corresponding block, and  $n = \sum_i n_i$ . For each block, the independent SVD allows to drop the vectors less relevant for the global response (estimated by the dropped singular value ratio, as presented in [29]). This step generates a set of projectors,  $V_i \in \mathbb{R}^{n_i \times q_i}$ , with  $q_i \ll n_i$  the reduced size for the  $i$ -th block of the global system matrix. These projectors can be placed in the diagonal blocks of an overall projector, that can be used for reducing the initial global matrices to an order  $q = \sum_i q_i$ . This block diagonal projector allows a block structure (and thus sub-domain) preservation, increasing the sparsity of the ROM with respect to that of the standard projection. This sparsity increase is particularly noticeable in the case of the sensitivities (if a Taylor series is used as base representation), as the block parameter dependence is maintained (e.g. in the static matrix)

$$\hat{G}_{ij}(\lambda_{\{ij\}}) = V_i^T G_{ij}(\lambda_{\{ij\}}) V_j. \quad (4.20)$$

The total cost for the procedure can be approximated by

$$O(Pn^\alpha + Pn^\beta + (Pm)^2 \sum_i n_i). \quad (4.21)$$



**Fig. 4.2** (Top:) Frequency response of the L-shape example. The original, both the nominal and the Taylor series for a fixed parameter value, of order 313, and the reductions via PRIMA, CORE, Passive Parameterized Time-Domain Macro Models (PP TDM), and Variational PMTBR (VPMTBR), of different orders. (Bottom:) Relative error of the reduction models with respect to the original Taylor series approximation.

## 4.1.4 Examples

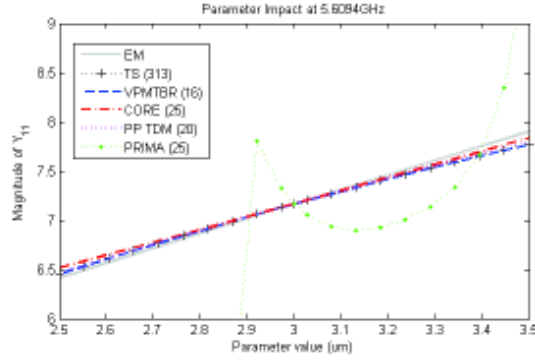
### 4.1.4.1 L-shape

As a first example we present a simple L-shape interconnect structure depending on the width of the metal layer. Figure 4.2 shows the frequency response for a fixed parameter value, of the nominal system, the Taylor series approximation (both of order 313), and the reduction models obtained with several parametric approaches:

- Nominal reduction of the Taylor Series, via PRIMA, of order 25,
- Multi-dimensional moment matching, via CORE, of order 25,
- Multi-dimensional moment matching, via Passive Parameterized Time-Domain Macro Models technique, of order 20,
- and Multi-dimensional sampling, via Variational PMTBR, of order 16.

Figure 4.3 shows the same example, but in this case the response of the systems with respect of the parameter variation, for a given frequency point. It is clear that the parametric Model Order Reduction techniques are able to capture the parametric behavior, whereas the nominal approach (PRIMA) fails to do so, even for high order.





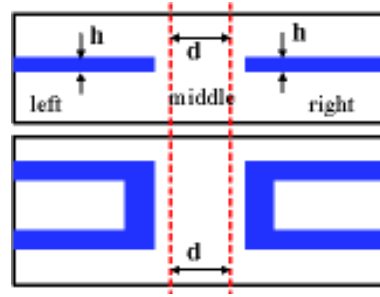
**Fig. 4.3** Parameter impact on the response of the L-shape example. The EM model for several parameter values (of order 313), the Taylor series approximation (of order 313 as well), and the reductions via PRIMA, CORE, PP TDM, and VPMTBR, of different orders.

**Table 4.1** Characteristics of the Examples

Ex	Domain	Dofs	Terminals (EH,MH,IT)	ROM Dofs
U-shape	Left	785	77 (42, 34, 1)	85
	Middle	645	152 (84, 68, 0)	90
	Right	785	77 (42, 34, 1)	85
	Complete	2215	2 (0, 0, 2)	260
Double Spiral	$Var_1$	49125	2 (0, 0, 2)	142
	$Var_2$	54977	2 (0, 0, 2)	165
	Complete	104102	2 (0, 0, 2)	307

#### 4.1.4.2 U-Coupled

This is a simple test case, which has two U-shape conductors; each of the conductors ends represent one port, having one terminal voltage excited (intentional terminal, IT) and one terminal connected to ground. A clear illustration of the setting is given by Figure 4.4. The distance ( $d$ ) separating the conductors and the thickness ( $h$ ) of the corresponding metal layer are parameterized. The complete domain is partitioned into three sub-domains, each of them connected to the others via a set of hooks (both electric, EH, and magnetic, MH). The domain hierarchy and parameter dependence are kept after the reduction, via Block Structure Preserving approaches. The Full Wave EM model was obtained via Finite Integration Technique (FIT) [7], and its matrices present a Block Structure that follows the domain partitioning. Table 4.1 shows the characteristics of the original system. Each sub-domain is affected by a parameter. The left and right sub-domains contain the conductors, and thus are affected by the metal thickness  $h$ . The middle domain width varies with the distance between the two conductors, and thus is affected by parameter  $d$ . For each param-



**Fig. 4.4** Topology of the U-shape: (Up) cross view, (Down) top view. Parameters: distance between conductors,  $d$ , and thickness of the metal,  $h$ .

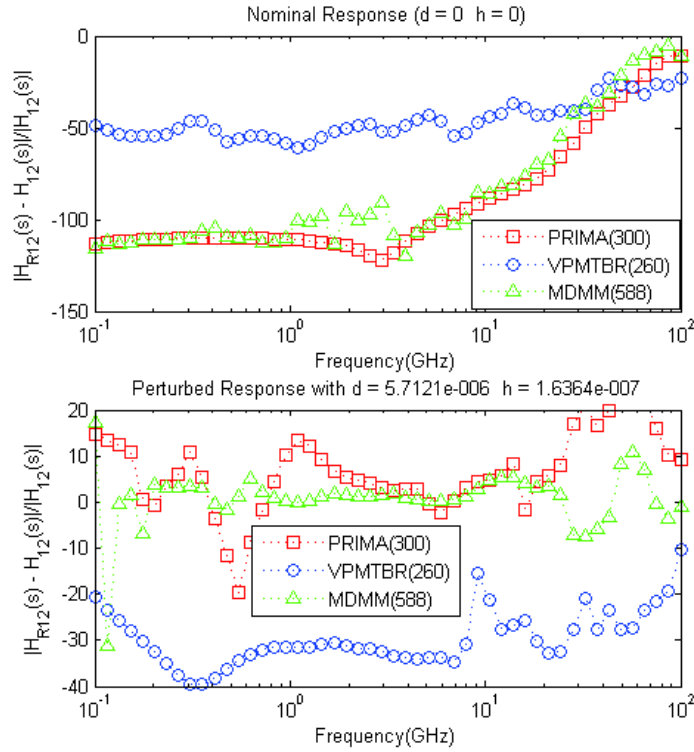
eter the first order sensitivity is taken into account, and a first order Taylor Series (TS) formulation is taken as the original system.

For the reduction we apply three techniques. First, a Nominal Block Structure Preserving (BSP) PRIMA [36], with a single expansion point and matching 50 moments, is applied. This leads to a 100-vector generated basis, that after BSP expansion produces a 300-dofs Reduced Order Model (ROM). Second, a BSP procedure coupled with a Multi-Dimensional Moment Matching (MDMM) approach [15], is tried. The basis will match 40 moments with respect to the frequency, and 30 moments with respect to each parameter. The orthonormalized basis has 196 vectors, that span a BSP ROM of size 588. Third, the proposed BSP VPMTBR, with 60 multidimensional samples, and a relative tolerance of 0.001 for each block, is studied. This process generates different reduced sizes for each block: 85, 90 and 85, with a global size of 260.

Figure 4.5 shows the relative error in the frequency transfer function at a parameter set point for the three ROMs w.r.t. the Taylor series. PRIMA and MDMM approaches fail to capture the behavior with the order set, but the proposed approach performs much better even for a lower order. Figure 4.6 shows the response change with the variation of parameter  $d$  at a single frequency point (Parameter Impact). PRIMA and MDMM only present accuracy for the nominal point, whereas the proposed method maintains the accuracy for the parameter range.

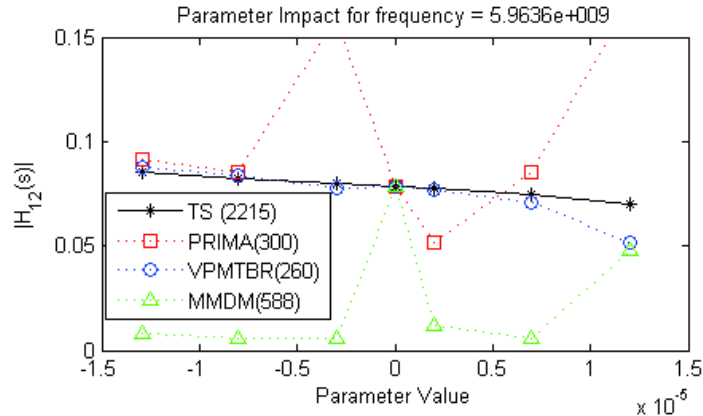
#### 4.1.4.3 Double Spiral

This is an industrial example, composed by two square integrated spiral inductors in the same configuration as the previous example (See Figure 4.7). The complete domain has two ports, and 104102 Dofs. The example also depends on the same two parameters, the distance  $d$  between spirals, and the thickness  $h$  of the corresponding metal layer. In this case a single domain is used, but the BSP approach is applied on the inner structure provided by the different variables in the FIT method (electric and magnetic grid). For the reduction, the proposed BSP VPMTBR methodology is benchmarked against a nominal BSP PRIMA (400 dofs) methodology, and com-

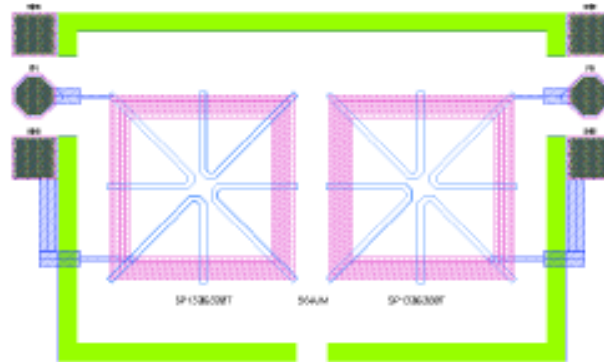


**Fig. 4.5** U-Coupled: Relative Error (dB) in  $|H_{12}(s)|$  for (Up) the nominal response, and (Down) the perturbed response at a single parameter set. The curves represent: BSP PRIMA, BSP VPMTBR, and BSP MDMM.

pared with the original Taylor Series formulation. The ROM size in this case is 142 and 165 respectively for the blocks. The results are presented in the Figures 4.8 and 4.9. Figure 4.8 shows the frequency relative error of the ROMs with respect to the original Taylor Series. PRIMA, although accurate for the nominal response, fails to capture the parametric behavior, whereas the proposed method succeeds in modeling such behavior. This is also the conclusion that can be drawn from the parameter impact in Figure 4.9.



**Fig. 4.6** U-Coupled: Variation of  $|H_{12}|$  vs. the variation of the parameter  $d$  at 59.6 GHz for the original TS and the three BSP ROMs.

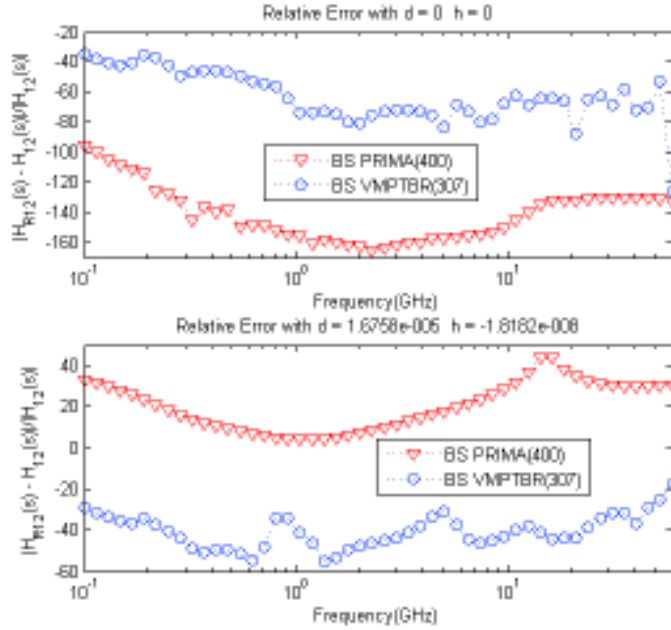


**Fig. 4.7** Layout configuration of the Double Spiral example (view from the top).

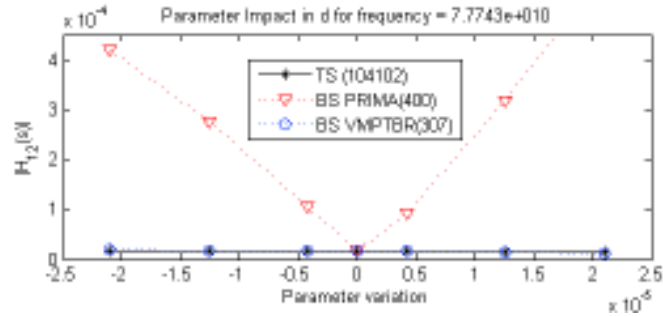
### 4.1.5 Conclusions

We conclude that Parametric Model Order Reduction techniques are essential for addressing parameter variability in the simulation of large dynamical systems. Representation of the state space based on Taylor series expansion with respect to the parameters provide the flexibility and accuracy required by efficient simulation. This representation approach can be combined with projection-based methods to generate structural equivalent reduced models.

Single-point based moment-matching approaches are suitable for small variations and local approximations, but usually suffer from several drawbacks when applied to EM based models operating in a wide frequency range. Multi-point based approaches, although computationally more expensive, are more reliable and generate more compressed models. Thus, the generation cost can be amortized in the simulation stages.



**Fig. 4.8** Double Spiral: Relative Error (dB) in  $|H_{12}(s)|$  for (Up) the nominal response, and (Down) the perturbed response at a single parameter set.



**Fig. 4.9** Double Spiral:  $|H_{12}|$  vs. the variation of the parameter  $d$  at a frequency point for the original TS and the ROMs: PRIMA, and VPMTBR.

Combination of the projection methodologies with Block Structure Preserving approaches can be done efficiently in parametric environments. Further advantages can be obtained in this case, such as different compression order for each block based on its relevance in the global behavior, higher degree of sparsification of the nominal matrices, and in particular, of the sensitivities, and the maintenance of the block domain hierarchy and block parameter dependence after reduction.

## References for Section 4.1

1. ACHAR, R., NAKHLA, M.S.: Simulation of high-speed interconnects. *Proceedings of the IEEE*, 89(5), pp. 693–728 (2001)
2. ANTOUNAS, A.C.: *Approximation of Large-Scale Dynamical Systems*. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2005)
3. BENNER, P., MEHRMANN, V., SORENSEN, D.C. (eds.): *Dimension Reduction of Large-Scale Systems*. Lecture Notes in Computational Science and Engineering 45, Springer-Verlag, Berlin/Heidelberg, Germany (2005)
4. BI, Y.: Effects of parameter variations on integrated circuits. MSc.-Thesis, Delft University of Technology / Technical University Eindhoven, The Netherlands (2007)
5. BI, Y., VAN DER KOLK, K.-J., IOAN, D., VAN DER MEIJS, N.P.: Sensitivity computation of interconnect capacitances with respect to geometric parameters. In: *Proc. IEEE International Conference on Electrical Performance of Electronic Packaging (EPEP) 2008*, pp. 209–212 (2008)
6. CIUPRINA, G., IOAN, D., NICULAE, D., FERNÁNDEZ VILLENA, J., SILVEIRA, L.M.: Parametric models based on sensitivity analysis for passive components. *Intelligent Computer Techniques in Applied Electromagnetics*, Studies in Computational Intelligence series 119, Springer Berlin / Heidelberg, pp. 231–239 (2008)
7. CIUPRINA, G., IOAN, D., MIHALACHE, D., SEEBACHER, E.: Domain partitioning based parametric models for passive on-chip components. In: J. Roos, L. Costa (eds.), *Scientific Computing in Electrical Engineering 2008*, Mathematics in Industry 14, Springer, pp. 37–44 (2010)
8. DANIEL, L., SIONG, O.C., LOW, S.C., LEE, K.H., WHITE, J.K.: A multiparameter moment-matching model-reduction approach for generating geometrically parametrized interconnect performance models. *IEEE Trans. Computer-Aided Design of Integr. Circuits and Systems*, 23, pp. 678–693 (2004)
9. DAVIS, T.A.: *Direct methods for sparse linear systems*. The Fundamentals of Algorithms. SIAM, Philadelphia, USA (2006)
10. DAVIS, T.A., PALAMADAI NATARAJAN, E.: Algorithm 907: KLU, a direct sparse solver for circuit simulation problems. *ACM Transactions on Mathematical Software*, 37-3, Article 36 (2010).
11. ELFADEL, I.M., LING, D.L.: A block rational Arnoldi algorithm for multipoint passive model-order reduction of multiport RLC networks. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 1997*, pp. 66–71 (1997)
12. FELDMANN, P., FREUND, R.W.: Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(5), pp. 639–649 (1995)
13. FREUND, R.W.: Sprim: Structure-preserving reduced-order interconnect macro-modeling. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 2004*, pp. 80–87 (2004)
14. FERNÁNDEZ VILLENA, J., SCHILDERS, W.H.A., SILVEIRA, L.M.: Parametric structure-preserving model order reduction. In *IFIP International Conference on Very Large Scale Integration, 2007. VLSI - SoC 2007*, pp. 31–36 (2007)
15. GUNUPUDI, P.K., KHAZAKA, R., NAKHLA, M.S., SMY, T., CELO, D.: Passive parameterized time-domain macromodels for high-speed transmission-line networks. *IEEE Trans. On Microwave Theory and Techniques*, 51(12), pp. 2347–2354 (2003)
16. HEYDARI, P., PEDRAM, M.: Model reduction of variable-geometry interconnects using variational spectrally-weighted balanced truncation. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 2001*, pp. 586–591 (2001)
17. JAIMOUKHA, I.M., KASENALLY, E.M.: Krylov subspace methods for solving large Lyapunov equations. *SIAM Journal on Numerical Analysis*, 31, pp. 227–251 (1994)
18. KAMON, M., WANG, F., WHITE, J.: Generating nearly optimally compact models from Krylov-subspace based reduced-order models. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 47(4), pp. 239–248 (2000)

19. S. KULA: *Reduced order models of interconnects in high frequency integrated circuits*. Ph.D.-Thesis, Politehnica Univ. of Bucharest (2009)
20. LI, J.-R., WANG, F., WHITE, J.: Efficient model reduction of interconnect via approximate system Grammians. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 1999*, pp. 380–383 (1999)
21. LI, P., LIU, F., LI, X., PILEGGI, L.T., NASSIF, S.R.: Modeling interconnect variability using efficient parametric model order reduction. In: *Proc. Design, Automation and Test in Europe Conference and Exhibition (DATE) 2005*, pp. 958–963 (2005)
22. LI, X., LI, P., PILEGGI, L.: Parameterized interconnect order reduction with Explicit-and-implicit multi-parameter moment matching for inter/intra-die variations. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 2005*, pp. 806–812 (2005)
23. LI, Y., BAI, Z., SU, Y., ZENG, X.: Model order reduction of parameterized interconnect networks via a two-directional Arnoldi process. *IEEE Trans. Computer-Aided Design of Integr. Circuits and Systems*, 27(9), pp. 1571–1582 (2008)
24. LIU, Y., PILEGGI, L.T., STROJWAS, A.J.: Model order reduction of RC(L) interconnect including variational analysis. In: *Proc. 36<sup>th</sup> ACM/IEEE Design Automation Conference (DAC) 1999*, pp. 201–206 (1999)
25. MOORE, B.: Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, AC-26(1), pp. 17–32 (1981)
26. EL-MOSELHY, T.A., ELFADEL, I.M., DANIEL, L.: A capacitance solver for incremental variation-aware extraction. In: *Proc IEEE/ACM International Conference on Computer Aided-Design (ICCAD) 2008*, pp. 662–669 (2008)
27. ODABASIOGLU, A., CELIK, M., PILEGGI, L.T.: PRIMA: passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. Computer-Aided Design of Integr. Circuits and Systems*, 17(8), pp. 645–654 (1998)
28. PHILLIPS, J., DANIEL, L., SILVEIRA, L.M.: Guaranteed passive balancing transformations for model order reduction. *IEEE Trans. Computer-Aided Design of Integr. Circuits and Systems*, 22(8), pp. 1027–1041 (2003)
29. PHILLIPS, J.: Variational interconnect analysis via PMTBR. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 2004*, pp. 872–879 (2004)
30. PHILLIPS, J.R., SILVEIRA, L.M.: Poor Man’s TBR: A simple model reduction scheme. *IEEE Trans. Computer-Aided Design of Integr. Circuits and Systems*, 24(1), pp. 43–55 (2005)
31. PILLAGE, L.T., ROHRER, R.A.: Asymptotic waveform evaluation for timing analysis. *IEEE Trans. Computer-Aided Design of Integr. Circuits and Systems*, 9(4), pp. 352–366 (1990)
32. SAAD, Y.: *Iterative Methods for Sparse Linear Systems*. Pws Publishing Co. (1996)
33. SCHILDERS, W.H.A., VAN DER VORST, H.A., ROMMES, J. (eds.): *Model Order Reduction: Theory, Research Aspects and Applications*. Mathematics in Industry, Vol. 13, Springer (2008)
34. SILVEIRA, L.M., KAMON, M., ELFADEL, I., WHITE, J.K.: A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits. In: *Proc. International Conference on Computer Aided-Design (ICCAD) 1996*, pp. 288–294 (1996)
35. YANG, F., ZENG, X., SU, Y., ZHOU, D.: RLCSYN: RLC equivalent circuit synthesis for structure-preserved reduced-order model of interconnect. In: *Proc. International Symposium on Circuits and Systems 2007*, pp. 2710–2713 (2007)
36. YU, H., HE, L., TAN, S.X.D.: Block structure preserving model order reduction. In *BMAS - IEEE Behavioral Modeling and Simulation Workshop*, pp. 1–6 (2005).
37. ZHU, Z., PHILLIPS, J.: Random sampling of moment graph: a stochastic Krylov-reduction algorithm. In: *Proc. Design, Automation and Test in Europe Conference and Exhibition (DATE) 2007*, pp. 1502–1507 (2007)

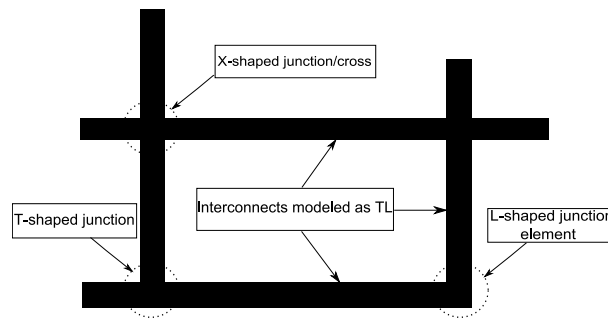


Fig. 4.10 Decomposition of the interconnect net in 2D TLs and 3D junctions.

## 4.2 Robust Procedures for Parametric Model Order Reduction of High Speed Interconnects

<sup>2</sup>Due to higher integration and increasing of running frequency, full Electromagnetic Models (EM) are needed for an accurate prediction of the real behavior of integrated passives and interconnects in currently designed chips [44]. In general, if on-chip interconnects are sorted with respect to their electric length, they may be categorized in three classes: short, medium and long. While the short interconnects have simple circuit models with lumped parameters, the extracted model of the interconnects longer than the wave length has to consider the effect of the distributed parameters, as well. Fortunately, the long interconnects have usually the same cross-sectional geometry along their extension. If not, they may be decomposed in straight parts connected by junction components (Fig.4.10). The former are represented as transmission lines (TLs) whereas the latter are modeled as common passive 3D components.

Due to the fact that the lithographic technology is pushed today to work at its limit, the variability of geometrical and physical properties cannot be neglected. That is why, to obtain robust devices, the variability analysis is necessary even during the design stage [43, 54].

This Section proposes a robust specialized technique to extract reduced parametric compact models, described as parametric SPICE like netlists, for long interconnects modeled as transmission lines with several field effects such as skin effect and substrate losses. The technique uses an EM formulation based on partial differential equations (PDE), which is discretized to obtain a finite state space model. Next, a variability analysis of the geometrical data is carried out. Finally, a method to extract an equivalent parametric circuit is proposed. The procedure is validated by applying it on a study case for which experimental data is available.

<sup>2</sup> Section 4.2 has been written by: Gabriela Ciuprina, Alexandra Ștefănescu, Sebastian Kula and Daniel Ioan. For additional topics see also the Ph.D.-Theses of the second author [59] and of the third author [55].



### 4.2.1 Field problem formulation: 3D - PDE models

Long interconnects and passive components with significant high frequency field effects, have to be modeled by taken into consideration Full Wave (FW) electromagnetic field equations. Typical examples of such parasitic effects are: skin effect and proximity, substrate losses, propagation retardation and crosstalk. Only Maxwell equations in FW regime

$$\begin{aligned} \operatorname{curl} H &= J + \frac{\partial D}{\partial t}, & \operatorname{div} B &= 0, \\ \operatorname{curl} E &= -\frac{\partial B}{\partial t}, & \operatorname{div} D &= \rho, \end{aligned} \quad (4.22)$$

complemented with the constitutive equations which describe the material behavior:

$$B = \mu H, \quad D = \varepsilon E, \quad J = \sigma E, \quad (4.23)$$

can model these effects. While material constants are known for each subdomain (Si, Al, SiO<sub>2</sub>), vectorial fields  $B, H, E, D : \Omega \times [0, T] \rightarrow \mathbb{R}^3$  and the scalar field  $\rho : \Omega \times [0, T] \rightarrow \mathbb{R}$  are the unknowns of the problem. They can be univocal determined in the simple connected set  $\Omega$ , which is the computational domain, for zero initial conditions ( $B = 0, D = 0$  for  $t = 0$ ), if appropriate boundary conditions are imposed.

According to authors' knowledge, the best boundary conditions which allow the field-circuit coupling are those given by the electric circuit element (ECE) formulation [53]. Considering  $S'_1, S'_2, \dots, S'_n \subset \partial\Omega$  a disjoint set of surfaces, called terminals (Fig.4.2.1), the following boundary conditions are assumed:

$$n \cdot \operatorname{curl} E = 0 \quad \text{on} \quad \partial\Omega, \quad (4.24)$$

$$n \cdot \operatorname{curl} H = 0 \quad \text{on} \quad \partial\Omega \setminus \cup_{k=1}^n S'_k \quad (4.25)$$

$$n \times E = 0 \quad \text{on} \quad \cup_{k=1}^n S'_k \quad (4.26)$$

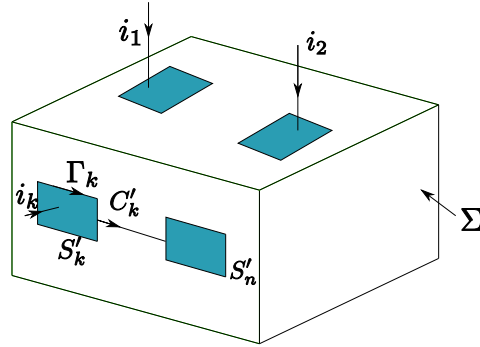
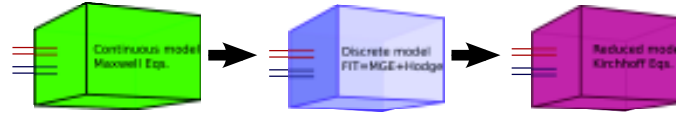


Fig. 4.11 ECE - Electric Circuit Element with multiple terminals.



**Fig. 4.12** Three levels of abstraction for a component model and its corresponding equations.

Condition (4.24) interdicts the magnetic coupling between the domain and its environment, (4.25) interdicts the galvanic coupling and the capacitive coupling through the boundary excepting for the terminals and (4.26) interdicts the variation of the electric potential over the terminal, thus allowing the connection of the device to exterior electric circuit nodes. For each terminal,  $k = 1, \dots, n$  the voltage and the current can be univocal defined:

$$u_k = \int_{C_k \subset \partial\Omega} E \cdot dr, \quad i_k = \int_{\partial S'_k} H \cdot dr, \quad (4.27)$$

where  $C'_k$  is an arbitrary path on the device boundary  $\partial\Omega$ , that starts on  $S'_k$  and ends on  $S'_n$ , where, by convention, the  $n$ -th terminal is considered as reference, i.e.  $u_n = 0$ . If we assume that the terminals are excited in voltage, then  $u_k, k = 1, 2, \dots, n-1$  are input signals and  $i_k, k = 1, 2, \dots, n-1$  are output signals. Equations (4.24)÷(4.26) define a multiple input multiple output (MIMO) linear system with  $n-1$  inputs and  $n-1$  outputs, but with a state space of infinite dimension. In the weak form of Maxwell's equations, state variables,  $H, E$  belong to the Sobolev space  $H(\text{curl}, \Omega)$  [39]. Uniqueness theorem of the ECE field problem [53] generates the correct formulation of the transfer function  $\mathbf{Y}(s) : \mathfrak{C} \rightarrow \mathfrak{C}^{(n-1) \times (n-1)}$ , which represents the matrix of the terminals admittance for a complex frequency  $s$ . The relation

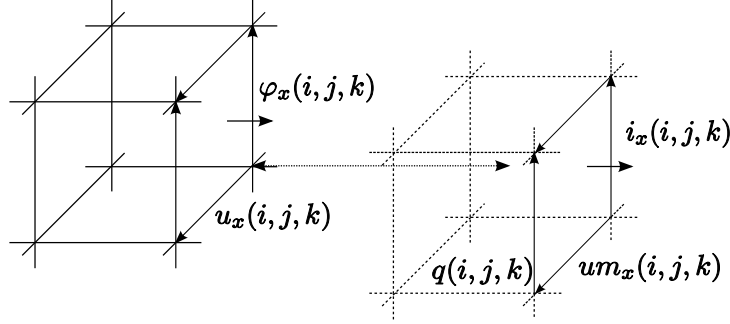
$$\mathbf{i} = \mathbf{Y}\mathbf{u} \quad (4.28)$$

defines a linear transformation in the frequency domain of the terminal voltages vector  $\mathbf{u} \in \mathfrak{C}^{n-1}$  to the currents vector  $\mathbf{i} \in \mathfrak{C}^{n-1}$ .

### 4.2.2 Numeric discretization and state space models

PDE models are too complex for designers needs. The approach we propose for the extraction of the electric models is schematically represented in Fig. 4.12. The left block corresponds to the formulation described in the previous section.

The next important step in the EM modeling is the discretization of the PDEs. One of the simplest methods to carry out this, is based on the indexFinite Integration Technique (FIT)Finite Integration Technique (FIT), a numerical method able to solve field problems based on spatial discretization "without shape functions". Two staggered orthogonal (Yee type) grids are used as discretization mesh [41]. The centers of the primary cells are the nodes of the secondary cells. The degrees of freedom



**Fig. 4.13** Dofs for FIT numerical method in the two dual grids cells.

(dofs) used by FIT are not local field components as in FEM or in FDTD, but global variables, i.e. electric and magnetic voltages  $\mathbf{u}_e, \mathbf{u}_m$ , electric currents  $\mathbf{i}$ , and magnetic and electric fluxes  $\phi, \psi$  assigned to the grid elements: edges and faces, respectively. They are associated to these grids elements in a coherent manner (Fig.4.13).

By applying the global form of electromagnetic field equations on the mesh elements (elementary faces and their borders), a system of differential algebraic equations (DAE), called Maxwell Grid Equations (MGE) is obtained:

$$\text{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \Rightarrow \int_{\Gamma} \mathbf{E} d\mathbf{r} = -\int \int_{S_{\Gamma}} \frac{\partial \mathbf{B}}{\partial t} d\mathbf{A} \Rightarrow \mathbf{C} \mathbf{u}_e = -\frac{d\phi}{dt} \quad (4.29)$$

$$\Leftrightarrow \text{div} \mathbf{B} = 0 \Rightarrow \int \int_{\Sigma} \mathbf{B} d\mathbf{A} = 0 \Rightarrow \mathbf{D}' \phi = 0 \quad (4.30)$$

$$\text{curl} \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \Rightarrow \int_{\Gamma} \mathbf{H} d\mathbf{r} = \int \int_{S_{\Gamma}} (\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}) d\mathbf{A} \Rightarrow \mathbf{C}' \mathbf{u}_m = \mathbf{i} + \frac{d\psi}{dt} \quad (4.31)$$

$$\Leftrightarrow \text{div} \mathbf{D} = \rho \Rightarrow \int \int_{\Sigma} \mathbf{D} d\mathbf{A} = \int \int \int_{D_{\Sigma}} \rho d\mathbf{v} \Rightarrow \mathbf{D} \psi = \mathbf{q} \quad (4.32)$$

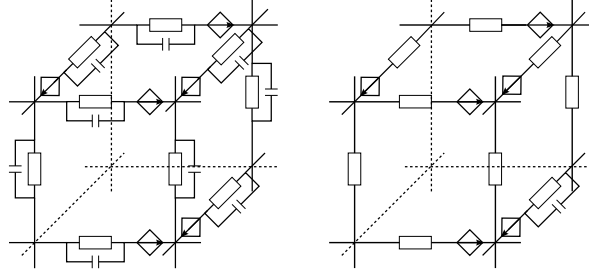
$$\Leftrightarrow \text{div} \mathbf{J} = -\frac{\partial \rho}{\partial t} \Rightarrow \int \int_{\Sigma} \mathbf{J} d\mathbf{A} = -\int \int \int_{D_{\Sigma}} \frac{\partial \rho}{\partial t} d\mathbf{v} \Rightarrow \mathbf{D} \mathbf{i} = -\frac{d\mathbf{q}}{dt} \quad (4.33)$$

FIT combines MGE with Hodge's linear transform operators, which approximate the material behavior (4.23):

$$\phi = \mathbf{G}_m \mathbf{u}_m, \quad \psi = \mathbf{C}_e \mathbf{u}_e, \quad \mathbf{i} = \mathbf{G}_e \mathbf{u}_e. \quad (4.34)$$

The main characteristics of the FIT method are:

- There is no discretization error in the MGE fundamental equations (4.29) ÷ (4.33). All numerical errors are hold by the discrete Hodge operators (4.34).
- An equivalent FIT circuit (Fig.4.14), having MGE+Hodge as equations may be easily build. The graphs of the two constituent mutually coupled sub-circuits are exactly the two dual discretization grids; therefore the complexity of the equivalent circuit has a linear order with respect to the number of grid-cells [47].
- MGE are:



**Fig. 4.14** Electric (left) and magnetic (right) equivalent FIT circuits

- **sparse**: matrices  $G_m, C_e$  and  $G_e$  are diagonal and matrices  $C, D$  have maximum 6 non-zero entries per row,
- **metric-free**: matrices  $\mathbf{C}$  - the discrete-curl and  $\mathbf{D}$  - the discrete-div operators have only 0, +1 and -1 as entries,
- **mimetic**: in Maxwell equations curl and div operators are replaced by their discrete counterparts  $\mathbf{C}$  and  $\mathbf{D}$ , and
- **conservative**: the discrete form of the discrete charge conservation equation is a direct consequence of both Maxwell and as well as of the MGE equations.

Due to these properties the numerical solutions have **no spurious modes**.

Considering FIT equations (4.29), (4.31), (4.34) with the discrete forms of boundary conditions (4.24) ÷ (4.27) a linear time-invariant system is defined having the same input-output quantities as (4.28), but the state equations:

$$C \frac{dx}{dt} + Gx = Bu, \quad i = Lx, \quad (4.35)$$

where  $x = [u_m^T, u_e^T, i^T]^T$  is the state space vector, consisting of electric voltages  $u_e$  defined on the electric grid used by FIT, magnetic voltages  $u_m$  defined on the magnetic grid and output quantities  $i$ . Equations can be written such that only two semi-state space matrices ( $C$  and  $G$ ) are affected by geometric parameters (denoted by  $\alpha$  in what follows). Considering all terminals voltage-excited, the number of inputs is always equal to the number of outputs. Since output currents are components of the state vector, the matrix  $L = B^T$  is merely a selection matrix.

For instance, the structure of the matrices in the case of voltage excitation is the following:

$$C = \begin{bmatrix} G_m(\alpha) & 0 & 0 \\ 0 & -C_i(\alpha) & 0 \\ 0 & 0 & 0 \\ 0 & C_{SI}(\alpha) & 0 \\ 0 & C_{TE}(\alpha) & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 & B_1 & B_2 & 0 \\ B_1^T & -G_i(\alpha) & 0 & 0 \\ 0 & 0 & B_{SI} & 0 \\ 0 & G_{SI}(\alpha) & 0 & 0 \\ 0 & G_{TE}(\alpha) & -S_E & 0 \\ 0 & P_E & 0 & 0 \end{bmatrix} \quad (4.36)$$

There are six sets of rows, corresponding to the six sets of equations. The first group of equations is obtained by writing Faraday's law for inner elementary electric loops.  $G_m$  is a diagonal matrix holding the magnetic conductances that pass

through the electric loops. The block  $[B_1 B_2]$  has only 0, 1, -1 entries, describing the incidence of inner branches and branches on the boundary to electric faces. The second group corresponds to Ampere's law for elementary magnetic loops.  $C_i$  and  $G_i$  are diagonal matrices, holding the capacitances and electric conductances of the inner branches. The third group represents Faraday's law for electric loops on the boundary.  $B_{SI}$  has only 0, 1, -1 entries, describing the incidence of electric branches included in the boundary to the electric boundary faces. The fourth row is obtained from the current conservation law for all nodes on the boundary excepting for the nodes on the electric terminals.  $G_{SI}$  and  $C_{SI}$  hold electric conductances and capacitances directly connected to boundary. The fifth set of equations represents current conservation for electric terminals.  $G_{TE}$  and  $C_{TE}$  hold electric conductances and capacitances that are directly connected to electric terminals.  $S_E$  is the connexion matrix between electric branches and terminals path. The last row is the discrete form of (4.27), obtained by expressing the voltages of electric terminals as sums of voltages along open paths from terminals to ground,  $P_E$  being a topological matrix that holds the paths that connect electric terminals to ground.

Thus, the top left square block of  $C$  is diagonal and the top left square bloc of  $G$  is symmetric. The size of this symmetric bloc corresponds to the useful magnetic branches and to the useful inner electric branches. Its size is dominant over the size of the matrix, therefore, solving or reduction strategies that take into consideration this particular structure are useful.

The discretized state-space system given by (4.35) describes the input output relationship in the frequency domain

$$i = Yu, \quad (4.37)$$

similar to (4.28), but having as transfer (circuit) function:

$$Y = L(sC + G)^{-1} B \quad (4.38)$$

which is a rational function with a finite number of poles.

In conclusion, the discretization of the continuous model leads to a model represented by a MIMO linear time invariant system described by the state equations of finite size. Even if this is an important step ahead in the extraction procedure, the state space dimension is still too large for designer's needs, therefore a further modeling step aiming an order reduction is required.

### 4.2.3 Transmission Lines - 2D+1D Models

In this section, aiming to reduce the model extraction effort, we will exploit the particular property of interconnects of having invariant transversal section along their extent. We assume that the field has a similar structure as a transversal electromagnetic wave that propagates along the line. The typical interconnect configuration (Fig.4.15) considered consists of  $n$  parallel conductors having rectangular cross sec-

tion, permeability  $\mu = \mu_0$ , permittivity  $\varepsilon = \varepsilon_0$  and conductivity  $\sigma_k, k = 1, 2, \dots, n$ , placed in a  $\text{SiO}_2$  layer ( $\sigma_d, \varepsilon_d$ , possibly dependent on  $y$ ) placed above a silicon substrate ( $\sigma_s, \varepsilon_s$ ).

If the field is decomposed in its **longitudinal** (oriented along the line, which is assumed to lie along the Oz axis) and the **transversal components** (oriented in the xOy plane)

$$\mathbf{E} = \mathbf{E}_t + \mathbf{k}E_z, \quad \mathbf{J} = \mathbf{J}_t + \mathbf{k}J_z, \quad \mathbf{H} = \mathbf{H}_t + \mathbf{k}H_z, \quad (4.39)$$

then Maxwell's Equations can be separated into two groups:

$$\begin{aligned} \text{curl}_{xy} \mathbf{H}_t &= \mathbf{k} \left( J_z + \varepsilon \frac{\partial E_z}{\partial t} \right), & \text{div}_{xy} (\mu \mathbf{H}_t) &= -\frac{\partial (\mu H_z)}{\partial z}, \\ \text{curl}_{xy} \mathbf{E}_t &= -\mathbf{k} \mu \frac{\partial H_z}{\partial t}, & \text{div}_{xy} (\varepsilon \mathbf{E}_t) &= \rho - \frac{\partial (\varepsilon E_z)}{\partial z}, \end{aligned} \quad (4.40)$$

called **transversal equations** and

$$\begin{aligned} \frac{\partial \mathbf{E}_t}{\partial z} - \text{grad}_{xy} E_t &= -\mu \frac{\partial}{\partial t} (\mathbf{H}_t \times \mathbf{k}); \\ \frac{\partial \mathbf{H}_t}{\partial z} - \text{grad}_{xy} H_z &= \mathbf{J}_t \times \mathbf{k} + \varepsilon \frac{\partial}{\partial t} (\mathbf{E}_t \times \mathbf{k}); \end{aligned}$$

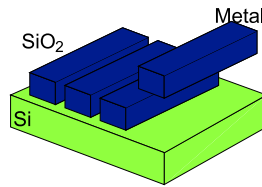
called **propagation equations**.

The following hypotheses are adopted:

- The volume charge density  $\rho$  and the displacement current density  $\frac{\partial \mathbf{E}}{\partial t}$  are neglected both in conductors and in the substrate.
- The following "longitudinal" terms  $E_z = 0, H_z = 0$  are canceled in the transversal equations, neglecting the field generated by eddy currents.
- The longitudinal conduction current is neglected in dielectric  $J_z = 0$ , but not in the conductors.
- Since the conductances  $\sigma_k$  of the conductors are much bigger than the dielectric conductance  $\sigma_d$ , the transversal component of the electric field is neglected in the line conductors and in the substrate:

$$\mathbf{E}_t = \frac{1}{\sigma_k} \mathbf{J}_t = 0. \quad (4.41)$$

Under these hypotheses the transversal equations have the following form (where (k) = conductor, (s) = substrate, (d) = dielectric):



**Fig. 4.15** Typical interconnect configuration.

$$\begin{aligned} \text{curl}_{xy} \mathbf{H}_t &= \begin{cases} \mathbf{k}J_z, & \text{in (k) and (s)} \\ 0, & \text{in (d)} \end{cases} & \text{div}_{xy}(\mu \mathbf{H}_t) &= 0 \\ \text{curl}_{xy} \mathbf{E}_t &= 0, & \text{div}_{xy}(\varepsilon \mathbf{E}_t) &= 0, \end{aligned} \quad (4.42)$$

identical with the **steady state electromagnetic field equations**. For this reason, the electric field admits a scalar electric potential  $V(x, y, z, t)$ , whereas the magnetic field admits a vector magnetic potential  $\mathbf{A}(x, y, z, t) = \mathbf{k}A(x, y, z, t)$  with longitudinal orientation, so that:

$$\mathbf{E}_t = -\text{grad}_{xy} V, \quad (4.43)$$

$$\mathbf{H}_t = \frac{1}{\mu} [\mathbf{k} \times (\text{curl} \mathbf{A} \times \mathbf{k})] = -\mathbf{k} \times \frac{1}{\mu} \text{grad}_{xy} (A \times \mathbf{k}). \quad (4.44)$$

Thus, **the propagation equations** become:

$$\begin{aligned} \text{grad}_{xy} \left[ \frac{\partial A}{\partial t} + \frac{\partial V}{\partial z} + E_z \right] &= 0, \\ -\text{grad}_{xy} H_z &= \mathbf{k} \times \left[ \frac{1}{\mu} \text{grad}_{xy} \left( \frac{\partial A}{\partial z} \right) + \sigma \text{grad}_{xy} V + \varepsilon \text{grad}_{xy} \left( \frac{\partial V}{\partial t} \right) \right]. \end{aligned} \quad (4.45)$$

By assuming an asymptotic behavior of potentials, the integration of the propagation equations yields to:

$$\begin{aligned} E_z &= \frac{1}{\sigma} J_z = -\frac{\partial V}{\partial z} - \frac{\partial A}{\partial t}, \\ H_z &= -\int_C \left[ \frac{1}{\mu} \frac{\partial}{\partial \mathbf{n}} \left( \frac{\partial A}{\partial z} \right) + \sigma \frac{\partial V}{\partial \mathbf{n}} + \varepsilon \frac{\partial}{\partial \mathbf{n}} \left( \frac{\partial V}{\partial t} \right) \right] ds, \end{aligned} \quad (4.46)$$

where  $C$  is a curve in the plane  $z=\text{constant}$ , which starts from the infinity and stops in the computation point of the field  $H_z$ ,  $\mathbf{n}$  is the normal to the curve, oriented so that the line element is

$$d\mathbf{s} = ds \mathbf{k} \times \mathbf{n}. \quad (4.47)$$

From (4.41) it follows that the potential  $V$  is constant on every transversal cross-section of the conductors and zero in the substrate:

$$V|_{S_k} = V_k(z, t), \quad V_s = 0. \quad (4.48)$$

From relations (4.42) and (4.43) it follows that, in the transversal plane, the electric field has the same distribution as an electrostatic field. By using the uniqueness theorem of the electrostatic field it results that the function  $V(x, y, z, t)$  is uniquely determined by the potentials of the conductors  $V_k$ . Consequently, due to the linearity, the per unit length (p.u.l.) charge of conductors and the current loss through the dielectric are:

$$q_k(z, t) = -\int_{\partial S_k} \varepsilon_d \frac{\partial V}{\partial \mathbf{n}} ds = \sum_{m=1}^n c_{km} V_m(z, t); \quad (4.49)$$

$$i_{gk}(z, t) = -\int_{\partial S_k} \sigma_d \frac{\partial V}{\partial \mathbf{n}} ds = \sum_{m=1}^n g_{km} V_m(z, t), \quad (4.50)$$

where  $c_{km}$  is the p.u.l. capacitance, and  $g_{km}$  is the p.u.l. conductance between the conductor  $k$  and the conductor  $m$ .

By integrating  $E_z$  equation from (4.46) over the surface  $S_k$  and  $H_z$  equation from (4.46) along the path  $\partial S_k$  which bounds this surface, the following propagation equations for potentials are obtained:

$$-\frac{\partial \tilde{v}_k}{\partial z} = r_k^0 i_k + \frac{\partial \tilde{a}_k}{\partial t}; \quad -\frac{\partial i_k}{\partial z} = i_{gk} + \frac{\partial q_k}{\partial t}, \quad (4.51)$$

where  $r_k^0 = 1/(\sigma_k A_{S_k})$  is the p.u.l. d.c. resistance of the conductor  $k$ , and

$$\begin{aligned} \tilde{v}_k(z, t) &= \frac{1}{A_{S_k}} \int_{S_k} V(x, y, z, t) dx dy = v_k(z, t), \\ \tilde{a}_k &= \frac{1}{A_{S_k}} \int_{S_k} A(x, y, z, t) dx dy \end{aligned} \quad (4.52)$$

are the average values of the two potentials on the cross-section of the conductor  $k$ .

By computing the average values of the magnetic potential as in [58] and by substituting (4.49), (4.50) in (4.51) the following expressions are obtained in zero initial conditions:

$$-\frac{\partial v_k}{\partial z} = r_k^0 i_k + \sum_{m=1}^n l_{km}^0 \frac{\partial i_m}{\partial t} + \sum_{m=1}^n \frac{\partial}{\partial t} \int_0^t \left( \frac{dl_{km}}{dt} \right)_{t-\tau} i_m(\tau, t) d\tau, \quad (4.53)$$

$$-\frac{\partial i_k}{\partial z} = \sum_{m=1}^n \left( g_{km} v_m + c_{km} \frac{\partial v_m}{\partial t} \right), \quad (4.54)$$

where  $l_{km}^0$  are the p.u.l. external inductances (self inductances for  $k = m$  and mutual inductances for  $k \neq m$ ) of the conductors ( $k$ ) and ( $m$ ) where the return current is distributed on the surface of the substrate, and  $l_{km}(t)$  are "transient p.u.l. inductances", defined as the average values on  $S_k$  of the vector potential  $A$  obtained in zero initial conditions by a unity step current injected in conductor ( $m$ ).

For zero initial conditions for the currents  $i_m(z, 0) = 0$ , for the potential  $v_m(z, 0) = 0$  and for the field  $B_k^0(s) = 0$ , the Laplace transform of (4.53) and (4.54) can be written as:

$$-\frac{dv_k(z, s)}{dz} = \sum_{m=1}^n Z_{km}(s) i_m(z, s), \quad -\frac{di_k(z, s)}{dz} = \sum_{m=1}^n Y_{km}(s) v_m(z, s), \quad (4.55)$$

which is identical to the operational form of the classical Transmission Lines (TLs) Telegrapher's equations, but where the p.u.l. inductances depend on  $s$  (implicitly on the frequency in a time-harmonic regime). In order to extract these dependencies, a magneto-quasi-static (MQS) field problem has to be solved.



#### 4.2.4 Numeric extraction of line parameters

Models with various degrees of fineness can be established for TLs. The coarsest ones are circuit models with lumped parameters, such as the pi equivalent circuit for a single TL shown in Fig.4.16. As expected, the characteristic of such a circuit is appropriate only at low frequencies, over a limited range, and for short lines. Even chaining similar cells, the result is not appropriate.

At high frequencies, the distributed effects have to be considered as an important component of the model. Proper values for the line parameters can be obtained only by simulating the electromagnetic field. The extraction of line parameters is the main step in TLs modeling since the behavior of a line with a given length can be computed from them. For instance, for a multiconductor transmission line, from the per unit length parameters matrices  $\mathbf{R}$ ,  $\mathbf{L}$ ,  $\mathbf{C}$  and  $\mathbf{G}$  the transfer matrix for a line of length  $l$  can be computed as

$$\mathbf{T} = \exp[(\mathbf{D} + j\omega\mathbf{E})l], \quad \text{where} \quad \mathbf{D} = \begin{bmatrix} \mathbf{0} & -\mathbf{R} \\ -\mathbf{G} & \mathbf{0} \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} \mathbf{0} & -\mathbf{L} \\ -\mathbf{C} & \mathbf{0} \end{bmatrix}. \quad (4.56)$$

From them, other parameters (impedance, admittance or scattering) can be computed. The simplest method to extract constant matrices of the line resistance  $\mathbf{R}$ , capacitance  $\mathbf{C}$  and inductance  $\mathbf{L}$ , respectively, is to solve the field equations numerically in steady-state electric conduction (EC), electrostatics (ES) and magnetostatics (MS) regimes. Empirical formulas may also be found in the literature, such as the ones given in [56] for the line capacitance. None of them take the frequency dependence of p.u.l. parameters into account.

A first attempt to take into consideration the frequency effect, which becomes important at high frequencies, is to compute the skin depth in the conductor and to use a better approximation for the resistance. In [50] we proposed a much more accurate estimation of frequency dependent line parameters based on the numerical modeling of the EM field including the semiconductor substrate. The previous section is the theoretical argument of this approach in which two complementary problems are solved, the first one describing the transversal behavior of the line from which  $\mathbf{Y}_l(\omega) = \mathbf{G}(\omega) + j\omega\mathbf{C}(\omega)$  is consequently extracted, and the second one describing the longitudinal behavior of the line from which  $\mathbf{Z}_l(\omega) = \mathbf{R}(\omega) + j\omega\mathbf{L}(\omega)$  is extracted.

Since the first field problem is dedicated to the computation of the transversal capacitances between wires and their loss conductances, according to the previous section, the natural choice is to solve a 2D problem of the transversal electro-quasi-

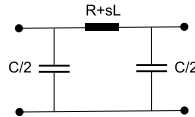
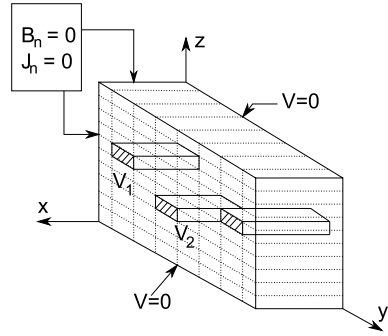
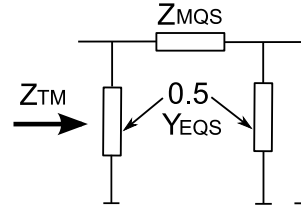


Fig. 4.16 The coarsest model for a single transmission line: a pi equivalent circuit.



**Fig. 4.17** Boundary conditions for the full wave - transversal magnetic problem.



**Fig. 4.18** The pi equivalent circuit for a simulated short line segment. Parameters are evaluated from field simulations.

static (EQS) field in dielectrics, considering the line wires as perfect conductors with given voltage. The boundary conditions are of Dirichlet type  $V = 0$  on the lower electrode, and open boundary conditions (e.g. Robin, SDI or appropriate ELOB [48]) on the other three sides. A dual approach, such as dFIT [49] allows a robust and accurate parameter extraction.

The second field problem focuses on the longitudinal electric and the generated transversal magnetic field. Consequently, a short line-segment (with only one cell layer) is considered. The magneto-quasi-static (MQS) regime of the EM field is appropriate for the extraction of  $\mathbf{Z}_l(\omega)$ . However, for our simulations we used a our FIT solver for Full Wave (FW) ECE problems. The magnetic grid is 2D, thus ensuring the TM mode of propagation.

In order to eliminate the transversal distribution of the electric field, the lower electrode is prolonged over the entire far-end cross-section of the rectangular computational domain, which thus has perfect electric conductor (PEC) boundary conditions  $\mathbf{E}_t = \mathbf{0}$  on two of their faces. On the three lateral faces, open-absorbing boundary conditions are the natural choice, whereas on the near-end cross-section the natural boundary conditions are those of the Electric Circuit Element (ECE):  $B_n = 0$ ,  $\mathbf{n} \times \text{curl} \mathbf{H} = \mathbf{0}$  excepting for the wire traces, where  $\mathbf{E}_t = \mathbf{0}$ . These conditions ensure the correct definition of the terminals voltages, and consequently of the impedance/admittance matrix (Fig.4.17).

These boundary conditions are the field representation of the line segment with short-circuit at the far-end, whereas the 2D EQS problem is the field representation of the segment-line with open far-end.

The transversal component is finally subtracted from the FW-TM simulation to obtain an accurate approximation of the line impedance, as given by

$$\mathbf{Z}_{MQS} = \left( \mathbf{Z}_{TM}^{-1} - \frac{1}{2} \mathbf{Y}_{EQS} \right)^{-1}. \quad (4.57)$$

This subtraction is carried out according to a pi-like equivalent net for the simulated short segment (Fig.4.18). Finally, the line parameters are:

$$\mathbf{G}(\omega) = \text{Re}(\mathbf{Y}_l), \quad \mathbf{C}(\omega) = \text{Im}(\mathbf{Y}_l)/\omega, \quad \mathbf{R}(\omega) = \text{Re}(\mathbf{Z}_l), \quad \mathbf{L}(\omega) = \text{Im}(\mathbf{Z}_l)/\omega, \quad (4.58)$$

where

$$\mathbf{Y}_l = \mathbf{Y}_{EQS}/\Delta l, \quad \mathbf{Z}_l = \mathbf{Z}_{MQS}/\Delta l, \quad (4.59)$$

where  $\Delta l$  is the length of the considered line-segment and  $\mathbf{Z}_{TM}$  is the impedance matrix extracted from the TM field solution.

This numerical approach to extract the line parameters, named the *two fields method*, is more robust and may be applied without difficulties to multi-wire lines. The obtained values of the line parameters are frequency dependent, taking into consideration proximity and skin effects as well as losses induced in the conducting substrate.

#### 4.2.5 Variability Analysis of Line Parameters

The simplest way to analyze the parameter variability is to compute first order sensitivities. These are derivatives of the device characteristics with respect to the design parameters. The sensitivities of the line parameters are essential to estimate the impact of small variations on the device behavior. Moreover, the sensitivity of the terminal behavior of interconnects can also be estimated.

For instance, in the case of a single TL, having the global admittance given by

$$\mathbf{Y} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} \frac{\cosh \gamma l}{Z_c \sinh \gamma l} & -\frac{1}{Z_c \sinh \gamma l} \\ -\frac{1}{Z_c \sinh \gamma l} & \frac{\cosh \gamma l}{Z_c \sinh \gamma l} \end{bmatrix} \quad (4.60)$$

the sensitivities of the terminal admittance with respect to a parameter can be computed as:

$$\frac{\partial Y_{11}}{\partial \alpha} = \frac{l}{Z_c} \frac{\partial \gamma}{\partial \alpha} - \frac{\cosh \gamma l}{Z_c^2 \sinh \gamma l} \frac{\partial Z_c}{\partial \alpha} - \frac{l \cosh^2 \gamma l}{Z_c \sinh^2 \gamma l} \frac{\partial \gamma}{\partial \alpha} \quad (4.61)$$

$$\frac{\partial Y_{12}}{\partial \alpha} = \frac{1}{Z_c^2 \sinh \gamma l} \frac{\partial Z_c}{\partial \alpha} + \frac{l \cosh}{Z_c \sinh^2 \gamma l} \frac{\partial \gamma}{\partial \alpha} \quad (4.62)$$

where the sensitivities of

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} \quad \text{and} \quad Z_c = \sqrt{(R + j\omega L)/(G + j\omega C)}$$

can be computed if the sensitivities of the p.u.l. parameters  $\partial R/\partial \alpha$ , etc. are known.

In the case of a multiconductor TL with  $n$  conductors the sensitivity of the admittance matrix  $\mathbf{Y}$  of dimension  $(2n \times 2n)$  is computed by means of the sensitivity of the transfer matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{bmatrix} \quad (4.63)$$

also of dimension  $(2n \times 2n)$ , knowing that

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} \end{bmatrix} = \begin{bmatrix} -\mathbf{T}_{12}^{-1}\mathbf{T}_{11} & \mathbf{T}_{12}^{-1} \\ \mathbf{T}_{22}\mathbf{T}_{12}^{-1}\mathbf{T}_{11} - \mathbf{T}_{21} & -\mathbf{T}_{22}\mathbf{T}_{12}^{-1} \end{bmatrix}. \quad (4.64)$$

In the formulas above, all the sub-blocks are of dimensions  $(n \times n)$ . For instance

$$\frac{\partial \mathbf{Y}_{11}}{\partial \alpha} = -\mathbf{T}_{12}^{-1} \frac{\partial \mathbf{T}_{12}}{\partial \alpha} \mathbf{T}_{12}^{-1} \mathbf{T}_{11} - \mathbf{T}_{12}^{-1} \frac{\partial \mathbf{T}_{11}}{\partial \alpha}, \quad (4.65)$$

$$\frac{\partial \mathbf{Y}_{12}}{\partial \alpha} = -\mathbf{T}_{12}^{-1} \frac{\partial \mathbf{T}_{12}}{\partial \alpha} \mathbf{T}_{12}^{-1}. \quad (4.66)$$

The transfer matrix  $\mathbf{T}$  is computed with (4.56) and its sensitivity is

$$\frac{\partial \mathbf{T}}{\partial \alpha} = \exp[(\mathbf{D} + j\omega\mathbf{E})l] \left( \frac{\partial \mathbf{D}}{\partial \alpha} + j\omega \frac{\partial \mathbf{E}}{\partial \alpha} \right), \quad (4.67)$$

where

$$\frac{\partial \mathbf{D}}{\partial \alpha} = \begin{bmatrix} \mathbf{0} & -\partial \mathbf{R} / \partial \alpha \\ -\partial \mathbf{G} / \partial \alpha & \mathbf{0} \end{bmatrix}, \quad \frac{\partial \mathbf{E}}{\partial \alpha} = \begin{bmatrix} \mathbf{0} & -\partial \mathbf{L} / \partial \alpha \\ -\partial \mathbf{C} / \partial \alpha & \mathbf{0} \end{bmatrix}. \quad (4.68)$$

Thus, the basic quantities needed to estimate the sensitivity of the admittance are the sensitivities of the p.u.l. parameters. By using a direct differentiation technique, as explained in [40] the sensitivities of the EQS and TM problems with respect to the parameters that vary, i.e.  $\partial \mathbf{Y}_{EQS} / \partial \alpha$  and  $\partial \mathbf{Z}_{TM} / \partial \alpha$  are computed. Then, the sensitivity of the MQS mode is computed by taking the derivative of (4.57):

$$\frac{\partial \mathbf{Z}_{MQS}}{\partial \alpha} = - \left( \mathbf{Z}_{TM}^{-1} - \frac{1}{2} \mathbf{Y}_{EQS} \right)^{-1} \left( -\mathbf{Z}_{TM}^{-1} \frac{\partial \mathbf{Z}_{TM}}{\partial \alpha} \mathbf{Z}_{TM}^{-1} - \frac{1}{2} \frac{\partial \mathbf{Y}_{EQS}}{\partial \alpha} \right) \left( \mathbf{Z}_{TM}^{-1} - \frac{1}{2} \mathbf{Y}_{EQS} \right)^{-1} \quad (4.69)$$

Finally, the sensitivities of the p.u.l. parameters are:

$$\frac{\partial \mathbf{R}}{\partial \alpha} = \frac{1}{l} \operatorname{Re} \left\{ \frac{\partial \mathbf{Z}_{MQS}}{\partial \alpha} \right\}, \quad \frac{\partial \mathbf{L}}{\partial \alpha} = \frac{1}{l\omega} \operatorname{Im} \left\{ \frac{\partial \mathbf{Z}_{MQS}}{\partial \alpha} \right\}, \quad (4.70)$$

$$\frac{\partial \mathbf{G}}{\partial \alpha} = \frac{1}{l} \operatorname{Re} \left\{ \frac{\partial \mathbf{Y}_{EQS}}{\partial \alpha} \right\}, \quad \frac{\partial \mathbf{C}}{\partial \alpha} = \frac{1}{l\omega} \operatorname{Im} \left\{ \frac{\partial \mathbf{Y}_{EQS}}{\partial \alpha} \right\}. \quad (4.71)$$

The values of the sensitivities thus obtained depend on the frequency as well.

### 4.2.6 Parametric Models based on Taylor Series

Continuous improvements in today's fabrication processes determine smaller chip sizes and smaller device geometries. Process variations induce changes in the properties of metallic interconnect between devices.

Simple parametric models are often obtained by truncating the Taylor series expansion for the quantity of interest. This requires the computation of the derivatives of the device characteristics with respect to the design parameters [54]. Let us assume that  $y(\alpha_1, \alpha_2, \dots, \alpha_n) = y(\alpha)$  is the device characteristic which depends on the design parameters  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]$ . The quantity  $y$  may be, for instance, the real or the imaginary part of the device admittance at a given frequency or any of the p.u.l. parameters. The parameter variability is thus completely described by the real function,  $y$ , defined over the design space  $S$ , a subset of  $\mathbb{R}^n$ . The nominal design parameters correspond to the particular choice  $\alpha_0 = [\alpha_{01} \ \alpha_{02} \ \dots \ \alpha_{0n}]$ .

#### 4.2.6.1 Additive Model (A)

If  $y$  is smooth enough then its truncated Taylor Series expansion is the best polynomial approximation in the vicinity of the expansion point  $\alpha_0$ . For one parameter ( $n = 1$ ), the additive model is the first order truncation of the Taylor series:

$$\hat{y}(\alpha) = y(\alpha_0) + \frac{\partial y}{\partial \alpha}(\alpha_0)(\alpha - \alpha_0). \quad (4.72)$$

If we denote by  $y(\alpha_0) = y_0$  the nominal value of the output function, by  $\frac{\partial y}{\partial \alpha}(y_0) \frac{\alpha_0}{y_0} = S_{\alpha}^y$  the relative first order sensitivity and by  $(\alpha - \alpha_0)/\alpha_0 = \delta \alpha$  the relative variation of the parameter  $\alpha$ , then the variability model based on (4.72) defines an *affine* [60] or *additive* model (A):

$$\hat{y}(\alpha) = y_0(1 + S_{\alpha}^y \delta \alpha). \quad (4.73)$$

To ensure a relative validity range of the first order approximation of the output quantity less a given threshold  $t_1$ , the absolute variation of the parameter must be less than

$$V_d = \sqrt{\frac{2y_0 t_1}{D_2}}, \quad (4.74)$$

where  $D_2$  is an upper limit of the second order derivative of the output quantity  $y$  with respect to parameter  $\alpha$  [40].

For the multiparametric case, one gets:

$$y(\alpha) = y(\alpha_0) + \nabla y(\alpha_0) \cdot (\alpha - \alpha_0) = y_0 + \sum_{k=1}^n \frac{\partial y}{\partial \alpha_k}(\alpha_0)(\alpha_k - \alpha_{0k}). \quad (4.75)$$

Similar with one parameter case, the relative sensitivities w.r.t. each parameter are denoted by  $\frac{\partial y}{\partial \alpha_k}(\alpha_0) \frac{\alpha_{0k}}{y_0} = S_{\alpha_k}^y$  and the relative variations of the parameters by  $\delta \alpha_k = (\alpha_k - \alpha_{0k})/\alpha_{0k}$ , the additive model (A) for  $n$  parameters being given by:

$$\hat{y}(\alpha) = y_0 \left( 1 + \sum_{k=1}^n S_{\alpha_k}^y \delta \alpha_k \right). \quad (4.76)$$

Thus, each new independent parameter taken into account adds a new term to the sum [50]. The additive model is simply a normalized standard version of a linearly truncated Taylor expansion.

Instead of using this truncated expansion may be numerically favorable to expand some transformation  $F(y)$  of  $y$  instead. Two particular choices for  $F$  have practical importance: identity and inversion as it will be indicated below.

#### 4.2.6.2 Rational Model (R)

The rational model is the additive model for the reverse quantity  $1/y$ . It is obtained from the first order truncation of the Taylor Series expansion for the function  $1/y$ . For  $n = 1$ , if we denote by  $r(\alpha) = \frac{1}{y(\alpha)}$ , it follows that:

$$\hat{r}(\alpha) = r(\alpha_0) + \frac{\partial r}{\partial \alpha}(\alpha_0)(\alpha - \alpha_0). \quad (4.77)$$

We define the relative first order sensitivity of the reverse circuit function:

$$\frac{\partial r}{\partial \alpha}(\alpha_0) \frac{\alpha_0}{r(\alpha_0)} = S_{\alpha}^r = S_{\alpha}^{1/y}. \quad (4.78)$$

Consequently, we obtain the rational model for  $n = 1$ :

$$y(\alpha) = \frac{y_0}{1 + S_{\alpha}^{1/y} \delta \alpha}. \quad (4.79)$$

It can be easily shown that the reverse relative sensitivity is  $S_{\alpha}^{\frac{1}{y}} = -S_{\alpha}^y$ . For the multiple parameter case, the rational model is:

$$\hat{y}(\alpha) = \frac{y_0}{1 + \sum_{k=1}^n S_{\alpha_k}^{1/y} \delta \alpha_k}. \quad (4.80)$$

If the circuit function  $y$  is for instance the admittance, its inverse  $1/y$  is the impedance. In the time domain, these two transfer functions correspond to a device excited in voltage or in current, respectively. Consequently, the choice between additive and rational models for the variability analysis of the circuit functions in frequency domain can be interpreted as a change in the terminal excitation mode in the time domain state representation. Choosing the appropriate terminal excitation, the validity range of the parametric model based on first order Taylor series approximation can be dramatically extended.

### 4.2.7 Parametric circuit synthesis

We have shown in [51] that one of the most efficient order reduction method for the class of problems we address is the Vector Fitting (VFIT) method proposed in [45], improved in [42, 46] and available at [61]. It finds the transfer function matching a given frequency characteristic. Thus, in the frequency domain, for the output quantity  $y(s)$ , this procedure finds the poles  $p_m$  (real or complex conjugate pairs), the residuals  $k_m$  and the constant terms  $k_0$  and  $k_\infty$  of a rational approximation of the output quantity (e.g an admittance):

$$y(s) \approx y_{VFIT}(s) = k_\infty + sk_0 + \sum_{m=1}^q \frac{k_m}{s - p_m}. \quad (4.81)$$

The resulting approximation has guaranteed stable poles and the passivity can be enforced in a post-processing step [42]. The transfer function (4.81) can be synthesized by using the Differential Equation Macromodel (DEM) [57]. Our aim is to extend DEM to take into consideration the parameterization.

To simplify the explanations, we assume a single input single output system, excited in voltage. It follows that the output current is given by (4.82), where  $x_m(s)$  is a new variable defined by (4.83).

$$i(s) = y(s)u(s) = k_\infty u(s) + sk_0 u(s) + \sum_{m=1}^q k_m x_m(s), \quad (4.82)$$

$$x_m(s) = \frac{u(s)}{s - p_m}. \quad (4.83)$$

By applying the inverse Laplace transformation to (4.82) and (4.83), relations (4.84) and (4.85) are obtained:

$$i(t) = k_\infty u(t) + k_0 \frac{du(t)}{dt} + \sum_{m=1}^q k_m x_m(t), \quad (4.84)$$

$$\frac{dx_m(t)}{dt} = p_m x_m(t) + u(t). \quad (4.85)$$

If we use the following matrix notations

$$A = \text{diag}(p_1, p_2, \dots, p_q), \quad b = [1 \ 1 \ \dots \ 1]^T, \quad (4.86)$$

$$c = [k_1 \ k_2 \ \dots \ k_q]^T \quad x = [x_1 \ x_2 \ \dots \ x_q]^T, \quad (4.87)$$

then equations of the system (4.84), (4.85) can be written in a compact form as

$$\frac{dx(t)}{dt} = Ax(t) + bu(t), \quad (4.88)$$

$$i(t) = k_\infty u(t) + k_0 \frac{du(t)}{dt} + cx(t). \quad (4.89)$$

### 4.2.7.1 Case of real poles

In the case in which all poles (and consequently, all the residuals) are real, eq. (4.84) can be synthesized by the circuit shown in Fig. 4.19 which consists of a capacitor having the capacitance  $k_0$ , in parallel with a resistor having the conductance  $k_\infty$ , in parallel with  $q$  voltage controlled current sources, their parameters being the residuals  $k_m$ .

Eq. (4.85) can be synthesized by the circuit in Fig. 4.20, where  $x_m$  is the voltage across a unity capacitor, connected in parallel with a resistor having the conductance  $-p_m$  and a voltage controlled current source, controlled by the input voltage  $u$ .

We would like to include the parametric dependence into the VFIT model and in the synthesized circuit. To keep the explanations simple, we assume that there is only one parameter that varies, i.e the quantity  $\alpha$  is a scalar. Assuming that keeping the order  $q$  is satisfactory for the whole range of the variation of this parameter, this means that (4.81) can be parameterized as:

$$y(s, \alpha) \approx y_{VFIT}(s, \alpha) = k_\infty(\alpha) + sk_0(\alpha) + \sum_{m=1}^q \frac{k_m(\alpha)}{s - p_m(\alpha)}. \quad (4.90)$$

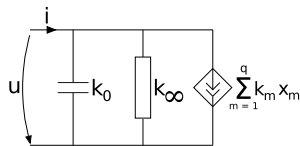
Without loss of generality, we can assume that the additive model is more accurate than the rational one. If not, the reverse quantity is used, which is equivalent, for our class of problems, to change the excitation of terminals from voltage excited to current excited, and use an additive model for the impedance  $z = y^{-1}$ . The additive model (4.73) can be written as

$$y(s, \alpha) \approx y_A(s, \alpha) = y(s, \alpha_0) + \frac{\partial y}{\partial \alpha}(s, \alpha_0)(\alpha - \alpha_0), \quad (4.91)$$

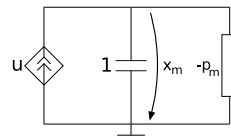
where here  $y$  is a matrix function. By combining (4.90) and (4.91) we obtain an approximate additive model based on VFIT:

$$y(s, \alpha) \approx y_{A-VFIT}(s, \alpha) = y_{VFIT}(s, \alpha_0) + \frac{\partial y_{VFIT}}{\partial \alpha}(s, \alpha_0)(\alpha - \alpha_0). \quad (4.92)$$

From (4.90) it follows that the sensitivity of the VFIT approximation needed in (4.92) is



**Fig. 4.19** Equivalent circuit for the output equation if all poles are real.



**Fig. 4.20** Sub-circuit corresponding to a real pole.



$$\frac{\partial y_{VFIT}}{\partial \alpha} = \frac{\partial k_\infty}{\partial \alpha} + s \frac{\partial k_0}{\partial \alpha} + \sum_{m=1}^q \left[ \frac{\partial k_m / \partial \alpha}{s - p_m} + \frac{k_m}{(s - p_m)^2} \frac{\partial p_m}{\partial \alpha} \right]. \quad (4.93)$$

The sensitivity  $\partial y / \partial \alpha$  can be evaluated with (4.61) for as many frequencies as required and thus the sensitivities of poles and residues in (4.93) can be computed by solving the linear system (4.93) by least square approximation. Finally, by substituting (4.93) and (4.90) in (4.92), the final parameterized and frequency dependent model is obtained:

$$y_{A-VFIT}(s, \alpha) = \left[ k_\infty + (\alpha - \alpha_0) \frac{\partial k_\infty}{\partial \alpha} \right] + s \left[ k_0 + (\alpha - \alpha_0) \frac{\partial k_0}{\partial \alpha} \right] + \sum_{m=1}^q \left[ \frac{k_m + (\alpha - \alpha_0) \frac{\partial k_m}{\partial \alpha}}{s - p_m} \right] + (\alpha - \alpha_0) \sum_{m=1}^q \left[ \frac{k_m}{(s - p_m)^2} \frac{\partial p_m}{\partial \alpha} \right]. \quad (4.94)$$

Expression (4.94) has the advantage that it has an explicit dependence with respect both to the frequency  $s = j\omega$  and parameter  $\alpha$ , is easy to implement and feasible to be synthesized as a net-list having components with dependent parameters, as explained below.

If we denote by

$$k_*(\alpha) = k_* + (\alpha - \alpha_0) \frac{\partial k_*}{\partial \alpha}, \quad (4.95)$$

where  $k_* = k_*(\alpha_0)$  then eq. (4.94) can be written as

$$y_{A-VFIT}(s, \alpha) = y_1(s, \alpha) + y_2(s, \alpha), \quad (4.96)$$

where

$$y_1(s, \alpha) = k_\infty(\alpha) + s k_0(\alpha) + \sum_{m=1}^q \frac{k_m(\alpha)}{s - p_m}, \quad (4.97)$$

$$y_2(s, \alpha) = (\alpha - \alpha_0) \sum_{m=1}^q \frac{k_m}{(s - p_m)^2} \frac{\partial p_m}{\partial \alpha}. \quad (4.98)$$

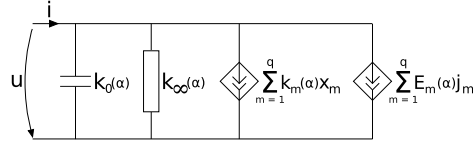
The output current is thus

$$i(s, \alpha) = y_1(s, \alpha)u(s) + y_2(s, \alpha)u(s), \quad (4.99)$$

where the first term can be synthesized with a circuit similar to the one in Fig. 4.19 but where the  $k_*$  parameters depend on  $\alpha$ , and the second term

$$i_2(s, \alpha) = (\alpha - \alpha_0) \sum_{m=1}^q \frac{k_m}{(s - p_m)^2} \frac{\partial p_m}{\partial \alpha} u(s) \quad (4.100)$$

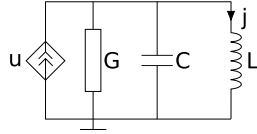
adds  $q$  new parallel branches to the circuit (Fig. 4.21). It is useful to write (4.100) as



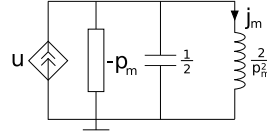
**Fig. 4.21** Parameterized circuit corresponding to the output equation.

$$i_2(s, \alpha) = \sum_{m=1}^q E_m(\alpha) \frac{u(s)}{\left(\frac{s}{p_m} - 1\right)^2}, \quad \text{where} \quad E_m(\alpha) = \frac{(\alpha - \alpha_0) k_m}{p_m^2} \frac{\partial p_m}{\partial \alpha}. \quad (4.101)$$

The part that depends on  $s$  in (4.101) can be synthesized by a second order circuit, such as the one in Fig. 4.22.



**Fig. 4.22** Second order subcircuit, with a voltage controlled current source.



**Fig. 4.23** Second order subcircuit corresponding to a real pole.

The current through the coil is

$$j(s) = \frac{u(s)}{s^2 LC + sLG + 1}. \quad (4.102)$$

To obtain the expression in (4.101) it is necessary that  $LC = 1/p_m^2$ ,  $LG = -2/p_m$ , for instance, we can chose  $G = -p_m$ ,  $L = 2/p_m^2$ ,  $C = 1/2$ . Thus, the parameterized circuit is given by the sub-circuits in Fig. 4.21, Fig. 4.20 and Fig. 4.23. The circuit that corresponds to the output equations has new branches with current controlled current sources. Only this sub-circuit contained parameterized components.

Another possibility to derive a parameterized circuit is to do as follows. In (4.100) we denote by

$$\frac{1}{(s - p_m)^2} \frac{\partial p_m}{\partial \alpha} u(s) = f_m(s), \quad (4.103)$$

and by

$$(s - p_m) f_m(s) = g_m(s). \quad (4.104)$$

Relationships (4.103) and (4.104) are equivalent to

$$s g_m(s) = p_m g_m(s) + \frac{\partial p_m}{\partial \alpha} u(s), \quad (4.105)$$

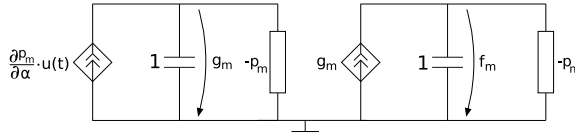
$$s f_m(s) = p_m f_m(s) + g_m(s), \quad (4.106)$$

which correspond in the time domain to

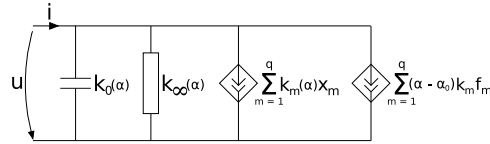
$$\frac{dg_m(t)}{dt} = p_m g_m(t) + \frac{\partial p_m}{\partial \alpha} u(t), \tag{4.107}$$

$$\frac{df_m(t)}{dt} = p_m f_m(t) + g_m(t). \tag{4.108}$$

Eqs. (4.107) and (4.108) can be synthesized with the subcircuit shown in Fig. 4.24.



**Fig. 4.24** Subcircuit corresponding to the second order term (second approach).



**Fig. 4.25** Parameterized circuit corresponding to the output equation (second approach).

In this case the circuit that corresponds to the output equation is the one in Fig. 4.25. In brief, the parameterized reduced order circuit can be either the one in Fig. 4.21, Fig. 4.20 and Fig. 4.23 or in Fig. 4.25, Fig. 4.20 and Fig. 4.24. In both approaches only the circuit that corresponds to the output equation is parameterized. The second approach has the advantage that can be generalized for a transfer function having complex poles as well.

#### 4.2.7.2 Case of complex poles

##### Nominal Differential Equation Macromodel

If some of the  $q$  poles are complex, then they appear in conjugated pairs since they are the roots of the characteristic equation corresponding to a real matrix. We assume for the beginning that the transfer function has only one pair of complex conjugate poles:  $p = a + jb$  and  $p^* = a - jb$ . In this case the transfer function is

$$y(s) = \frac{k_1}{s - p} + \frac{k_2}{s - p^*} = \frac{(s - a)(k_1 + k_2) + jb(k_1 - k_2)}{(s - a)^2 + b^2}. \tag{4.109}$$

The numerator can be a real polynomial in  $s$  only if  $k_1$  and  $k_2$  are complex conjugated residues:  $k_1 = c + jd$ ,  $k_2 = c - jd$ . In this case, the matrices in (4.86) are

$$A = \begin{bmatrix} a+jb & 0 \\ 0 & a-jb \end{bmatrix}, \quad b = [1 \ 1]^T \quad c = [c+jd \ c-jd] \quad x = [x_1 \ x_2]^T. \quad (4.110)$$

In order to obtain a real coefficient equation, a matrix transformation is introduced. The system (4.88) becomes

$$V \frac{dx(t)}{dt} = VAV^{-1}Vx(t) + Vbu(t), \quad (4.111)$$

$$i(t) = k_\infty u(t) + k_0 \frac{du(t)}{dt} + cV^{-1}Vx(t), \quad (4.112)$$

where

$$V = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{j}{\sqrt{2}} \\ \frac{j}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{j}{\sqrt{2}} \\ -\frac{j}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}. \quad (4.113)$$

Let

$$\hat{x} = Vx = [\hat{x}_1 \ \hat{x}_2]^T, \quad \hat{A} = VAV^{-1} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad (4.114)$$

$$\hat{b} = Vb = [-\sqrt{2} \ 0], \quad \hat{c} = cV^{-1} = [-\sqrt{2}c \ \sqrt{2}d]. \quad (4.115)$$

The transformation  $\hat{A} = VAV^{-1}$  is a *similarity transformation*, preserving the eigenvalues of the matrix and thus the characteristic polynomial of the system.

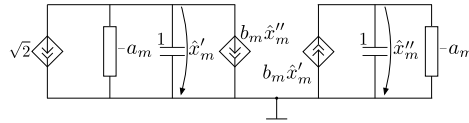
The two equations corresponding to the complex conjugated pair of poles

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p & 0 \\ 0 & p^* \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \quad (4.116)$$

become after applying the similarity transformation

$$\frac{d}{dt} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix} u(t). \quad (4.117)$$

If there are several pairs of complex conjugated poles, equation (4.117) will be true for any of these pairs and, by renaming  $p \rightarrow p_m$ ,  $\hat{x}_1 \rightarrow \hat{x}'_m$ ,  $\hat{x}_2 \rightarrow \hat{x}''_m$ ,  $a \rightarrow a_m$ ,  $b \rightarrow b_m$ , the synthesized circuit is shown in Fig. 4.26.



**Fig. 4.26** Sub-circuit corresponding to a pair of complex conjugate poles.

In general, if the system has  $q$  poles out of which  $q_r$  are real and  $q_c = (q - q_r)/2$  are pairs of complex conjugate poles, then the synthesis will be done as follows: for each real pole  $m = 1, \dots, q_r$ , let  $k_m$  be the residue corresponding to the pole; for

each pair of complex conjugate poles  $m = 1, \dots, q_c$  let the pole be  $p'_m = a_m + jb_m$ , with the corresponding residue  $k'_m = c_m + jd_m$ . An equivalent circuit for the output equation is shown in Fig. 4.27. It consists of the following elements connected in

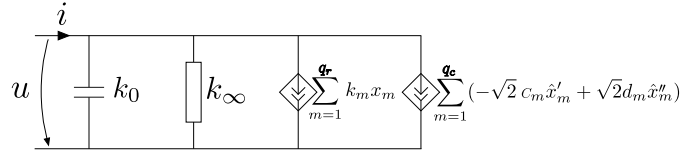


Fig. 4.27 Sub-circuit corresponding to a pair of complex conjugate poles.

parallel:

- a capacitance  $k_0$ ;
- a conductance  $k_\infty$ ,
- $q_r$  voltage controlled current sources (having the parameter  $k_m$ , controlled by the voltages  $x_m$ ),
- $q_c$  voltage controlled current sources (having the parameter  $-\sqrt{2}c_m$ , controlled by the voltages  $\hat{x}'_m$ )
- $q_c$  voltage controlled current sources (having the parameter  $\sqrt{2}d_m$ , controlled by the voltages  $\hat{x}''_m$ ).

The voltages  $x_m$  are defined on the  $q_r$  subcircuits that correspond to real poles (Fig. 4.20) and the voltages  $\hat{x}'_m, \hat{x}''_m$  are defined on the  $q_c$  subcircuits that correspond to the pair of complex conjugate poles (Fig. 4.26).

### Parametric DEM

To derive the parametric circuit in the case of complex poles, we could proceed as we did in the first approach for real poles. This would conduce to a transfer function of order 4, which is not obvious how it can be synthesized. The second approach can be extended to the case of complex poles, as follows.

Let's consider equations (4.107) and (4.108) written for a pair of complex conjugate poles  $p_1 = a + jb, p_2 = a - jb$ :

$$\frac{dg_1(t)}{dt} = p_1 g_1(t) + \frac{\partial p_1}{\partial \alpha} u(t), \quad (4.118)$$

$$\frac{df_1(t)}{dt} = p_1 f_1(t) + g_1(t), \quad (4.119)$$

$$\frac{dg_2(t)}{dt} = p_2 g_2(t) + \frac{\partial p_2}{\partial \alpha} u(t), \quad (4.120)$$

$$\frac{df_2(t)}{dt} = p_2 f_2(t) + g_2(t). \quad (4.121)$$

By using the matrix notations

$$g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad \frac{\partial p}{\partial \alpha} = \begin{bmatrix} \partial p_1 / \partial \alpha \\ \partial p_2 / \partial \alpha \end{bmatrix}, \quad A = \begin{bmatrix} p_1 & 0 \\ 0 & p_2 \end{bmatrix}, \quad (4.122)$$

it follows that (4.118)  $\div$  (4.121) can be written in a compact form as

$$\frac{dg(t)}{dt} = Ag(t) + \frac{\partial p}{\partial \alpha} u(t), \quad (4.123)$$

$$\frac{df(t)}{dt} = Af(t) + g(t), \quad (4.124)$$

and by applying the similarity transformation described in the previous section it follows that

$$\frac{d\hat{g}(t)}{dt} = VAV^{-1}\hat{g}(t) + V\frac{\partial p}{\partial \alpha}u(t), \quad (4.125)$$

$$\frac{d\hat{f}(t)}{dt} = VAV^{-1}\hat{f}(t) + \hat{g}(t), \quad (4.126)$$

where  $VAV^{-1}$  is given by (4.114). It is straightforward to derive that

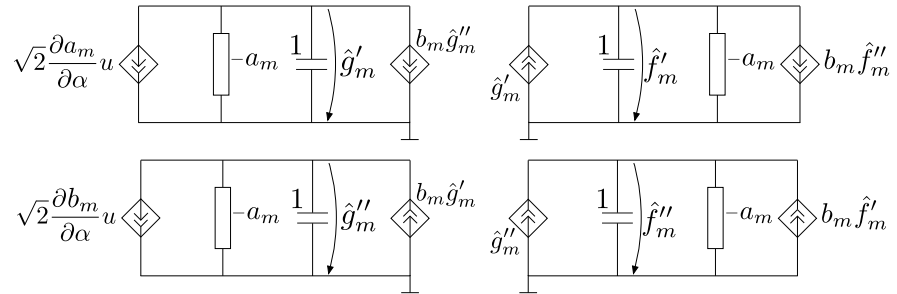
$$V\frac{\partial p}{\partial \alpha} = \begin{bmatrix} -\sqrt{2}\frac{\partial a}{\partial \alpha} & -\sqrt{2}\frac{\partial b}{\partial \alpha} \end{bmatrix}. \quad (4.127)$$

Thus, the equations (4.123) and (4.124) corresponding to the two complex-conjugated poles become after applying the similarity transformation

$$\frac{d}{dt} \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix} = \begin{bmatrix} a-b & \\ b & a \end{bmatrix} \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix} + \begin{bmatrix} -\sqrt{2}\partial a/\partial \alpha \\ -\sqrt{2}\partial b/\partial \alpha \end{bmatrix} u(t), \quad (4.128)$$

$$\frac{d}{dt} \begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \end{bmatrix} = \begin{bmatrix} a-b & \\ b & a \end{bmatrix} \begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \end{bmatrix} + \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix}. \quad (4.129)$$

If there are several pairs of complex conjugated poles, equations above will be true for any of these pairs and, by renaming  $p \rightarrow p_m$ ,  $\hat{g}_1 \rightarrow \hat{g}'_m$ ,  $\hat{g}_2 \rightarrow \hat{g}''_m$ ,  $\hat{f}_1 \rightarrow \hat{f}'_m$ ,  $\hat{f}_2 \rightarrow \hat{f}''_m$ ,  $a \rightarrow a_m$ ,  $b \rightarrow b_m$ , the synthesized circuit is shown in Fig. 4.28.



**Fig. 4.28** Sub-circuit corresponding to a pair of complex conjugate poles.

The new terms added in the output equations are

$$\begin{aligned} i_2(s, \alpha) &= (\alpha - \alpha_0) \begin{bmatrix} k & k^* \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = (\alpha - \alpha_0) \begin{bmatrix} k & k^* \end{bmatrix} V^{-1} \hat{f} = \\ &= (\alpha - \alpha_0) \begin{bmatrix} -\sqrt{2}c & \sqrt{2}d \end{bmatrix} \hat{f} = -\sqrt{2}c(\alpha - \alpha_0)\hat{f}_1 + \sqrt{2}d(\alpha - \alpha_0)\hat{f}_2. \end{aligned} \quad (4.130)$$

In general, if the system has  $q$  poles out of which  $q_r$  are real and  $q_c = (q - q_r)/2$  are pairs of complex conjugate poles, then the parametric synthesis will be done as follows: for each real pole  $m = 1, \dots, q_r$ , let  $k_m$  be the residue corresponding to the pole; for each pair of complex conjugate poles  $m = 1, \dots, q_c$  let the pole be  $p'_m = a_m + jb_m$ , with the corresponding residue  $k'_m = c_m + jd_m$ . The equivalent circuit for the parametric output equation is shown in Fig. 4.29. It consists of the

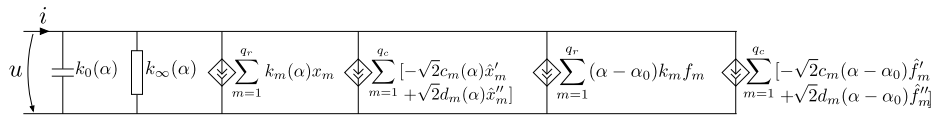


Fig. 4.29 Parametric sub-circuit corresponding to the output equation.

following elements connected in parallel:

- a parameterized capacitance  $k_0(\alpha) = k_0 + (\alpha - \alpha_0)\partial k_0/\partial\alpha$ ,
- a parameterized conductance  $k_\infty(\alpha) = k_\infty + (\alpha - \alpha_0)\partial k_\infty/\partial\alpha$ ,
- $q_r$  voltage controlled current sources (having as parameter the parameterized value  $k_m(\alpha) = k_m + (\alpha - \alpha_0)\partial k_m/\partial\alpha$ , controlled by the voltages  $x_m$ ),
- $q_c$  voltage controlled current sources (having as parameter the parameterized value  $-\sqrt{2}c_m(\alpha)$ , controlled by the voltages  $\hat{x}'_m$ ),
- $q_c$  voltage controlled current sources (having the parameter  $\sqrt{2}d_m(\alpha)$ , controlled by the voltages  $\hat{x}''_m$ ),
- $q_r$  voltage controlled voltage sources (having the parameter  $(\alpha - \alpha_0)k_m$ , controlled by the voltages  $f_m$ ),
- $q_c$  voltage controlled current sources (having the parameter  $-\sqrt{2}c_m(\alpha - \alpha_0)$ , controlled by the voltages  $\hat{f}'_m$ ),
- $q_c$  voltage controlled current sources (having the parameter  $\sqrt{2}d_m(\alpha - \alpha_0)$ , controlled by the voltages  $\hat{f}''_m$ ).

The voltages  $x_m$  are defined on the  $q_r$  subcircuits that correspond to real poles (Fig. 4.20), the voltages  $\hat{x}'_m, \hat{x}''_m$  are defined on the  $q_c$  subcircuits that correspond to the pair of complex conjugate poles (Fig. 4.26), the voltages  $f_m$  are defined on the  $q_r$  subcircuits that correspond to real poles (Fig. 4.24), the voltages  $\hat{f}'_m$  and  $\hat{f}''_m$  are defined on the  $q_c$  subcircuits that correspond to the complex poles (Fig. 4.28).

### 4.2.8 Case Study

In order to validate our approach and to evaluate different parametric models which can be extracted by the proposed procedure, several experiments have been performed on a test structure that consists of a microstrip (MS) transmission line having one Aluminum conductor embedded in a  $\text{SiO}_2$  layer. The line has a rectangular cross-section, parameterized by several parameters (Fig. 4.30). The return path is the grounded surface placed at  $y = 0$ . The nominal values used are:  $h_1 = 1\mu\text{m}$ ,  $h_2 = 0.69\mu\text{m}$ ,  $h_3 = 10\mu\text{m}$ ,  $a = 130.5\mu\text{m}$ ,  $p_1 = h_1$ ,  $p_2 = h_2$ ,  $p_3 = 3\mu\text{m}$ ,  $x_{\text{max}} = 264\mu\text{m}$ . In order to comply with designer's requirements, the model should include the field propagation along the line, taking into consideration the distributed parameters and the high frequency effects.

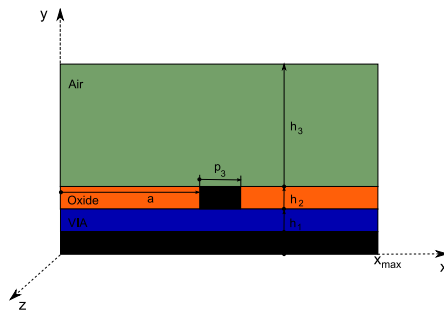


Fig. 4.30 Stripline parameterized structure.

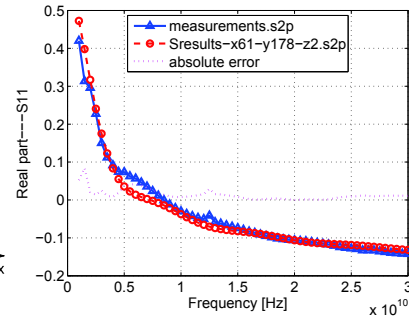


Fig. 4.31 Frequency characteristic  $\text{Re}(S_{11})$ : numerical model vs measurements.

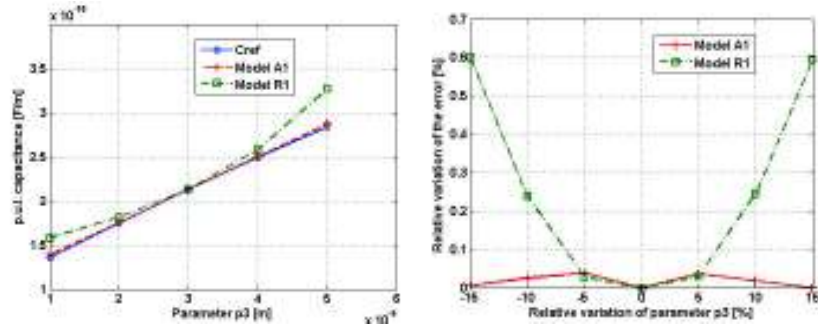
#### 4.2.8.1 Validation of the nominal model

The first step of the validation refers to the simulation of the nominal case for which measurements (S parameters) are available from the European project FP5/Codestar (<http://www.magwel.com/codestar/>). By using dFIT + dELOB [50], at low frequencies, the following values are obtained:

$$R = 18.11\text{k}\Omega/\text{m}, \quad L = 322\text{nH}/\text{m}, \quad C = 213\text{pF}/\text{m}, \quad (4.131)$$

which are coherent with the values obtained from the measurements at low frequencies, and validates the grid used and the extension of the boundary used in the numerical model. Then, by using the method described in section 4.2.4 the dependence of p.u.l. parameters with respect to the frequency was computed. The comparison between the resulting S parameters and the measurements is shown in Fig. 4.31 and it validates the nominal model. The sensitivities of the p.u.l. parameters are computed using the CHAMY software [62], by direct differentiation method applied





**Fig. 4.32** Left: Reconstruction of the p.u.l. C from Taylor Series first order expansion; Right: Relative error w.r.t. the relative variation of parameter  $p_3$ .

to the state space equations [40]. They could also be computed by Adjoint Field Technique (AFT) [38, 52].

#### 4.2.8.2 Parametric Models

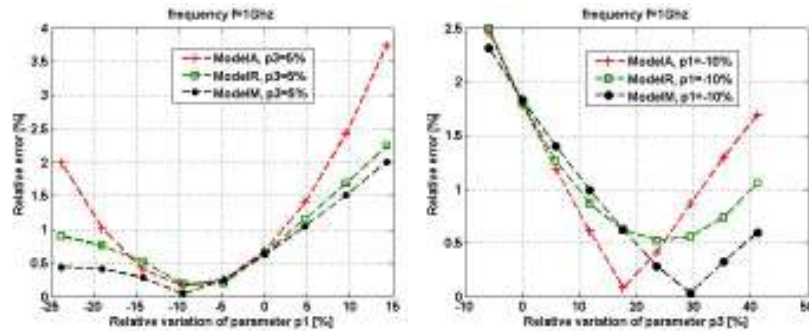
In this section, the accuracy of several parametric models for the line capacitance is investigated.

The first sets of tests considered only one parameter that varies, namely the width of the line,  $p_3$ . The nominal value chosen was  $p_3 = 3\mu\text{m}$  and samples in the interval  $[1, 5]\mu\text{m}$  were considered. The reference result was obtained by simulated the samples separately (each sample was discretized and solved). These were compared with the approximate values obtained from models A and R (Fig. 4.32). As expected intuitively, the dependence w.r.t.  $p_3$  is almost linear and the A model is better than the R model. Considering the relative variation of the parameters less than 15% (which is the typical limit for the technological variations nowadays) the relative variation of the output parameter is obtained (Fig. 4.32, right). The errors of both affine and rational first order models for p.u.l. parameters are given in Table 4.2. Model A based on the first order Taylor series approximation has a maximal error for technologic variations 1.78% for p.u.l. resistance when  $p_3$  is variable, whereas model R has an approximation error of only 0.6% for the same range of the technological variations for p.u.l. capacitance when  $p_3$  is variable. Using (4.74) one can be easily identify which is the best model for any case.

The second set of tests considered two parameters that vary simultaneously:  $p_1$  and  $p_3$ . For reference, a set of samples in  $[0.8, 1.2]\mu\text{m} \times [2, 4]\mu\text{m}$  were considered. The p.u.l. capacitance was approximated using the additive, rational and multiplicative models described above. In this case, a new model M is computed using an additive model for  $p_3$  and a rational one for  $p_1$ , which is the best choice. Fig. 4.33, left compares the relative variation of the errors w.r.t. a relative variation of parameter  $p_1$  for a variation of  $p_3$  of 5%. Model M provides lower errors (maximum error is 2%) than models A (3.7%) and R (2.2%). Fig. 4.33, right illustrates that in the

**Table 4.2** Maximal errors [%] of p.u.l. parameters for technology variation of  $\pm 15\%$ .

Parameter	Quantity	Affine (A1)	Rational (R1)
$p_1$	$L$	0.11	0.15
	$C$	0.65	0.25
$p_3$	$R$	1.78	0.22
	$L$	0.34	0.04
	$C$	0.035	0.6

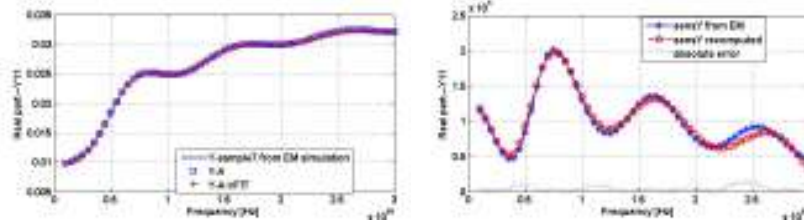
**Fig. 4.33** Left: Relative error w.r.t. the relative variation of parameter  $p_1$ , for a variation of  $p_3$  of 5%; Right: Relative error w.r.t. the relative variation of parameter  $p_3$ , for a variation of  $p_1$  of 10%

range from 20% to 40% model M is the best one if we look at the variation w.r.t.  $p_3$  for a variation of  $p_1$  of 10%.

Thus, by using the appropriate multiplicative models in the modeling of the technological variability, the necessity of higher order approximations can be eliminated.

### 4.2.8.3 Frequency dependent parametric models

In this case, the parameter considered variable is  $h_2$ . The sensitivity of the admittance with respect to this parameter has been calculated according to (4.61), using EM field solution. By applying Vector Fitting, a transfer function with 8 poles has been obtained. This conduced to an overdetermined system of size (236, 26) which has been solved with an accuracy (relative residual) of 3.7% (Fig.4.34-left). Finally, the relative error of the A-VFIT model is 1.09 % compared to the relative error of the A model which is 0.95 % for a relative variation of the parameter of 10 % (in Fig. 4.34-right the three curves are on top of each other).



**Fig. 4.34** Left: Variation of the admittance sensitivity with respect to the frequency; Right: Reference simulation vs. answer obtained from the frequency dependent parametric model (4.94).

### 4.2.9 Conclusions

The paper describes an effective procedure to extract reduced order parametric models of on-chip interconnects allowing model order reduction in coupled field (PDE) - circuit (DAE) problems. These models consider all EM field effects at high frequency, described by 3D-FW Maxwell equations. The proposed procedure is summarized by the following steps:

- Step 1 - Solve two field problems (2D EQS and FW-TM) and compute frequency dependent p.u.l. parameters and their sensitivities with respect to the geometric parameters that vary;
- Step 2 - Compute admittance for the real length of the line and its sensitivities with respect to the variable parameters;
- Step 3 - Choose the A/R variation model, i.e. the appropriate terminal excitation (admittance or impedance);
- Step 4 - Apply Vector Fitting for the nominal case in order to extract a rational model for the circuit function with respect to the frequency;
- Step 5 - Compute sensitivities of poles and residues of the circuit function by solving a least square problem;
- Step 6 - Assemble the frequency dependent parametric model by using the compact expression (4.94) or by synthesis of a SPICE like parametric netlist having frequency constant parameters.

Step 1 is dedicated to the extraction of the frequency dependent p.u.l. line parameters in a more robust and flexible way than the inversion of the equation of the short line segment. It is based on the solving of two field problems: 2D-EQS field which describes the transversal effects such as capacitive coupling whereas EMQS-TM field describes the longitudinal effects such as inductive, skin effect and eddy currents. The longitudinal propagation is described by the classic TL equations, but with frequency dependent p.u.l. line parameters.

Then (step 2), variability models for TL structures considering the dependency of p.u.l. parameters w.r.t. geometric parameters, at a given frequency were analyzed. A detailed study of the line sensitivity was made by using numeric techniques. For one parameter case, the proposed methods avoid the evaluation of higher order sensitivities, but keeping a high level of accuracy by introducing models based on a rational

approximation in the frequency domain. The multi-parametric case has been analyzed, in addition, a multiplicative parametric model (M) has been proposed. This is based on the assumption that the quantity of interest can be expressed with separated variables, for which A and/or R models are used. Model M is sometimes better than A and R models obtained from Taylor Series expansion. Its specific terms (products of first order sensitivities) can thus approximate higher order, cross-terms of Taylor Series. In order to automatically select the best first order model for a multi-parametric problem, the validity ranges of direct and reversed quantities have to be evaluated (step 3). Once we establish the best model (A or R) for each parameter, the M model will be easily computed by multiplication of individual submodels. Our numerical experiments with the proposed algorithm in all particular structures we investigated prove that the technological variability (e.g.  $\pm 20\%$  variation of geometric parameters, which is typical for the technology node of 65 nm) can be modeled with acceptable accuracy (relative errors under 5%) using only first order parametric models for line parameters. This is one of the most important results of our research.

Next, a rational approximation in the frequency domain, obtained with Vector Fitting (step 4) is combined with a first order Taylor Series approximation. The sensitivities of poles, residues and constant terms are computed by solving an over-determined system of linear equations (step 5). The main advantage of this approach is that the final result is amenable to be synthesized with a small parameterized circuit (step 6). This method relies on the differential equation macromodel which is extended in order to take into account the variability. It also assumes that a first order Taylor Series expansion for the parameter that varies is accurate enough for the frequency range of interest. As shown in our previous work, there is a specific excitation type of terminals for which this assumption is acceptable for a certain frequency range. The passivity of the obtained circuit is guaranteed by the fact that the transfer function used as input for the synthesis procedure is passive as it is obtained by a fitting procedure with passivity enforcement.

Thus, the proposed method allows one to obtain parameterized reduced order circuits, having equivalent behavior as on-chip interconnects. These equivalent circuits described in SPICE language are extracted by considering all electromagnetic field effects in interconnects at very high frequency. This method applied to extract the reduced order model of the system described by PDE is a robust and efficient one, being experimentally validated.

The advantages of the proposed approach are:

- its high accuracy, due to the consideration of all field effects at high frequency;
- fast model extraction due to the reduced order of degrees of freedom in the numerical approach;
- high efficiency of the model order reduction step due to the use of Vector Fitting; in all interconnect studied cases, extracted models with an order less than 10 had an acceptable accuracy for designers.
- simple geometric variability models based only on first order sensitivities, with extended valability domain due to the appropriate excitation;
- appropriate variation model for frequency and length of interconnects, due to the transmission lines approach;

- the reduced SPICE models are simple and compact, containing ideal linear elements with lumped frequency independent constant parameters: capacitances, resistances and voltage controlled current sources; these element parameters have very simple affine variation in the case of the geometric variability.

The proposed method was successfully applied to model technological variability, without being necessary the use of higher order sensitivities.

**Acknowledgements** The authors gratefully acknowledge the financial support offered during several years by the following complementary projects: FP6/COMSON, FP6/Chameleon, FP5/Codestar, CEEX/nEDA, UEFISCSU/IDEI 609/16.01.2009 and POSDRU/89/1.5/S/62557.

## References for Section 4.2

38. BI, Y., VAN DER MEIJS, N., IOAN, D.: Capacitance sensitivity calculation for interconnects by adjoint field technique. In: *Proc. 12th IEEE Workshop on Signal Prop. on Interconnects (SPI-2008)*, pp. 1–4 (2008)
39. BOSSAVIT, A.: Most general non-local boundary conditions for the Maxwell equations in a bounded region. *COMPEL: The Int. J. for Computation and Mathematics in Electrical and Electronic Engineering*, **19**(2), pp. 239–245 (2000)
40. CIUPRINA, G., IOAN, D., NICULAE, D., FERNANDEZ VILLENA, J., SILVEIRA, L.: Parametric models based on sensitivity analysis for passive components. In: S. Wiak, A. Krawczyk, and I. Dolezel (eds.), *Intelligent Computer Techniques in Applied Electromagnetics, Studies in Computational Intelligence*, **119**, Springer, pp. 231–239 (2008)
41. CLEMENS, M., WEILAND, T.: Discrete electromagnetism with the finite integration technique. *Progress in Electromagnetics Research (PIER)*, **32**, pp. 65–87 (2001)
42. DESCHRIJVER, D., MROZOWSKI, M., DHAENE, T., DE ZUTTER, D.: Macromodeling of multiport systems using a fast implementation of the vector fitting method. *IEEE Microwave and Wireless Components Letters*, **18**(6), pp. 383–385 (2008)
43. FERRANTI, F., ANTONINI, G., DHAENE, T., KNOCKAERT, L.: Passivity-preserving parameterized model order reduction for PEEC based full wave analysis. In: *Proc. 14th IEEE Workshop on Signal Prop. on Interconnects*, pp. 65–68 (2010)
44. GOEL, A.: High Speed VLSI Interconnections. Wiley Series in Microwave and Optical Engineering (2008)
45. GUSTAVSEN, B., SEMLYEN, A.: Rational approximation of frequency domain responses by vector fitting. *IEEE Trans. Power Delivery*, **14**(3), pp. 1052–1061 (1999)
46. GUSTAVSEN, B.: Improving the pole relocating properties of vector fitting. *IEEE Trans. Power Delivery*, **21**(3), pp. 1587–1592 (2006)
47. IOAN, D., CIUPRINA, G., RADULESCU, M.: Algebraic sparsefied partial equivalent electric circuit - ASPEEC. In: A.M. Anile, G. Ali, and G. Mascali (eds.): *Scientific Computing in Electrical Engineering*, Series Mathematics in Industry **9**, Springer, pp. 45–50 (2006)
48. IOAN, D., CIUPRINA, G., RADULESCU, M.: Absorbing boundary conditions for compact modeling of on-chip passive structures. *COMPEL: The Int. J. for Computation and Mathematics in Electrical and Electronic Engineering*, **25**(3), pp. 652–659 (2006)
49. IOAN, D., CIUPRINA, G., RADULESCU, M., SEEBACHER, E.: Compact modeling and fast simulation of on-chip interconnect lines. *IEEE Trans. on Magnetics*, **42**(4), pp. 547–550 (2006)
50. IOAN, D., CIUPRINA, G., KULA, S.: Reduced order models for HF interconnect over lossy semiconductor substrate. In: *Proc. 11th IEEE Workshop on SPI*, pp. 233–236 (2007)

51. IOAN, D., CIUPRINA, G.: Reduced order models of on-chip passive components and interconnects, workbench and test structures. In: W.H.A. Schilders, H.A. van der Vorst, J. Rommes (eds.): *Model Order Reduction: Theory, Research Aspects and Applications*, Mathematics in Industry **13**, Springer-Verlag, pp. 447–467 (2008)
52. IOAN, D., CIUPRINA, G., SCHILDERS, W.: Parametric models based on the adjoint field technique for RF passive integrated components. *IEEE Trans. on Magnetics*, **44**(6), pp. 1658–1661 (2008)
53. IOAN, D., SCHILDERS, W., CIUPRINA, G., VAN DER MEIJS, N., SCHOENMAKER, W.: Models for integrated components coupled with their environment. *COMPEL: The Int. J. for Computation and Mathematics in Electrical and Electronic Engineering*, **27**(4), pp. 820–828 (2008)
54. KINZELBACH, H.: Statistical variations of interconnect parasitics: Extraction and circuit simulation. In: *Proc. 10th IEEE Workshop on Signal Prop. on Interconnects (SPI)*, pp. 33–36 (2006)
55. S. KULA: *Reduced order models of interconnects in high frequency integrated circuits*. Ph.D.-Thesis, Politehnica Univ. of Bucharest (2009)
56. VAN DER MEIJS, N., FOKKEMA, J.: VLSI circuit reconstruction from mask topology. *INTEGRATION, the VLSI Journal*, **2**(3), pp. 85–119 (1984)
57. PALENIUS, T., ROOS, J.: Comparison of reduced-order interconnect macromodels for time-domain simulation. *IEEE Trans. Microwave Theory and Techniques*, **52**(9), pp. 2240–2250 (2004)
58. RĂDULEȚ, R., TIMOTIN, A., ȚUGULEA, A.: The propagation equations with transient parameters for long lines with losses. *Rev. Roum. Sci. Tech.*, **15**(4), pp. 585–599 (1979)
59. ȘTEFĂNESCU, A.: *Parametric Models for Interconnections from Analogue High Frequency Integrated Circuits*. Ph.D.-Thesis, Politehnica Univ. of Bucharest (2009)
60. ȘTEFĂNESCU, A., IOAN, D., CIUPRINA, G.: Parametric models of transmission lines based on first order sensitivities. In: J. Roos, L.R.J. Costa (eds): *Scientific Computing in Electrical Engineering 2008*, Mathematics in Industry **14**, Springer, pp. 29–36 (2010)
61. THE VECTOR FITTING WEB SITE, <http://www.energy.sintef.no/produkt/VECTFIT/index.asp>
62. CHAMELEON-RF WEBSITE, <http://www.chameleon-rf.org>

### 4.3 Model Order Reduction and Sensitivity Analysis

<sup>3</sup>Several types of parameters  $\mathbf{p} = (\mathbf{d}, \mathbf{s}, \theta)$  influence the behaviour of electronic circuits and have to be taken into account when optimizing appropriate performance functions  $f(\mathbf{p})$ : design parameters  $\mathbf{d}$ , manufacturing process parameters  $\mathbf{s}$ , and operating parameters  $\theta$ . The impact of changes of design parameters, e. g., the width and length of transistors or the values of resistors, plays a key role in the design of integrated circuits. Deviations from the nominal values defined in the design phase arise in the manufacturing process. Hence, to guarantee that the physical circuit shows the performances that were specified, the design has to be robust with respect to variations in the manufacturing phase. It has to be analysed how sensitive to parameter changes an integrated circuit and its performance is.

The manufacturing process parameters have a statistical impact, f.i., for the oxide

---

<sup>3</sup> Section 4.3 has been written by: Michael Striebel, Roland Pulch, E. Jan W. ter Maten, Zoran Ilievski, and Wil H.A. Schilders. Of parts of this Section extensions can be found in the Ph.D.-Thesis of the fourth author [95].

thickness threshold. Examples of operating parameters are temperature and supply voltage.

Sensitivity can ease calculations on statistics (for instance by including the sensitivity in calculating the standard deviation of quantities that nearly linearly depend on independent normal distributed parameters [91]: if  $F(p) \approx F_0 + Ap$  with  $p_i \sim N(0, \sigma_i)$  then  $\sigma^2(F_i) \approx \sum_j a_{ij}^2 \sigma_j^2$  ( $\sigma(F_i)$ , and  $\sigma_j$  being the standard deviation of  $F_i$  and  $p_j$ , resp.).

For optimizing one wants to minimize a performance function  $f(\mathbf{p})$  while also several constraints have to be satisfied. The performance function  $f(\mathbf{p})$  and the constraint functions  $c(\mathbf{p})$  can be costly to evaluate and are subject to noise (for instance due to numerical integration effects). For both, the dependency on  $\mathbf{p}$  can be highly nonlinear. Here there is interest in derivative free optimization [102], or to response surface model techniques [79, 80, 92, 115]. Partly these approaches started because in circuit simulation, sensitivities of  $f(\mathbf{p})$  and  $c(\mathbf{p})$  with respect to  $\mathbf{p}$  are not always provided (several model libraries do not yet support the calculation of sensitivities). However, when the number of parameters increases adjoint sensitivity methods become of interest [73, 74]. For transient integration of linear circuits this is described in [76, 77]. In [96] a more general procedure is described that also applies to nonlinear circuits and retains efficiency by exploiting (nonlinear) techniques from Model Order Reduction.

A special sensitivity problem arises in verification of a design after layouting. During the verification the original circuit is extended by a huge number of ‘parasitics’, linear elements that generate additional couplings to the system. To reduce their effect the dominant parasitics should be detected in order to modify the layout.

Adjoint equations are also used for goal achievement. One example is in global error estimation in numerical integration [75, 99].

In this Section we describe adjoint techniques for sensitivity analysis in the time domain and indicate how MOR techniques like POD (Proper Orthogonal Decomposition) may fit here. Next we give a short introduction into Uncertainty Quantification which techniques provide an alternative way to perform sensitivity analysis. Here pMOR (parameterized MOR) techniques can be exploited.

### 4.3.1 Recap MNA and time integration of circuit equations

Modified Nodal Analysis (MNA) is commonly applied to model electrical circuits [86, 90]. Including the parameterization the dynamical behaviour of a circuit is then described by network equations of the general form

$$\frac{d}{dt} \mathbf{q}(\mathbf{x}(t, \mathbf{p}), \mathbf{p}) + \mathbf{j}(\mathbf{x}(t, \mathbf{p}), \mathbf{p}) = \mathbf{s}(t, \mathbf{p}). \quad (4.132a)$$

The state variables  $\mathbf{x}(t, \mathbf{p}) \in \mathbb{R}^n$ , i. e., the potentials at the network’s nodes and the currents through inductors and current sources, are the unknowns in this system. They depend implicitly on the parameters gathered in the vector  $\mathbf{p} \in \mathbb{R}^{n_p}$ ,

because the voltage-charge and current-flux relations of capacitors and inductors, subsumed in  $\mathbf{q}(\cdot, \cdot) \in \mathbb{R}^n$ , the voltage-current relations of resistive elements, appearing in  $\mathbf{j}(\cdot, \cdot) \in \mathbb{R}^n$  and the source terms  $\mathbf{s}(\cdot, \cdot) \in \mathbb{R}^n$ , i. e., the excitation of the circuit, may depend on the parameterization. The elements' characteristics  $\mathbf{q}, \mathbf{j}$  are usually nonlinear in the state variables  $\mathbf{x}$ , e. g., when transistors or diodes are present in the design at hand.

If, however, all elements behave linear with respect to  $\mathbf{x}$ , the MNA equations are of the form

$$\mathbf{C}(\mathbf{p})\dot{\mathbf{x}}(t, \mathbf{p}) + \mathbf{G}(\mathbf{p})\mathbf{x}(t, \mathbf{p}) = \mathbf{s}(t, \mathbf{p}), \quad (4.132b)$$

where  $\dot{\mathbf{x}}$  denotes total differentiation,  $(d/dt)\mathbf{x}(t, \mathbf{p})$  with respect to time.  $\mathbf{C}(\mathbf{p})$  and  $\mathbf{G}(\mathbf{p})$  are real  $n \times n$ -matrices that might depend nonlinearly on the parameters  $\mathbf{p}$ .

Usually the network equations (4.132) state a system of Differential-Algebraic Equations (DAEs), i. e.,  $(\partial/\partial\mathbf{x})\mathbf{q}(\mathbf{x}, \mathbf{p})$  (or  $\mathbf{C}(\mathbf{p})$ ) does not have full rank along the solution trajectory  $\mathbf{x}(t, \mathbf{p})$ .

In transient analysis the network equations (4.132) are solved on a time-interval  $[t_0, t_{\text{end}}] \subset \mathbb{R}$ , where the parameter vector  $\mathbf{p}$  is fixed and a (consistent) initial value  $\mathbf{x}_{0, \mathbf{p}} := \mathbf{x}(t_0, \mathbf{p}) \in \mathbb{R}^n$  is chosen. As the system can usually not be solved exactly, numerical integration, e. g., BDF (backward differentiation formulas) or RK (Runge-Kutta) methods are used to compute approximations  $\mathbf{x}_{i, \mathbf{p}} \approx \mathbf{x}(t_i, \mathbf{p})$  to the state variables on a discrete timegrid  $\{t_0, \dots, t_l, \dots, t_K = t_{\text{end}}\}$ . For a timestep  $h$  from  $t_{l-1}$  to  $t_l = t_{l-1} + h$ , multistep methods, like the BDF schemes, approximate the time derivative  $\frac{d}{dt}\mathbf{q}(\mathbf{x}(t_l, \mathbf{p}), \mathbf{p})$  by some  $k$ -stage operator  $\rho\mathbf{q}(\mathbf{x}_l, \mathbf{p}) := \alpha\mathbf{q}(\mathbf{x}_l, \mathbf{p}) + \beta$ , where  $\alpha = \alpha(h) \in \mathbb{R}$  is the integration coefficient and  $\beta \in \mathbb{R}^n$  is made up of history terms  $\mathbf{q}(\mathbf{x}_\mu, \mathbf{p})$  at the timepoints  $t_\mu$  for  $\mu = l-k, \dots, l-1$ . For the backward Euler method, e. g., we have

$$\rho\mathbf{q}(\mathbf{x}_l, \mathbf{p}) := \underbrace{\frac{1}{h}\mathbf{q}(\mathbf{x}_l, \mathbf{p})}_{=\alpha} - \underbrace{\frac{1}{h}\mathbf{q}(\mathbf{x}_{l-1}, \mathbf{p})}_{=\beta}.$$

This results at each discretisation point  $t_l \in \{t_0, t_1, \dots, t_K\}$  in a nonlinear equation for the state variable  $\mathbf{x}_{l, \mathbf{p}}$  of the form

$$\alpha\mathbf{q}(\mathbf{x}_{l, \mathbf{p}}, \mathbf{p}) + \beta + \mathbf{j}(\mathbf{x}_{l, \mathbf{p}}, \mathbf{p}) = \mathbf{s}(t_l, \mathbf{p}). \quad (4.133)$$

This nonlinear problem is usually solved with some Newton-Raphson technique, where in each underlying iteration  $v = 1, 2, \dots$  a linear system with a system matrix of the form

$$\mathbf{J}(\mathbf{x}_{l, \mathbf{p}}^{(v)}) = \left( \alpha \frac{\partial \mathbf{q}(\cdot, \mathbf{p})}{\partial \mathbf{x}} + \frac{\partial \mathbf{j}(\cdot, \mathbf{p})}{\partial \mathbf{x}} \right) (\mathbf{x}_{l, \mathbf{p}}^{(v)}) \quad (4.134)$$



appears. Typically, simplified Newton-Raphson iterations may be applied. i.e., only the evaluation at  $\mathbf{x}_{t,\mathbf{p}}^{(1)}$  is involved. Note, also when applying a onestep method, like an RK-scheme, linear systems, made up of the Jacobian (4.134) arise.

### 4.3.2 Sensitivity Analysis

We encounter that the state variables  $\mathbf{x}(t, \mathbf{p})$  implicitly depend on the the parameters  $\mathbf{p} \in \mathbb{R}^{n_p}$ . Hence, one is interested in how sensitive the behavior of the circuit with respect to variations in the parameters is. Thinking about "behavior of the system" we can basically have in mind :

- (i) How do the state variables vary with varying parameters?
- (ii) How do measures derived from the state variables, e. g., the power consumption change with varying parameter?

Furthermore, due to usually nonlinear dependence of the element characteristics  $\mathbf{q}$  and  $\mathbf{j}$  or  $\mathbf{C}$  and  $\mathbf{G}$ , respectively, on the parameters, we are interested in variations around a nominal value  $\mathbf{p}_0 \in \mathbb{R}^{n_p}$ .

In the following we will give a brief overview on the different kinds of sensitivities and how they can be treated numerically. For further reading we refer to the PhD thesis by Ilievsky [95] and the papers by Daldoss et al [78], Hovecar et al. [93], Cao et al. [73, 74] and Ilievski et al. [96].

#### 4.3.2.1 State Sensitivity

In (transient) state sensitivity one is interested, how the trajectories of the state variables  $\mathbf{x}$  vary with respect to the parameters  $\mathbf{p}$  around the nominal setting  $\mathbf{p}_0$ . Hence, the goal is to compute

$$\chi_{\mathbf{p}_0}(t) := \left. \frac{d\mathbf{x}(t, \mathbf{p})}{d\mathbf{p}} \right|_{\mathbf{p}=\mathbf{p}_0} \in \mathbb{R}^{n \times n_p}, \quad \text{for all } t \in [t_0, t_{\text{end}}]. \quad (4.135)$$

As described by Daldoss et al. [78] we linearize the nonlinear network equations (4.132b) around the nominal parameter set  $\mathbf{p}_0$ , i. e., we differentiate with respect to  $\mathbf{p}$ . We assume that the element functions  $\mathbf{q}$ ,  $\mathbf{j}$  are sufficiently smooth such that we can exchange the order of differentiation (Schwarz's theorem) and get:

$$\frac{d}{dt} \left[ \mathbf{C}_x(t) \cdot \chi_{\mathbf{p}_0}(t) \right] + \mathbf{G}_x(t) \cdot \chi_{\mathbf{p}_0}(t) = \mathbf{S}_p(t) - \left( \frac{d}{dt} \mathbf{C}_p(t) + \mathbf{G}_p(t) \right) \quad (4.136)$$

$$\begin{aligned}
\text{with } \mathbf{C}_x(t) &:= \frac{\partial \mathbf{q}}{\partial \mathbf{x}}(\mathbf{x}(t, \mathbf{p}_0), \mathbf{p}_0), & \mathbf{C}_p(t) &:= \frac{\partial \mathbf{q}}{\partial \mathbf{p}}(\mathbf{x}(t, \mathbf{p}_0), \mathbf{p}_0), \\
\mathbf{G}_x(t) &:= \frac{\partial \mathbf{j}}{\partial \mathbf{x}}(\mathbf{x}(t, \mathbf{p}_0), \mathbf{p}_0), & \mathbf{G}_p(t) &:= \frac{\partial \mathbf{j}}{\partial \mathbf{p}}(\mathbf{x}(t, \mathbf{p}_0), \mathbf{p}_0), \\
\mathbf{S}_p(t) &:= \frac{\partial \mathbf{s}}{\partial \mathbf{p}}(t, \mathbf{p}_0),
\end{aligned}$$

where  $\mathbf{C}_x(t), \mathbf{G}_x(t) \in \mathbb{R}^{n \times n}$  and  $\mathbf{C}_p(t), \mathbf{G}_p(t), \mathbf{S}_p(t) \in \mathbb{R}^{n \times n_p}$  and  $\mathbf{x}(t, \mathbf{p}_0)$  solves the network problem (4.132a).

The initial sensitivity value  $\chi_{\mathbf{p}_0}(t_0) =: \chi_{\mathbf{p}_0,0} =: \chi_{\mathbf{p}_0}^{\text{DC}}$  can easily be calculated as the sensitivity of the DC-solution  $\mathbf{x}(0, \mathbf{p}_0) =: \mathbf{x}_{\mathbf{p}_0}^{\text{DC}}$  of the network equation (4.132a), satisfying

$$\mathbf{j}(\mathbf{x}_{\mathbf{p}_0}^{\text{DC}}, \mathbf{p}_0) = \mathbf{s}(t_0, \mathbf{p}_0). \quad (4.137)$$

Obviously, the DAE (4.136) states a *linear* time varying dynamical system for the state sensitivity  $\chi_{\mathbf{p}_0}$ , even when (4.132a) was nonlinear. We assume that we have used the backward Euler method to solve the network problem (4.132). Using the same time grid for solving the state sensitivity problem (4.136), we advance from time point  $t_{l-1}$  to  $t_l = t_{l-1} + h$ , i. e., we compute  $\chi_{\mathbf{p}_0,l} \approx \chi_{\mathbf{p}_0}(t_l)$  again with the backward Euler by solving

$$\mathbf{M}_l \chi_{\mathbf{p}_0,l} = \mathbf{rhs}_l \quad (4.138)$$

with

$$\begin{aligned}
\mathbf{M}_l &:= \frac{1}{h} \mathbf{C}_x(t_l) + \mathbf{G}_x(t_l) \approx \frac{\partial \mathbf{q}}{\partial \mathbf{x}}(\mathbf{x}_l, \mathbf{p}_0, \mathbf{p}_0) + \frac{\partial \mathbf{j}}{\partial \mathbf{x}}(\mathbf{x}_l, \mathbf{p}_0, \mathbf{p}_0), \\
\mathbf{rhs}_l &:= \mathbf{S}_p(t_l) - \left( \frac{d}{dt} \mathbf{C}_p(t_l) + \mathbf{G}_p(t_l) \right) + \frac{1}{h} \mathbf{C}_x(t_{l-1}) \cdot \chi_{\mathbf{p}_0,l-1} \\
&\approx \mathbf{S}_p(t_l) - \left( \frac{1}{h} \left( \frac{\partial \mathbf{q}}{\partial \mathbf{p}}(\mathbf{x}_l, \mathbf{p}_0, \mathbf{p}_0) - \frac{\partial \mathbf{q}}{\partial \mathbf{p}}(\mathbf{x}_{l-1}, \mathbf{p}_0, \mathbf{p}_0) \right) \right. \\
&\quad \left. + \frac{\partial \mathbf{j}}{\partial \mathbf{p}}(\mathbf{x}_l, \mathbf{p}_0, \mathbf{p}_0) \right) - \frac{1}{h} \frac{\partial \mathbf{q}}{\partial \mathbf{x}}(\mathbf{x}_{l-1}, \mathbf{p}_0, \mathbf{p}_0) \cdot \chi_{\mathbf{p}_0,l-1}.
\end{aligned} \quad (4.139)$$

We note, that the partial derivatives with respect to  $\mathbf{x}$  have already been computed in the transient analysis and are available, if they have been stored. Especially, the system matrix  $\mathbf{M}_l$  is the same as we have used in applying the backward Euler method in the underlying simulation: within the Newton iterations these were the system matrices in the steps where convergence was recorded. Hence, also the decomposition of this matrix is available, such that the system could be solved efficiently. For schemes other than the backward Euler, we can also state, that the solution of the transient sensitivity problem (4.136) needs ingredients that are available (if they have been stored) from the solution of the network problem with the same method and the same step size. A reasoning for this and details on step size control and error estimation can be found in the paper by Daldoss et al. [78].

However, the sensitivities of the element functions  $\mathbf{q}, \mathbf{j}$  and  $\mathbf{s}$  have to be calculated.

In total, the evaluation of the right-hand side  $\mathbf{rhs}_l$  requires  $\mathcal{O}(n_p \cdot n^2) + \mathcal{O}(n_p \cdot n)$  operations [96]. As in addition a lot of data has to be stored, computing the state sensitivities for circuits containing a large number of parameters is not tractable.

#### 4.3.2.2 Observation Function Sensitivity

Often one is not interested in the sensitivity  $\chi_{\mathbf{p}_0}(t)$  of the states of the parameter dependent network problem (4.132) but rather in the sensitivity of some performance figures of the system, like e. g., power consumption. These measures can usually be described by some *observation function*  $\Gamma(\mathbf{x}, \mathbf{p}) \in \mathbb{R}^{n_o}$  of the form

$$\Gamma(\mathbf{x}, \mathbf{p}) = \int_{t_0}^{t_{\text{end}}} \mathbf{g}(\mathbf{x}, t, \mathbf{p}) dt, \quad (4.140)$$

where the function  $\mathbf{g} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_o}$  is such that the partial derivatives  $\partial \mathbf{g} / \partial \mathbf{x}$  and  $\partial \mathbf{g} / \partial \mathbf{p}$  exist and are bounded. Note that at the left-hand side of (4.140)  $\mathbf{x} = \mathbf{x}(\cdot, \mathbf{p})$ , which is a whole waveform in time.

The sensitivity of the observation function  $\Gamma : \mathbb{R}^n \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_o}$  around some nominal parameterset  $\mathbf{p}_0 \in \mathbb{R}^{n_p}$ , clearly is

$$\frac{d\Gamma}{d\mathbf{p}}(\mathbf{x}(\mathbf{p}_0), \mathbf{p}_0) = \frac{\partial \Gamma}{\partial \mathbf{x}}(\mathbf{x}(\mathbf{p}_0), \mathbf{p}_0) \frac{\partial \mathbf{x}}{\partial \mathbf{p}}(\mathbf{x}(\mathbf{p}_0), \mathbf{p}_0) + \frac{\partial \Gamma}{\partial \mathbf{p}}(\mathbf{x}(\mathbf{p}_0), \mathbf{p}_0) \in \mathbb{R}^{n_o \times n_p}. \quad (4.141)$$

For problems where the sensitivity of a few observables, i. e., where  $n_o$  is small but the system depends on a large number  $n_p$  of parameters, the *adjoint method*, introduced by Cao et al. in [73, 74] is an attractive approach. In the mentioned papers, the observation sensitivity problem is derived for implicit differential equations of the form

$$\mathbf{F}(\mathbf{x}, \dot{\mathbf{x}}, t, \mathbf{p}) = \mathbf{0}. \quad (4.142)$$

Here we derive the observation sensitivity problem for problem (4.132a) as we usually encounter in circuit simulation. The idea however, follows the idea presented by Cao et al. in the papers mentioned.

The observation function's sensitivity is not calculated directly. Instead, an intermediate quantity  $\lambda$ , defined by a dynamical system, the *adjoint model* [82] of the parent problem, is calculated.

#### 4.3.2.3 Adjoint System for Sensitivity Analysis

Instead of considering the definition (4.140) of the observation function  $\Gamma$  directly, we define an augmented observation function

$$Y(\mathbf{x}, \mathbf{p}) := \Gamma(\mathbf{x}, \mathbf{p}) - \int_{t_0}^{t_{\text{end}}} \lambda^T(t) \left[ \frac{d}{dt} \mathbf{q}(\mathbf{x}(t, \mathbf{p}), \mathbf{p}) + \mathbf{j}(\mathbf{x}(t, \mathbf{p}), \mathbf{p}) - \mathbf{s}(t, \mathbf{p}) \right] dt, \quad (4.143)$$

which arises from coupling the dynamics and the observation function  $\Gamma$  by a Lagrangian multiplier  $\lambda(t) \in \mathbb{R}^{n \times n_o}$  that we will define more precisely furtheron. If  $\mathbf{x}(t, \mathbf{p})$  solves the network equations (4.132a) for  $\mathbf{p} = \mathbf{p}_0$  it holds  $Y(\mathbf{x}, \mathbf{p}_0) = \Gamma(\mathbf{x}, \mathbf{p}_0)$  and also the sensitivities coincide:

$$\frac{d\Gamma}{d\mathbf{p}}(\mathbf{x}, \mathbf{p}_0) = \frac{dY}{d\mathbf{p}}(\mathbf{x}, \mathbf{p}_0).$$

Note, that where it is clear from the context we omit in the following the specifications of the evaluation points, e. g.,  $(\mathbf{x}, \mathbf{p}_0)$ .

By the definitions (4.140) and (4.143) of the observation function and the augmented observation function, respectively, we get

$$\frac{d\Gamma}{d\mathbf{p}} = \int_{t_0}^{t_{\text{end}}} \left( \frac{\partial \mathbf{g}}{\partial \mathbf{p}} + \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{p}} \right) dt - \int_{t_0}^{t_{\text{end}}} \lambda^T(t) \left( \frac{d}{dt} \frac{d\mathbf{q}}{d\mathbf{p}} + \frac{d\mathbf{j}}{d\mathbf{p}} - \frac{d\mathbf{s}}{d\mathbf{p}} \right) dt. \quad (4.144)$$

We have a closer look at the second integral and apply integration by parts:

$$\int_{t_0}^{t_{\text{end}}} \lambda^T(t) \left( \frac{d}{dt} \frac{d\mathbf{q}}{d\mathbf{p}} \right) dt = \left[ \lambda^T \frac{d\mathbf{q}}{d\mathbf{p}} \right]_{t_0}^{t_{\text{end}}} - \int_{t_0}^{t_{\text{end}}} \frac{d\lambda^T}{dt} \frac{d\mathbf{q}}{d\mathbf{p}} dt.$$

Recombining this with the observation sensitivity (4.144) and expanding the total derivatives with respect to  $\mathbf{p}$  we see

$$\begin{aligned} \frac{d\Gamma}{d\mathbf{p}} = & - \left[ \lambda^T(t) \left( \mathbf{C}_x(t) \chi_{\mathbf{p}_0}(t) + \mathbf{C}_p(t) \right) \right]_{t_0}^{t_{\text{end}}} \\ & + \int_{t_0}^{t_{\text{end}}} \left( \frac{d\lambda^T}{dt} \mathbf{C}_x(t) - \lambda^T \mathbf{G}_x(t) + \gamma_x(t) \right) \cdot \chi_{\mathbf{p}_0}(t) dt \\ & + \int_{t_0}^{t_{\text{end}}} \left( \gamma_p(t) + \frac{d\lambda^T}{dt} \mathbf{C}_p(t) - \lambda^T [\mathbf{G}_p - \mathbf{S}_p(t)] \right) dt, \end{aligned} \quad (4.145)$$

where  $\mathbf{C}_x, \mathbf{G}_x, \mathbf{C}_p, \mathbf{G}_p$  and  $\mathbf{S}_p$  are the quantities defined in Eq. (4.136) and

$$\gamma_x(t) := \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}(t, \mathbf{p}_0)) \in \mathbb{R}^{n_o \times n}, \quad \gamma_p(t) := \frac{\partial \mathbf{g}}{\partial \mathbf{p}}(\mathbf{x}(t, \mathbf{p}_0)) \in \mathbb{R}^{n_o \times n_p}$$

are the partial derivatives of the kernel of the observable and can thus be computed, if the solution trajectory  $\mathbf{x}(t, \mathbf{p}_0)$  is known.

In the present form (4.145) the calculation of the observation function's sensitivity still demands to know the development of the state sensitivities  $\chi_{\mathbf{p}_0}(t)$ . As the above considerations are valid for any smooth  $\lambda(t) \in \mathbb{R}^{n \times n_o}$  we may choose this parameter

such that the state sensitivity disappears in the equation. We have already seen that the sensitivity  $\chi_{\mathbf{p}_0}^{\text{DC}}$  of the circuit's operating point  $\mathbf{x}_{\mathbf{p}_0}^{\text{DC}}$  can easily be calculated. Hence, choosing  $\lambda$  such that

$$\mathbf{C}_x^T(t) \frac{d\lambda}{dt} - \mathbf{G}_x(t)^T \lambda = -\gamma_x^T(t) \quad (4.146a)$$

$$\text{and } \lambda^T(T) \mathbf{C}_x(T) = \mathbf{0}, \quad (4.146b)$$

the calculation of the observable function's sensitivity reduces to evaluating

$$\begin{aligned} \frac{d\Gamma}{d\mathbf{p}} = & \lambda^T(t_0) \left( \mathbf{C}_x(t_0) \chi_{\mathbf{p}_0}^{\text{DC}} + \mathbf{C}_p(t_0) \right) - \lambda^T(t_{\text{end}}) \mathbf{C}_p(t_{\text{end}}) \\ & + \int_{t_0}^{t_{\text{end}}} \left( \gamma_p(t) + \frac{d\lambda^T}{dt} \mathbf{C}_p(t) - \lambda^T [\mathbf{G}_p - \mathbf{S}_p(t)] \right) dt. \end{aligned} \quad (4.147)$$

Equation (4.146a) inherits the basic structure of the underlying network problem (4.132a). Therefore, this equation defining the Lagrangian  $\lambda(t)$ , is usually a DAE system. This linear system is called the *adjoint system* to the underlying network equation. For DAEs of index up to 1, the choice (4.146b) defines a consistent initial value [74]. For systems of index larger and equal to 2, the consistent initialisation is more difficult. As an initial value for  $\lambda$  is specified for the end of the interval  $[t_0, t_{\text{end}}]$  of interest, the adjoint equation (4.146a) is solved backwards in time.

Several kind of observation functions also need  $\dot{\mathbf{x}} = \frac{d}{dt} \mathbf{x}$  in addition to  $\mathbf{x}$ . For instance when considering jitter one is interesting in the time difference between two subsequent times  $\tau_1$  and  $\tau_2$  when a specific unknown reaches or crosses a given value  $c$  (with equal signs of the time derivative). For the frequency of the jitter we have  $f = 1/T = 1/(\tau_2 - \tau_1)$ . Let the specific unknown be  $x_i(t, p)$ . The time moment  $\tau$  for which  $x_i(\tau, p) = c$  may be determined by inverse interpolation between two time points  $t_1$  and  $t_2$  and known values  $x_i$  obtained by time integration such that  $x_i(t_1, p) < c < x_i(t_2, p)$ . Of course  $\tau$  depends on  $p$ , so more precisely we have  $x_i(\tau(p), p) = c$ . By differentiation we obtain:  $d(\tau)/dp = -[d(x_i)/dt]^{-1} d(x_i)/dp$ .

Hence we are also interested in a more general case than (4.140)

$$\mathbf{H}(\dot{\mathbf{x}}(\mathbf{p}), \mathbf{x}(\mathbf{p}), \mathbf{p}) = \int_{t_0}^{t_{\text{end}}} \mathbf{F}(\dot{\mathbf{x}}(t, \mathbf{p}), \mathbf{x}(t, \mathbf{p}), \mathbf{p}) dt. \quad (4.148)$$

By a similar analysis as presented in [96] for (4.144)-(4.145) we derive ( $\hat{\mathbf{x}} = \partial \dot{\mathbf{x}} / \partial \mathbf{p}$ )

$$\begin{aligned}
\frac{d}{d\mathbf{p}}\mathbf{H}(\hat{\mathbf{x}}(\mathbf{p}), \mathbf{x}(\mathbf{p}), \mathbf{p}) &= \int_{t_0}^{t_{\text{end}}} \left( \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \cdot \hat{\mathbf{x}} + \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \cdot \hat{\mathbf{x}} + \frac{\partial \mathbf{F}}{\partial \mathbf{p}} \right) dt \\
&= - \left( \zeta^T(t_{\text{end}}) \mathbf{C}_x(t_{\text{end}}) - \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}}(t_{\text{end}}) \right) \chi_{\mathbf{p}_0}(t_{\text{end}}) - \zeta^T(t_{\text{end}}) \mathbf{C}_p(t_{\text{end}}) \\
&\quad + \left( \zeta^T(t_0) \mathbf{C}_x(t_0) - \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}}(t_0) \right) \chi_{\mathbf{p}_0}(t_0) + \zeta^T(t_0) \mathbf{C}_p(t_0) \\
&\quad + \int_{t_0}^{t_{\text{end}}} \left( \left[ \frac{d\zeta^T}{dt} \mathbf{C}_x - \zeta^T \mathbf{G}_x - \frac{d}{dt} \left( \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \right) + \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right] \cdot \chi_{\mathbf{p}_0} \right. \\
&\quad \left. + \frac{d\zeta^T}{dt} \mathbf{C}_p - \zeta^T (\mathbf{G}_p - \mathbf{S}_p) + \frac{\partial \mathbf{F}}{\partial \mathbf{p}} \right) dt. \tag{4.149}
\end{aligned}$$

which holds for any  $\zeta(t) \in \mathbb{R}^{n \times n_o}$ . If  $\zeta$  is chosen such that

$$\mathbf{C}_x^T \frac{d\zeta}{dt} - \mathbf{G}_x^T \zeta = \frac{d}{dt} \left( \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \right)^T - \left( \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right)^T, \tag{4.150}$$

$$\text{with 'initial' value } \mathbf{C}_x^T \zeta(t_{\text{end}}) = \left( \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \right)^T(t_{\text{end}}), \tag{4.151}$$

a significant reduction occurs in (4.149) and  $\hat{x}(t)$  is not explicitly needed. This generalizes the result in [96] (see also [64]). Note that  $\chi_{\mathbf{p}_0}(0) = \hat{\mathbf{x}}(0, \mathbf{p}_0) = \hat{\mathbf{x}}_{\text{DC}}(\mathbf{p}_0)$ , which is the sensitivity of the DC-solution, which one needs to determine explicitly. Some efficiency is gained by calculating  $\zeta^T(0) \mathbf{C}_x(0) \hat{\mathbf{x}}_{\text{DC}} = [\mathbf{C}_x^T(0) \zeta(0)]^T \text{state}_{\text{DC}}$  (when  $n_p \gg 1$ ). Note however that (4.151) can be satisfied only when the right-hand side is in the range of  $\mathbf{C}_x^T$ . Because in (4.150)-(4.151) the right-hand sides are evaluated at  $\mathbf{x}(t, \mathbf{p})$ , in general, the solution  $\zeta$  will depend on  $\mathbf{p}$ , even in the case of constant matrices  $\mathbf{C}_x$  and  $\mathbf{G}_x$ . This is in contrast to [64].

Summing up, the *backward adjoint method* for computing the sensitivity of the observable  $\Gamma$  with respect to parameter variations around a nominal parameter setting  $\mathbf{p}_0$  is carried out by the following steps

1. solve the network DAE (4.132a) for  $\mathbf{x} = \mathbf{x}(t, \mathbf{p}_0)$ , on the interval  $[t_0, t_{\text{end}}]$ ;
2. solve the backward adjoint problem (4.146a), subject to the initial condition (4.146b) for  $\lambda$  on the interval  $[t_{\text{end}}, t_0]$ , i. e., backward in time;
3. compute the observable sensitivity  $d\Gamma / d\mathbf{p}$  using the expression (4.147).

Carrying out the backward adjoint method, one has to consider several aspects we do not address here. Amongst these are the evaluations of the partial derivatives like  $\mathbf{C}_x$  along the solution trajectory. On the one hand, these derivatives are usually not available as a closed function but are approximated by finite differences. On the other hand, the evaluation points, i. e., points on the trajectory  $\mathbf{x}(t, \mathbf{p}_0)$  are also available as approximations only. Furthermore, the integral in the formula (4.147) has to be approximated by a numerical quadrature. The nodes needed in the according scheme may not be met exactly during transient simulation and/or during the backward integration of the adjoint problem. For further reading on these problems we refer to [95, 96].

However, leaving all these aspects aside, one has to integrate two dynamical systems numerically. First the (nonlinear) forward problem (4.132a) for the states and then

the linear backward problem (4.146a). The contribution of the COMSON project for transient sensitivity analysis was to add model order reduction (MOR) to the process. More precisely, the idea elaborated during the project was to solve an order reduced variant of the backward problem where the data needed to apply the reduction is calculated from the forward solving phase. In the next section we will describe the very basic idea of MOR and give a brief introduction to the specific technique that was used in this project.

### 4.3.3 Model Order Reduction (with POD)

Solving a dynamical system with any numerical scheme implies to set up and solve a series of linear equations. In circuit simulation typically the dimension of these systems are in a range of  $10^5$  to  $10^9$ . Both the evaluation of the system matrices and right-hand sides, e. g.,  $\mathbf{M}_l$  and  $\mathbf{rhs}_l$  in (4.138)-(4.139), as well as solving the system, i. e., decomposing the system matrices, is computationally costly. However, in circuit design often the main interest is the analysis of how a circuit block processes an input signal, e. g., if some input signal is amplified or damped by the circuit. That means, one may not be interested in all  $n$  internal state variables but only in a limited selection. This concern is described by an input-output variant of the network model. For a linear network problem (4.132b), omitting the parameters for ease of notation, e. g., the corresponding input-output system reads

$$\begin{aligned}\mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{G}\mathbf{x}(t) &= \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{L}\mathbf{x}(t),\end{aligned}\tag{4.152}$$

where  $\mathbf{u}(t) \in \mathbb{R}^m$  and  $\mathbf{y}(t) \in \mathbb{R}^q$  are the input and the output of the system, injected to and extracted from the system by the matrices  $\mathbf{B} \in \mathbb{R}^{n \times m}$  and  $\mathbf{L} \in \mathbb{R}^{q \times n}$ .

As in an input-output setting, the states  $\mathbf{x}$  represent an auxiliary variable only. The idea of MOR is to replace the high-dimensional dynamical system (4.152) by

$$\begin{aligned}\widehat{\mathbf{C}}\dot{\mathbf{z}}(t) + \widehat{\mathbf{G}}\mathbf{z}(t) &= \widehat{\mathbf{B}}\mathbf{u}(t), \\ \tilde{\mathbf{y}}(t) &= \widehat{\mathbf{L}}\mathbf{z}(t),\end{aligned}\tag{4.153}$$

where  $\mathbf{z}(t) \in \mathbb{R}^r$  and the system matrices  $\widehat{\mathbf{C}}, \widehat{\mathbf{G}} \in \mathbb{R}^{r \times r}$ ,  $\widehat{\mathbf{B}} \in \mathbb{R}^{r \times m}$  and  $\widehat{\mathbf{L}} \in \mathbb{R}^{q \times r}$  are chosen such that  $r \ll n$  and  $\tilde{\mathbf{y}}(t) \approx \mathbf{y}(t)$ .

There are various methods to construct the reduced variant (4.153) from the full problem (4.152). We refer to Chapter 3 for an overview, as well as to [63, 67, 68, 70, 104, 106–108, 117] for further studies.

A large class of MOR methods are based on projection. These methods determine a subspace of dimension  $r$ , spanned by a basis of vectors  $\mathbf{v}_i \in \mathbb{R}^n$  ( $i = 1, \dots, r$ ). The original state vector  $\mathbf{x}(t)$  is approximated by an element of this subspace that can be written in the form  $\mathbf{V}\mathbf{z}(t)$ , where  $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_r) \in \mathbb{R}^{n \times r}$ . Hence, one replaces  $\mathbf{x}(t)$  by  $\mathbf{V}\mathbf{z}(t)$  in (4.152) and projects the equation onto the space subspaces spanned by

the columns of  $\mathbf{V}$  by a Galerkin approach. In this way, a dynamical system (4.153) emerges where the system matrices are given by

$$\widehat{\mathbf{C}} := \mathbf{V}^T \mathbf{C} \mathbf{V}, \quad \widehat{\mathbf{G}} := \mathbf{V}^T \mathbf{G} \mathbf{V}, \quad \widehat{\mathbf{B}} := \mathbf{V}^T \mathbf{B}, \quad \widehat{\mathbf{L}} := \mathbf{L} \mathbf{V}. \quad (4.154)$$

### 4.3.3.1 Proper Orthogonal Decomposition

While other MOR methods start operating from the matrices  $\mathbf{C}$ ,  $\mathbf{G}$ ,  $\mathbf{B}$  and  $\mathbf{L}$ , the method of Proper Orthogonal Decomposition (POD) constructs the matrix  $\mathbf{V}$ , whose columns span the reduced space the system (4.152) is projected on, from the space that is spanned by the trajectory  $\mathbf{x}(t)$ , i. e., the solution of the dynamical system. The method applies to nonlinear systems as well.

Recall, that our aim is to construct a reduced model for the backward adjoint problem (4.146). As this is a linear system from which we know the system matrices only after a solution of the underlying forward network problem (4.132a), POD seems to be the best choice for this task.

The mission POD fulfills is to find a subspace approximating a given set of data in an optimal least-squares sense. The basis of this approach is known also as *Principal Component Analysis* and *Karhunen-Loève theorem* from picture and data analysis. The mathematical formulation of POD [104, 105, 117] is as follows: Given a set of  $K$  datapoints  $\mathbf{X} := \{\mathbf{x}_1, \dots, \mathbf{x}_K\}$ , a subspace  $S \subset \mathbb{R}^n$  is searched for that minimizes

$$\|\mathbf{X} - \rho \mathbf{X}\|_2^2 := \frac{1}{K} \sum_{k=1}^K \|\mathbf{x}_k - \rho \mathbf{x}_k\|_2^2, \quad (4.155)$$

where  $\rho : \mathbb{R}^n \rightarrow S$  is the orthogonal projection onto  $S$ , which has  $\{\varphi_1, \dots, \varphi_r\}$  as an orthonormal basis of  $S$ .

This problem is solved, applying the Singular Value Decomposition (SVD) to the matrix  $\mathbf{X} := (\mathbf{x}_1, \dots, \mathbf{x}_K) \in \mathbb{R}^{n \times K}$ , which is called *snapshot matrix*, as its columns are (approximations to) the solution of the dynamical system (4.152) at timepoints  $t_1, \dots, t_K \in [t_0, t_{\text{end}}]$ . The SVD applied to the matrix  $\mathbf{X}$ , provides three matrices:

$$\begin{aligned} \Phi &\in \mathbb{R}^{n \times n} \quad \text{orthogonal,} \\ \Psi &\in \mathbb{R}^{K \times K} \quad \text{orthogonal,} \\ \Sigma &= \text{diag}(\sigma_1, \dots, \sigma_v) \in \mathbb{R}^{v \times v} \quad \text{with } \sigma_1 \geq \dots \geq \sigma_v > \sigma_{v+1} = \dots = \sigma_K = 0, \end{aligned}$$

such that

$$\mathbf{X} = \Phi \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \Psi^T, \quad (4.156)$$

where the columns of  $\Phi$  and  $\Psi$  are left and right eigenvectors, respectively, and  $\sigma_1, \dots, \sigma_v$  are the singular values of  $\mathbf{X}$ .

Then, for any  $r \leq v$ , taking  $\varphi_1, \dots, \varphi_r$  as the first  $r$  columns of the matrix  $\Phi$  is



optimal in the sense that it minimizes the projection mismatch (4.155).

Both cases,  $n \geq K$  and  $n \leq K$ , are allowed; in practise one often has  $n \gg K$ .

Finally, the MOR projection matrix  $\mathbf{V}$  in (4.154) is chosen made up of these basis vectors:

$$\mathbf{V} := (\varphi_1, \dots, \varphi_r) \in \mathbb{R}^{n \times r}.$$

To understand why the first  $r$  columns of  $\Phi$  solve the minimization problem (4.155) one can recall that for  $i = 1, \dots, n$  the  $i$ th column  $\varphi_i$  of  $\Phi$  is actually an eigenvector of the correlation (or covariance) matrix of the snapshots with  $\sigma_i^2$  as eigenvalue:

$$\mathbf{X}\mathbf{X}^T \varphi_i = \sigma_i^2 \varphi_i.$$

Intuitively the correlation matrix  $\mathbf{X}\mathbf{X}^T$  detects the principal directions in the data cloud that is made up of the snapshots  $\mathbf{x}_1, \dots, \mathbf{x}_K$ . The eigenvectors and eigenvalues can be thought of as directions and radii of axes of an ellipsoid that incloses the cloud of data. Then, the smaller the radii of one axis is, the less information is lost if that direction is neglected.

We abandon to explain the derivation of POD in detail here as in literature e. g., [63, 104, 117] this is well explained. For details on the accuracy of MOR with POD we refer to papers by Petzold et al [94, 105].

#### 4.3.4 The BRAM algorithm

In [96] it was observed that a forward analysis in time of (4.132a) automatically provides snapshots  $\mathbf{x}(t_i, \mathbf{p})$  at time points  $t_i$ . This can lead to a reduced system of equations for  $\zeta(t) = \mathbf{V}\tilde{\zeta}$  in (4.150)-(4.151)

$$\mathbf{V}^T \mathbf{C}_x^T \mathbf{V} \frac{d\tilde{\zeta}}{dt} - \mathbf{V}^T \mathbf{G}_x^T \mathbf{V} \tilde{\zeta} = \mathbf{V}^T \frac{d}{dt} \left( \frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{x}}} \right)^T - \mathbf{V}^T \left( \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right)^T, \quad (4.157)$$

$$\text{with 'initial' value } \mathbf{V}^T \mathbf{C}_x^T \mathbf{V} \tilde{\zeta}(t_{\text{end}}) = \mathbf{V}^T \left( \frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{x}}} \right)^T(t_{\text{end}}), \quad (4.158)$$

Then the overall algorithm is described in Algorithm 4.1, without the lines 5-7. Here it is assumed that the matrices are saved after the forward simulation. It is also assumed that for the adjoint system the same step sizes are used as in the forward run. If not, additional interpolation has to be taken into account to determine the reduced matrices at intermediate solutions and also effort has to be spent in LU-decomposition.

Apart from this discussion, the question is why this should work in general (apart from special cases in [96]). The solution  $\mathbf{x} \approx \mathbf{V}\tilde{\mathbf{x}}$  depends on the right-hand side  $\mathbf{s}$  of (4.132a). Clearly  $\mathbf{V}^T \mathbf{s}$  should contain the dominant behaviour of  $\mathbf{s}$ . If  $\mathbf{V}^T \left[ \frac{d}{dt} \left( \frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{x}}} \right)^T - \left( \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right)^T \right]$  behaves similarly when compared to the right-hand side of (4.150) we may

**Algorithm 4.1** BRAM: Backward Reduced Adjoint Method

- 
- 1: Integrate (4.132a) and store the solutions  $\mathbf{x}(t_i, \mathbf{p})$
  - 2: Build the snapshot matrix  $\mathbf{X} = [\mathbf{x}(t_0, \mathbf{p}), \dots, \mathbf{x}(t_N, \mathbf{p})]$  (where  $t_N = t_{\text{end}}$ )
  - 3: Determine the singular value decomposition  $\mathbf{X} = \Phi^T \Sigma \Psi$  and dominant singular values  $\sigma_1, \dots, \sigma_r$ .
  - 4: Determine the Proper Orthogonal Decomposition (POD) time-independent projection matrix  $\mathbf{V}$ , such that  $\mathbf{x} \approx \mathbf{V}\tilde{\mathbf{x}}$  and  $\frac{d}{dt}\mathbf{x} \approx \mathbf{V}\frac{d}{dt}\tilde{\mathbf{x}}$
  - 5: **if** (BRAM II) **then**
  - 6:   Include a second forward time integration, now for the reduced system of equations.
  - 7: **end if**
  - 8: Integrate (4.157) backward in time using reduced matrices  $\mathbf{V}^* \mathbf{C}_x^T \mathbf{V}$  and  $\mathbf{V}^T \mathbf{G}_x^T \mathbf{V}$  and the projected right-hand side  $\mathbf{V}^T \left[ \frac{d}{dt} \left( \frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{x}}} \right)^T - \left( \frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{x}}} \right)^T \right]$
- 

expect a similar good approximation for the solution  $\zeta \approx \mathbf{V}\tilde{\zeta}$ . Because the right-hand side of (4.157) does not depend on  $\zeta$  this can be checked in advance, before solving (4.157). In the case of power loss through a resistor we have  $\left( \frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{x}}} \right)^T = (\mathbf{A}\mathbf{x})^T$  (for some matrix  $\mathbf{A}$ ) and we have to check if  $(\mathbf{A}\mathbf{x})^T \approx \mathbf{V}^T \tilde{\mathbf{x}}^T \mathbf{V}^T \mathbf{A}^T$ . Another point of attention is that the projection matrix  $\mathbf{V}$  found implies that we more or less are looking to the sensitivity of the solution  $\tilde{\mathbf{x}}$  of

$$\frac{d}{dt} [\mathbf{V}^T \mathbf{q}(\mathbf{V}\tilde{\mathbf{x}}(t, \mathbf{p}), \mathbf{p})] + \mathbf{V}^T \mathbf{j}(\mathbf{V}\tilde{\mathbf{x}}(t, \mathbf{p}), \mathbf{p}) = \mathbf{V}^T \mathbf{s}(t, \mathbf{p}) \quad (4.159)$$

rather than for the solution  $x$  of (4.132a). By this it is clear that  $\mathbf{V}$  depends on  $\mathbf{p}$  and thus

$$\mathbf{x}(t, \mathbf{p}) \approx \mathbf{V}(\mathbf{p})\tilde{\mathbf{x}}(t, \mathbf{p}) \implies \frac{\partial \mathbf{x}}{\partial \mathbf{p}} \approx \frac{\partial \mathbf{V}}{\partial \mathbf{p}} \tilde{\mathbf{x}} + \mathbf{V} \frac{\partial \tilde{\mathbf{x}}}{\partial \mathbf{p}}. \quad (4.160)$$

The question is: can we ignore the first term at the right-hand side of (4.160). Here the last term represents the change inside the space defined by the span of the columns of  $\mathbf{V}$ . The first term represents the effect by the change of this space itself. One may expect that this term is smaller than the last term ('the first term will in general require more energy'), especially when the reduction is more or less determined by topology. In several tests we made, this first term indeed was much smaller than the other term.

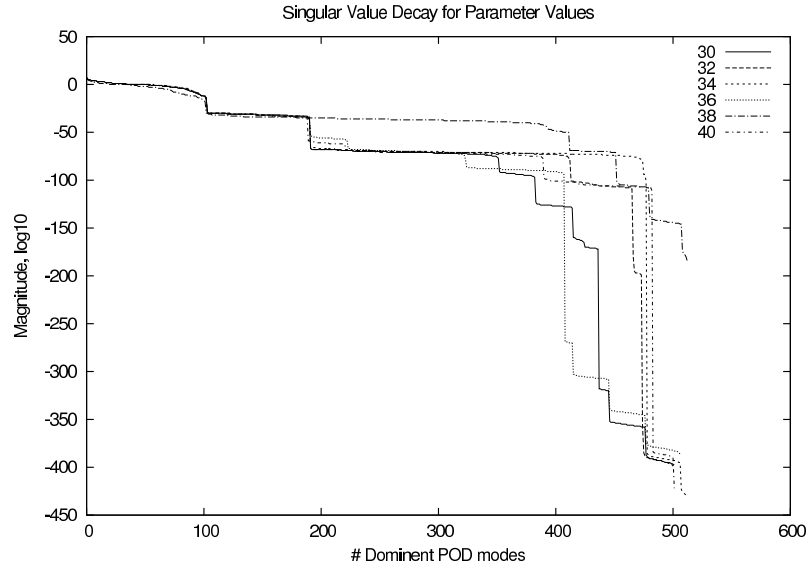
Note that we not intend to solve (4.159) by using a fixed projection matrix  $\mathbf{V}$ , valid for  $\mathbf{p} = \mathbf{p}_0$ , for several different values of  $\mathbf{p}$ . The danger of obtaining improper results when doing this was pointed out by [83]. Contrarily, we always apply an up-to-date matrix  $\mathbf{V}(\mathbf{p})$ . However, this example shows that  $\frac{\partial \mathbf{V}}{\partial \mathbf{p}}$  is not always negligible.

One can collect  $\tilde{\mathbf{V}} = [\mathbf{V}(\mathbf{p}_1), \dots, \mathbf{V}(\mathbf{p}_k)]$  and apply an additional SVD to  $\tilde{\mathbf{V}}$ . This procedure provides a larger, uniform, projection matrix  $\mathbf{V}$ .

In [95] the parameter dependency of the singular values for POD was analysed for a battery charger, for a ring oscillator, and for a car transceiver example. Also the nr of dominant singular values as function of  $\mathbf{p}$  was studied. Finally the angle between

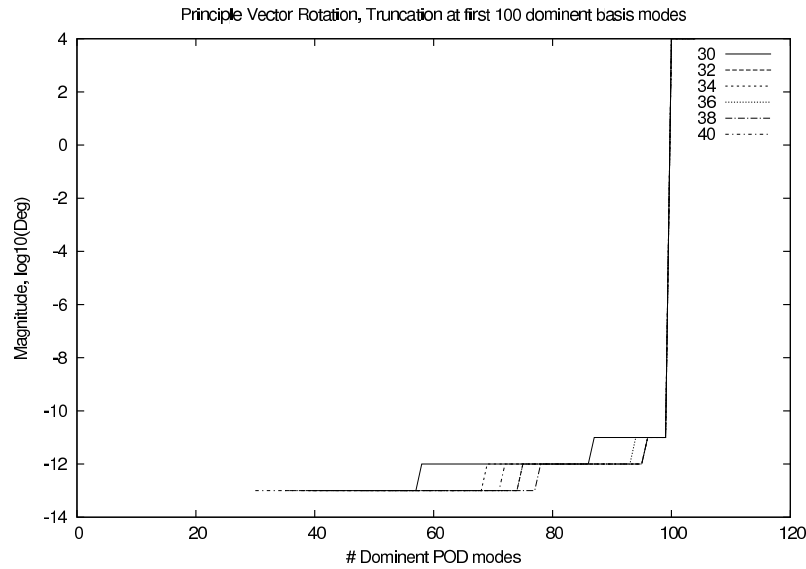
the subspaces for different  $\mathbf{p}$  was studied. Note that one can use a matlab function for this based on the algorithm by Knyazev-Argentati [98].

Finally, in [95] a modification was introduced in Algorithm 4.1 by introducing the lines 5-7. Note that the additional step 6 is cheap. We obtain the solution of the POD-reduced system. In [94, 104, 105, 117] error estimates are determined for the approximation error of the POD approximation. Actually, in step 8, BRAM II determines the sensitivity of the POD solution. In Fig. 4.35 [95] the singular values of



**Fig. 4.35** Singular values of POD after 3500 snapshots for a Li-ion charger for different values of the area of a capacitor.

POD after 3500 snapshots within a simulation from  $t_0 = 0\text{ms}$  and  $t_{\text{end}} = 200\text{ms}$  for a Li-ion charger for different values of the area of a capacitor. The parameter  $p$  took values  $p = 30, 32, 34, 36, 38, 40$ . Clearly the first 100 singular values are enough for a good reconstruction, which as a by-product also shows a high potential for the application of the BRAM methods as the dimension of the problem can be reduced by roughly a factor 35. In Fig. 4.36 [95] the angle in the rotation of the principle vector is studied, the nominal being for  $p = 30$ . The apparent jump to 90 degrees rotation near the cut off point is due to matrix diagonal zero padding introduced in the general case for principle vector analysis. These large 90 degree rotations are not due to principle vectors influenced by parameter changes and should not be taken into account.



**Fig. 4.36** Principle vector rotation as a function of the capacitor area for the problem in Fig. 4.35.

### 4.3.5 Sensitivity by Uncertainty Quantification

A modern approach to Uncertainty Quantification is to expand a solution  $\mathbf{x}(t, \mathbf{p})$  in a series of orthogonal polynomials in which the  $p$  is argument of the (multidimensional) polynomials and the  $t$  appears in the coefficients. If the  $\mathbf{p}$  are subject to variations such a representation is called a generalized Polynomial Chaos (gPC) expansion. Having established the expansion, this provides facilities similar like a response surface model: fast and accurate statistics and sensitivity.

In this section we shortly summarize some basic items. We also point out how a strategy for parameterized Model Order Reduction (pMOR) fits here. This strategy contains a generalization of one of the pMOR algorithms described in Section 4.1 of this Chapter.

We will denote parameters by  $\mathbf{p} = (p_1, \dots, p_q)^T$  and assume a probability space given  $(\Omega, \mathcal{A}, \mathcal{P})$  with  $\mathcal{P} : \mathcal{A} \rightarrow \mathbb{R}$  (measure; in our case the range will be  $[0, 1]$ ) and  $\mathbf{p} : \Omega \rightarrow Q \subseteq \mathbb{R}^q, \omega \mapsto \mathbf{p}(\omega)$ . Here we will assume that the  $p_i$  are independent. For a function  $f : Q \rightarrow \mathbb{R}$ , the mean or expected value is defined by

$$\langle f \rangle = \int_{\Omega} f(\mathbf{p}(\omega)) d\mathcal{P}(\omega) = \int_Q f(\mathbf{p}) \rho(\mathbf{p}) d\mathbf{p}. \quad (4.161)$$

The specific probability distribution density is defined by the function  $\rho(\mathbf{p})$ . A bilinear form  $\langle f, g \rangle$  is defined by

Distribution	Polynomial	Weight function	Support range
Gaussian	Hermite $H_n(p)$	$e^{-\frac{p^2}{2}}$	$(-\infty, \infty)$
Uniform	Legendre $P_n(p)$	1	$[-1, 1]$
Beta	Jacobi $P_n^{\alpha, \beta}(p)$	$(1-p)^\alpha(1+p)^\beta$	$[-1, 1]$
Exponential	Laguerre $L_n(p)$	$e^{-p}$	$[0, \infty)$
Gamma	Generalized Laguerre $L_n^{(\alpha)}(p)$	$p^\alpha e^{-p}$	$[0, \infty)$

**Table 4.3** One-dimensional orthogonal polynomials related to well-known probability density functions.

$$\langle f, g \rangle = \int_Q f(\mathbf{p}) g(\mathbf{p}) \rho(\mathbf{p}) d\mathbf{p} = \langle f g \rangle. \quad (4.162)$$

The last form is convenient when products of more functions are involved. Similar definitions hold for vector- or matrix-valued functions  $\mathbf{f}: Q \rightarrow \mathbb{R}^{m \times n}$ .

We assume a complete orthonormal basis of polynomials  $(\phi_i)_{i \in \mathbb{N}}$ ,  $\phi_i: \mathbb{R}^q \rightarrow \mathbb{R}$ , given with  $\langle \phi_i, \phi_j \rangle = \delta_{ij}$  ( $i, j, \geq 0$ ). When  $q = 1$ ,  $\phi_i$  has degree  $i$ . To treat a uniform distribution (i.e., for studying effects caused by robust variations) one can use Legendre polynomials; for a Gaussian distribution one can use Hermite polynomials [100, 119, 120]. Some one-dimensional polynomials are mentioned in Table 4.3. A polynomial  $\phi_{\mathbf{i}}$  on  $\mathbb{R}^q$  can be defined from one-dimensional polynomials:  $\phi_{\mathbf{i}}(\mathbf{p}) = \prod_{d=1}^q \phi_{i_d}(p_d)$ . Actually  $\mathbf{i}$  orders a vector  $\mathbf{i} = (i_1, \dots, i_q)^T$ ; however we will simply write  $\phi_i$ , rather than  $\phi_{\mathbf{i}}$ . An example is given in (4.163), using Legendre polynomials. Note that, due to normalization,  $L_0(p) = 1/\sqrt{2}$ ,  $L_1(p) = \sqrt{3}/2 p$ ,  $L_2(p) = \frac{1}{2} \sqrt{\frac{5}{3}} (3p^2 - 1)$  - see also [87]. In [88] one finds algorithms how to efficiently generate orthogonal polynomials from a given weight function.

$$\begin{aligned} \phi_0(\mathbf{p}) &= L_0(p_1) L_0(p_2), \\ \phi_1(\mathbf{p}) &= L_1(p_1) L_0(p_2), \\ \phi_2(\mathbf{p}) &= L_0(p_1) L_1(p_2), \\ \phi_3(\mathbf{p}) &= L_2(p_1) L_0(p_2), \\ \phi_4(\mathbf{p}) &= L_1(p_1) L_1(p_2), \\ \phi_5(\mathbf{p}) &= L_0(p_1) L_2(p_2). \end{aligned} \quad (4.163)$$

We will denote a dynamical system by

$$\mathbf{F}(\mathbf{x}(t, \mathbf{p}), t, \mathbf{p}) = 0, \quad \text{for } t \in [t_0, t_1]. \quad (4.164)$$

Here  $\mathbf{F}$  may contain differential operators. The solution  $\mathbf{x} \in \mathbb{R}^n$  depends on  $t$  and on  $\mathbf{p}$ . In addition initial and boundary values are assumed. In general these may depend on  $\mathbf{p}$  as well.

A solution  $\mathbf{x}(t, \mathbf{p}) = (x_1(t, \mathbf{p}), \dots, x_n(t, \mathbf{p}))^T$  of the dynamical system becomes a random process. We assume that second moments  $\langle x_j^2(t, \mathbf{p}) \rangle$  are finite, for all  $t \in [t_0, t_1]$  and  $j = 1, \dots, n$ . We express  $\mathbf{x}(t, \mathbf{p})$  in a Polynomial Chaos expansion

$$\mathbf{x}(t, \mathbf{p}) = \sum_{i=0}^{\infty} \mathbf{v}_i(t) \phi_i(\mathbf{p}), \quad (4.165)$$

where the coefficient functions  $\mathbf{v}_i(t)$  are defined by

$$\mathbf{v}_i(t) = \langle \mathbf{x}(t, \mathbf{p}), \phi_i(\mathbf{p}) \rangle. \quad (4.166)$$

Continuity/smoothness follow from the solution  $\mathbf{x}(t, \mathbf{p})$  and similarly the construction of expected values and variances.

A finite approximation  $\mathbf{x}^m(t, \mathbf{p})$  to  $\mathbf{x}(t, \mathbf{p})$  is defined by

$$\mathbf{x}^m(t, \mathbf{p}) = \sum_{i=0}^m \mathbf{v}_i(t) \phi_i(\mathbf{p}). \quad (4.167)$$

For long time range integration  $m$  may have to be chosen larger than for short time ranges. Further below we will describe how the coefficient functions  $\mathbf{v}_i(t)$  can be efficiently approximated.

For functions  $\mathbf{x}(t, \mathbf{p})$  that depend smoothly on  $\mathbf{p}$  convergence rates for  $\|\mathbf{x}(t, \cdot) - \mathbf{x}^m(t, \cdot)\|$ , in the norm associated with (4.162), are known. For instance, for one-dimensional functions  $x(p)$  that depend on a scalar parameter  $p$  such that  $x^{(1)}, \dots, x^{(k)}$  are continuous (i.e., derivatives w.r.t.  $p$ ), one has

$$\|x(\cdot) - x_H^m(\cdot)\|_{L_\rho^2} \leq C \frac{1}{m^{k/2}} \|x^{(k)}(\cdot)\|_{L_\rho^2}, \quad (\text{Hermite expansion [65]}), \quad (4.168)$$

$$\|x(\cdot) - x_L^m(\cdot)\|_{L_\rho^2} \leq C \frac{1}{m^k} \sqrt{\sum_{i=0}^k \|x^{(i)}(\cdot)\|_{L_\rho^2}^2}, \quad (\text{Legendre expansion [120]}), \quad (4.169)$$

Here the  $L_\rho^2$ -norms include the weighting/density function  $\rho(\cdot)$ . Note that the upperbound in (4.169) actually involves a Sobolev-norm. In [72] one also finds upperbounds using seminorms (that involve less derivatives).

For more general distributions  $\rho(\cdot)$  convergence may not be true. For instance, polynomials in a lognormal variable are not dense in  $L_\rho^2$ . For convergence one needs to require that the probability measure is uniquely determined by its moments [81]. One at least needs that the expected value of each polynomial has to exist. This has a practical impact. The imperfections in a manufacturing process cause some variability in the components of an electronic circuit. To address the variability, corresponding parameters or functions are replaced by random variables or random fields for uncertainty quantification. However, the statistics of the parameters often do not obey traditional probability distributions like Gaussian, uniform, beta or others. In such a case one may have to construct probability distributions or probability density functions, respectively, which approximate the true statistics at a sufficient accuracy. Thereby, one has to match corresponding data obtained from measurements and observations of electronic devices. The resulting probability distribution functions should be continuous and all moments of the random variables should be finite such that a broad class of methods like, e.g., Polynomial Chaos, is applicable.

The integrals (4.166) can be computed by (quasi) Monte Carlo, or by multi-dimensional quadrature. We assume quadrature grid points  $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^K$  and quadrature weights  $w_k, 1 \leq k \leq K$ , such that

$$\langle \mathbf{x}(t, \mathbf{p}), \phi_i(\mathbf{p}) \rangle \approx \sum_{k=1}^K w_k \mathbf{x}(t, \mathbf{p}^k) \phi_i(\mathbf{p}^k). \quad (4.170)$$

We solve (4.164) for  $\mathbf{x}(t, \mathbf{p}^k), k = 1, \dots, K$  ( $K$  deterministic simulations). Here any suitable numerical solver for (4.164) can be used. In fact (4.170) is a (discrete) inner-product with weighting function  $w_K(\mathbf{p}) = \sum_{k=1}^K w_k \delta(\mathbf{p} - \mathbf{p}^k)$ . This approach is called Stochastic Collocation [100, 119, 120]. Afterwards we determine

$$\mathbf{v}_i(t) = \sum_{k=0}^K w_k \mathbf{x}(t, \mathbf{p}^k) \phi_i(\mathbf{p}^k), \quad \text{for each } i. \quad (4.171)$$

Here the Polynomial Chaos expansion is just a post-processing step.

Only for low dimensions  $q$ , tensor-product grids of Gaussian quadrature are used. Gaussian quadrature points are optimal for accuracy. In higher-dimensional cases ( $q > 1$ ) one prefers sparse grids [119, 120], like the Smolyak algorithm. Sparse grids may have options for refinement. Note that Gaussian points do not offer this refinement. Stroud-3 and Stroud-5 formulas [114] have become popular [118].

An alternative approach to Stochastic Collocation is provided by Stochastic Galerkin. After, inserting an expansion of the solution, in polynomials in  $\mathbf{p}$ , into the equations one orthogonally projects the residue of the equations to the subspace spanned by these polynomials. By this, one gets one big system of differential equations in which the  $\mathbf{v}_i$  are the unknowns [100, 119, 120]. In practise, Stochastic Collocation is much more easily combined with dedicated software for the simulation problem at hand than is the case with Stochastic Galerkin. Theoretically the last approach is more accurate. However, statistics obtained with Stochastic Collocation is very satisfactory.

We note that the expansion  $\mathbf{x}(t, \mathbf{p})$ , see (4.165), gives full detailed information when varying  $\mathbf{p}$ . From this the actual (and probably biased) range of solutions can be determined. These can be different from envelope approximations based on mean and variances.

Because of the orthogonality, the mean of  $\mathbf{x}(t, \mathbf{p})$  and of  $\mathbf{x}^m(t, \mathbf{p})$  are equal and are given by

$$\mathbf{E}_p[\mathbf{x}(t, \mathbf{p})] = \int_Q \mathbf{x}(t, \mathbf{p}) \rho(\mathbf{p}) d\mathbf{p} = \mathbf{v}_0(t) = \int_Q \mathbf{x}^m(t, \mathbf{p}) \rho(\mathbf{p}) d\mathbf{p}. \quad (4.172)$$

Using (4.171), we get an approximative value. The integrals in (4.172) involve all  $p_k$  together. One may want to consider effects of  $p_i$  and  $p_j$  separately. This restricts the parameter space  $\mathbb{R}^q$  to a one-dimensional subset with individual distribution densities  $\rho_i(p)$  and  $\rho_j(p)$ .

A covariance function of  $\mathbf{x}(t, \mathbf{p})$  can also be easily expressed

$$\begin{aligned}
R_{\mathbf{xx}}(t_1, t_2) &= \mathbf{E}_p[(\mathbf{x}(t_1, \mathbf{p}) - \mathbf{E}_p[\mathbf{x}(t_1, \mathbf{p})])^T (\mathbf{x}(t_2, \mathbf{p}) - \mathbf{E}_p[\mathbf{x}(t_2, \mathbf{p})])] \\
&= \int_Q (\mathbf{x}(t_1, \mathbf{p}) - \mathbf{E}_p[\mathbf{x}(t_1, \mathbf{p})])^T (\mathbf{x}(t_2, \mathbf{p}) - \mathbf{E}_p[\mathbf{x}(t_2, \mathbf{p})]) \rho(\mathbf{p}) d\mathbf{p} \\
&\approx \langle (\mathbf{x}^m(t_1, \mathbf{p}) - \mathbf{E}_p[\mathbf{x}^m(t_1, \mathbf{p})])^T (\mathbf{x}^m(t_2, \mathbf{p}) - \mathbf{E}_p[\mathbf{x}^m(t_2, \mathbf{p})]) \rangle \\
&= \langle (\sum_{i=1}^m \mathbf{v}_i^T(t_1) \phi_i(\mathbf{p})) (\sum_{j=1}^m \mathbf{v}_j(t_2) \phi_j(\mathbf{p})) \rangle \\
&= \sum_{i=1}^m \mathbf{v}_i^T(t_1) \mathbf{v}_i(t_2). \tag{4.173}
\end{aligned}$$

This outcome clearly depends on  $m$ . A (scalar) variance is given by

$$\text{Var}_p[\mathbf{x}(t, \mathbf{p})] = R_{\mathbf{xx}}(t, t) \approx \sum_{i=1}^m \mathbf{v}_i^T(t) \mathbf{v}_i(t) = \sum_{i=1}^m \|\mathbf{v}_i(t)\|^2 = \|\mathbf{V}_0(t)\|^2, \tag{4.174}$$

where  $\mathbf{V}_0^T(t) = (\mathbf{0}^T, \mathbf{v}_1^T(t), \dots, \mathbf{v}_m^T(t))^T$ . Note that this equals

$$\text{Var}_p[\mathbf{x}(t, \mathbf{p})] \approx \sum_{i=1}^m \sum_{d=1}^q v_{i,q}^2(t) = \sum_{d=1}^q \sum_{i=1}^m v_{i,q}^2(t) = \sum_{d=1}^q \text{Var}_p[x_d(t, \mathbf{p})]. \tag{4.175}$$

Having a gPC expansion the sensitivity (matrix) w.r.t.  $\mathbf{p}$  is easily obtained

$$\mathbf{S}_p(t, \mathbf{p}) = \left[ \frac{\partial \mathbf{x}(t, \mathbf{p})}{\partial \mathbf{p}} \right] \approx \sum_{i=0}^m \mathbf{v}_i(t) \frac{\partial \phi_i(\mathbf{p})}{\partial \mathbf{p}}. \tag{4.176}$$

One may restrict this to  $\mathbf{S}_p(t, \mu_p)$ , where  $\mu_p = \mathbf{E}[\mathbf{p}]$  and  $\frac{\partial \mathbf{x}(t, \mathbf{p})}{\partial \mathbf{p}}$  is the solution of the system that is differentiated w.r.t.  $\mathbf{p}$  at  $\mathbf{p} = \mu_p$ . For a scalar quantity  $x$  one can order according to a ‘stochastic influence’ based on

$$\max \left\{ \frac{\partial x}{\partial p_1} \sigma_{p_1}, \dots, \frac{\partial x}{\partial p_q} \sigma_{p_q} \right\}. \tag{4.177}$$

Here  $\sigma_{p_i}^2 = \text{Var}[p_i]$ . The sensitivity matrix also is subject to stochastic variations. With a gPC expansion one can determine a mean global sensitivity matrix by

$$\mathbf{S}_p(t) = \mathbf{E}_p \left[ \frac{\partial \mathbf{x}(t, \mathbf{p})}{\partial \mathbf{p}} \right] \approx \sum_{i=0}^m \mathbf{v}_i(t) \int_Q \frac{\partial \phi_i(\mathbf{p})}{\partial \mathbf{p}} \rho(\mathbf{p}) d\mathbf{p}. \tag{4.178}$$

Note that the integrals at the right-hand side can be determined in advance and stored in tables.



**Algorithm 4.2** pMOR Strategy in Uncertainty Quantification

- 
- 1: A set  $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^K$  is given in advance, together with frequencies  $s_1, s_2, \dots, s_K$ . In our case the  $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^K$  can come from quadrature points in Stochastic Collocation. Let  $\Psi^k = (s_k, \mathbf{p}^k)$ . Furthermore, let  $\mathbf{A} = s\mathbf{C}(\mathbf{p}) + \mathbf{G}(\mathbf{p})$  and  $\mathbf{A}\mathbf{X} = \mathbf{B}$ , and, similarly,  $\mathbf{A}_k = \mathbf{A}(\Psi^k) = s_k\mathbf{C}(\mathbf{p}^k) + \mathbf{G}(\mathbf{p}^k)$  and  $\mathbf{A}_k\mathbf{X}_k = \mathbf{B}$ .
  - 2: Assume that we have already found some part of the (orthonormal) basis,  $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_k)$
  - 3: For any  $\Psi^j$ , that was not selected before to extend the basis, the actual error formally is given by  $\mathbf{E}^j = \mathbf{X}(\Psi^j) - \sum_{i=1}^k \alpha_i(\Psi^j)\mathbf{v}_i$  and thus for the residue we have  $\mathbf{R}^j = \mathbf{A}_j\mathbf{E}^j = \mathbf{B} - \sum_{i=1}^k \alpha_i(\Psi^j)\mathbf{A}_j\mathbf{v}_i$ . In [85] one determines  $\mathbf{R} = \mathbf{B} \perp \text{Span}(\mathbf{A}_j\mathbf{V})$ , the residue after orthogonalization of  $\mathbf{B}$  against  $\text{Span}(\mathbf{A}_j\mathbf{V})$ . This step does not require evaluation of a solution.
  - 4: Let  $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_m)$ ,  $r_j = \sum_{i=1}^m \|\mathbf{R}_i\|$  and determine  $j_0$  such that  $r_{j_0} = \max_j r_j$ .
  - 5: **if** ( $r_{j_0} > \varepsilon$ ) **then**
  - 6:  $\mathbf{X}(\Psi_{j_0})$  may add most significantly rank to the space spanned by  $\mathbf{V}$ . Hence one now really evaluates  $\mathbf{X}_{j_0} = \mathbf{X}(\Psi_{j_0})$  and orthogonalizes this against  $\mathbf{V}$  and extends  $\mathbf{V}$  with this orthogonal complement. Thus  $\mathbf{X}_{j_0} = \mathbf{X}(\Psi_{j_0}) = [\mathbf{A}_k]^{-1}\mathbf{B} = [\mathbf{A}(\Psi^k)]^{-1}\mathbf{B}$  and  $\mathbf{V}_k = \mathbf{X}_{j_0} - \mathbf{V}(\mathbf{V}^T\mathbf{X}_{j_0})$  is the expansion to  $\mathbf{V}$ . One can use a rank-revealing QR for this step (which also includes a tolerance). Note that until now one collects only zero-moments (in the frequency expansion); for refinements see remarks at the end of this Section.
  - 7: Reduce the set of the  $\Psi^k$  with  $\Psi^{j_0}$ . Go to Step 2.
  - 8: **else**
  - 9: Decide for applying MOR on remainder.
  - 10: **if** (MOR) **then**
  - 11:     **if** (Expressions for  $\mathbf{C}(\mathbf{p})$  and  $\mathbf{G}(\mathbf{p})$  are explicitly known) **then**
  - 12:         Expand the matrices  $\mathbf{C}(\mathbf{p})$  and  $\mathbf{G}(\mathbf{p})$  in polynomials as in (4.181)-(4.182)
  - 13:         Apply the common projection matrix to get the reduced parameterized system.
  - 14:         Apply the collocation to the reduced system (and possibly re-evaluate for parameters used so far the solutions of the reduced system). The solutions of the reduced system at the re-evaluated parameters may be compared to the solutions of the non-reduced system to provide some error control. Note that the expanded expressions provide expressions for the reduced system.
  - 15:         **for all**  $\Psi^k$  **do**
  - 16:             Evaluate  $\mathbf{C}(\mathbf{p}^k)$  and  $\mathbf{G}(\mathbf{p}^k)$  of the reduced system.
  - 17:             Solve the reduced system.
  - 18:         **end for**
  - 19:         *One now has a parameterized reduced system.*
  - 20:     **else**
  - 21:         **for all**  $\Psi^k$  **do**
  - 22:             Evaluate  $\mathbf{C}(\mathbf{p}^k)$  and  $\mathbf{G}(\mathbf{p}^k)$  of the big system (in the CAD environment, say).
  - 23:             Apply the common projection matrix to get the reduced system.
  - 24:             Solve the reduced system.
  - 25:         **end for**
  - 26:     **end if**
  - 27:     Determine the gPC-expansion of the solution of the reduced system.
  - 28:     Perform statistics and/or determine sensitivity of the solution of the reduced system.
  - 29: **else**
  - 30:     Use the Krylov space found so far to efficiently solve all remaining solutions  $\mathbf{X}(\Psi_j)$ . Note that we can use the original expressions in (4.179).
  - 31:     Determine the gPC-expansion of the solution of the original system (4.179).
  - 32:     Perform statistics and/or determine sensitivity of the solution of the original system.
  - 33:     **end if**
  - 34: **end if**
-

In [85] (see also [84]) a parameterized system in the frequency domain

$$[s\mathbf{C}(\mathbf{p}) + \mathbf{G}(\mathbf{p})]\mathbf{x}(s, \mathbf{p}) = \mathbf{B}\mathbf{u}(s), \quad (4.179)$$

$$\mathbf{y}(s, \mathbf{p}) = \mathbf{B}^T \mathbf{x}(s, \mathbf{p}). \quad (4.180)$$

is considered. Here  $s$  is the (angular) frequency. For this system a parameterized MOR approach is proposed, which exploits an expansion of  $\mathbf{C}(\mathbf{p})$  and  $\mathbf{G}(\mathbf{p})$

$$\mathbf{C}(\mathbf{p}) = \sum_{l_1 \dots l_q = 0 \dots 0}^{k_1 \dots k_q} \Phi_{l_1 \dots l_q}(\mathbf{p}) \mathbf{C}_{l_1 \dots l_q}, \quad (4.181)$$

$$\mathbf{G}(\mathbf{p}) = \sum_{l_1 \dots l_q = 0 \dots 0}^{k_1 \dots k_q} \Phi_{l_1 \dots l_q}(\mathbf{p}) \mathbf{G}_{l_1 \dots l_q}, \quad (4.182)$$

$$\Phi_{l_1 \dots l_q}(\mathbf{p}) = p_1^{l_1} p_2^{l_2} \dots p_q^{l_q}. \quad (4.183)$$

In [71] the parameter variation in  $\mathbf{C}$  and  $\mathbf{G}$  did come from parameterized layout extraction of RC circuits.

In Algorithm 4.2 it is assumed that a set  $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^K$  is given in advance, together with frequencies  $s_1, s_2, \dots, s_K$ . Let  $\Psi^k = (s_k, \mathbf{p}^k)$ . Furthermore, let  $\mathbf{A} = s\mathbf{C}(\mathbf{p}) + \mathbf{G}(\mathbf{p})$  and  $\mathbf{A}\mathbf{X} = \mathbf{B}$ , and, similarly,  $\mathbf{A}_k = \mathbf{A}(\Psi^k) = s_k \mathbf{C}(\mathbf{p}^k) + \mathbf{G}(\mathbf{p}^k)$  and  $\mathbf{A}_k \mathbf{X}_k = \mathbf{B}$ .

A projection matrix  $\mathbf{V}$  (with orthonormal columns  $\mathbf{v}_i$ ) is determined such that  $\mathbf{X}(s, \mathbf{p}) \approx \tilde{\mathbf{X}}(s, \mathbf{p}) \equiv \mathbf{V}\hat{\mathbf{X}}(s, \mathbf{p}) \equiv \sum_{i=1}^{K'} \alpha_i(s, \mathbf{p}) \mathbf{v}_i$ . Algorithm 4.2 applies a strategy of which a key step is found in [85]. The extension of  $\mathbf{V}$  is similar to the recycling of Krylov subspaces [103] and used in MOR by [84]. The refinement introduced in [85] is in the selection from the remaining set (steps 5-6). Note that the residues deal with  $\mathbf{B}$  and with  $\mathbf{x}$  and not with the effect in  $\mathbf{y}$ . Hence, one may consider a two-sided projection here. The method of [85] was used in [71] (using expansions of the matrices in moments of  $\mathbf{p}$ ; note that used expressions from layout extraction were linear in  $\mathbf{p}$ ).

This procedure assumes that the evaluation of a matrix  $\mathbf{A}_k$  (and subsequent matrix vector multiplications) is much cheaper than determining a solution  $\mathbf{X}(\Psi_k)$ . Note also that after extending the basis  $\mathbf{V}$  in the next step the norms of the residues should reduce. This allows for some further efficiency in the algorithm [85]. Finally, we remark that the  $\mathbf{X}_k$  are zero order (block) moments at  $\Psi^k$ . After determining the LU-decomposition of  $\mathbf{A}_k$  one easily includes higher moments as well when extending the basis.

A main conclusion of this section is that for the Stochastic Collocation the expansions (4.181)-(4.183) are not explicitly needed by the algorithm. This facilitates dealing with parameters that come from geometry, like scaling [109–113]. The evaluation can completely be done within the CAD environment of the simulation tool - in which case the expressions remain hidden.

The selection of the next parameter introduces a notion of "dominancy" from an algorithmic point of view: this parameter most significantly needs extension of the

Krylov subspace. To invest for this parameter will automatically reduce work for other parameters (several may even drop out of the list because of small residues). If first order sensitivity matrices are available, like in  $\mathbf{C}(\mathbf{p}) = \mathbf{C}_0(\mathbf{p}_0) + \mathbf{C}'(\mathbf{p}_0)\mathbf{p}$  and in  $\mathbf{G}(\mathbf{p}) = \mathbf{G}_0(\mathbf{p}_0) + \mathbf{G}'(\mathbf{p}_0)\mathbf{p}$  one can apply a Generalized Singular Value Decomposition [89] to both pairs  $(\mathbf{C}_0^T(\mathbf{p}_0), [\mathbf{C}'^T(\mathbf{p}_0)])$  and  $(\mathbf{G}_0^T(\mathbf{p}_0), [\mathbf{G}'^T(\mathbf{p}_0)])$ . In [101] this was applied in MOR for linear coupled systems. The low-rank approximations for  $\mathbf{C}'(\mathbf{p}_0)$  and  $\mathbf{G}'(\mathbf{p}_0)$  (obtained by a Generalized SVD [89]) give way to increase the basis for the columns of  $\mathbf{B}$  of the source function. Note that by this one automatically will need MOR methods that can deal with many terminals [69, 97, 116]. In Algorithm 4.2 and in [85] the subspace generated by the basis  $\mathbf{V}$  is slightly increasing with each new  $\mathbf{p}_k$ . A different approach is to apply normal MOR for each  $\mathbf{p}_k$ , giving bases  $\mathbf{V}_k$ , and next determine  $\mathbf{V}$  by an SVD or rank-revealing QR-factorization of  $[\mathbf{V}_k, \dots, \mathbf{V}_K]$ . In [66] this approach is used to obtain a Piecewise  $\mathcal{H}_2$ -Optimal Interpolation pMOR Algorithm.

### References for Section 4.3

63. A.C. ANTOULAS: *Approximation of large-scale dynamical systems*. SIAM Publications, Philadelphia, PA, USA (2005)
64. H. ARMBRUSTER, U. FELDMANN, M. FRERICHS: Analysis based reduction using sensitivity analysis. *Proc. 10-th IEEE Workshop on Signal Propagation on Interconnects (SPI)*, pp. 29–32 (2006)
65. F. AUGUSTIN, A. GILG, M. PAFFRATH, P. RENTROP, U. WEVER: Polynomial chaos for the approximation of uncertainties: Chances and limits. *Euro. J. of Appl. Maths*, Vol. 19, pp. 149–190 (2008)
66. U. BAUR, C. BEATTIE, P. BENNER, S. GUGERCIN: Interpolatory projection methods for parameterized model reduction. *SIAM J. Comput.*, 33, pp. 2489–2518 (2011)
67. P. BENNER: Advances in balancing-related model reduction for circuit simulation. In: J. Roos, L.R.J. Costa (eds.): *Scientific Computing in Electrical Engineering SCEE 2008*. Vol. 14 Math. in Industry, Berlin/Heidelberg, Springer, pp. 469–482 (2010)
68. P. BENNER, M. HINZE, AND E.J.W. TER MATEN (eds.): *Model reduction for circuit simulation*. Lecture Notes in Electrical Engineering, Vol. 74. Springer-Verlag, Berlin, Germany (2011)
69. P. BENNER, A. SCHNEIDER: Model reduction for linear descriptor systems with many ports. In M. Günther, A. Bartel, M. Brunk, S. Schöps, M. Striebel (eds.): *Progress in industrial mathematics at ECMI 2010*, Mathematics in Industry, Springer, pp. 137–143 (2012)
70. P. BENNER, V. MEHRMANN, D. SORENSEN (eds.): *Dimension reduction of large-scale systems*. Lecture Notes in Computational Science and Engineering, Vol. 45. Springer-Verlag, Berlin, Germany (2005)
71. Y. BI, K.-J. VAN DER KOLK, J. FERNÁNDEZ VILLENA, L.M. SILVEIRA, N. VAN DER MEIJS: Fast statistical analysis of RC nets subject to manufacturing variabilities. *Proc. DATE 2011*, Grenoble, March 14–18 (2011)
72. C. CANUTO, M.Y. HUSSAINI, A. QUARTERONI, T.A. ZANG: *Spectral methods - Fundamentals in single domains*. Springer, Berlin (2010)
73. Y. CAO, S. LI, L. PETZOLD: Adjoint sensitivity analysis for differential-algebraic equations: algorithms and software. *SIAM J. Sci. Comput.*, 149, pp. 171–191 (2002)
74. Y. CAO, S. LI, L. PETZOLD, R. SERBAN: Adjoint sensitivity for differential-algebraic equations: the adjoint DAE system and its numerical solution. *SIAM J. Sci. Comput.*, Vol. 24-3, pp. 1076–1089 (2002)

75. Y. CAO, L.R. PETZOLD: A posteriori error estimation and global error control for ordinary differential equations by the adjoint method. *SIAM J. Sci. Comput.*, 26-2, pp. 359–374 (2004)
76. A.R. CONN, R.A. HARING, C. VISWESWARIAH, C.W. WU: Circuit optimization via adjoint Lagrangians. *Proc. ICCAD*, San Jose, CA, USA, Nov. 1997, pp. 281–288 (1997)
77. A.R. CONN, P.K. COULMAN, R.A. HARING, G.L. MORRILL, C. VISWESWARIAH, C.W. WU: JiffyTune: circuit optimization using time-domain sensitivities. *IEEE Trans. on CAD of ICs and Systems*, Vol. 17-12, pp. 1292–1309 (1998)
78. L. DALDOSS, P. GUBIAN, M. QUARANTELLI: Multiparameter time-domain sensitivity computation. *IEEE Trans. on Circuits and Systems - I: Fund. Theory and Applics*, Vol. 48-11, pp. 1296–1307 (2001)
79. D. ECHEVERRÍA CIAURRI: *Multi-level optimization: Space Mapping and Manifold Mapping*. Ph.D.-Thesis, University of Amsterdam, <http://dare.uva.nl/document/45897> (2007)
80. D. ECHEVERRÍA, D. LAHAYE, P.W. HEMKER: Space mapping and defect correction. In [108], pp. 157–176 (2008)
81. O.G. ERNST, A. MUGLER, H.-J. STARKLOFF, E. ULLMANN: On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46, pp. 317–339 (2012)
82. R.M. ERRICO: What is an adjoint model?. *Bulletin of the American Meteorological Society*, 78, 2577-2591 (1997)
83. L. FENG: Parameter independent model order reduction. *Maths. and Comput. in Simul.*, Vol. 68, No. 3, pp. 221–234 (2005)
84. L. FENG, P. BENNER: A robust algorithm for parametric model order reduction. *PAMM - Proc. Appl. Math. Mech.*, 7, pp. 1021501–1021502 (2007)
85. J. FERNÁNDEZ VILLENA, L.M. SILVEIRA: Multi-dimensional automatic sampling schemes for multi-point modeling methodologies. *IEEE Trans. on Comput.-Aided Design of Integr. Circuits and Systems (TCAD)*, 30-8, pp. 1141–1151 (2011)
86. J.G. FIJNVANDRAAT, S.H.M.J. HOUBEN, E.J.W. TER MATEN, J.M.F. PETERS: Time domain analog circuit simulation. *J. of Comp. and Appl. Maths.*, Vol. 185, pp. 441–459 (2006)
87. W. GAUTSCHI: *OPQ: A Matlab suite of programs for generating orthogonal polynomials and related quadrature rules*. <http://www.cs.purdue.edu/archives/2002/wxg/codes> (2002)
88. W. GAUTSCHI: Orthogonal polynomials (in Matlab). *J. of Comput. and Applied Maths.*, 178, pp. 215–234 (2005)
89. G.H. GOLUB, C.F. VAN LOAN: *Matrix Computations, Third edition*. The Johns Hopkins University Press, Baltimore, MD (1996)
90. M. GÜNTHER, U. FELDMANN, J. TER MATEN: Modelling an discretization of circuit problems. In: W.H.A. Schilders, E.J.W. ter Maten (Eds): *Handbook of Numerical Analysis Vol. XIII. Special Volume: Numerical Methods in Electromagnetics*. Elsevier Science, Chapter 6, pp. 523–659 (2005)
91. R. HÄUSLER, H. KINZELBACH: Sensitivity-based stochastic analysis method for power variations. *Proc. Analog '06*, VDE Verlag (2006)
92. P.W. HEMKER, D. ECHEVERRÍA: Manifold mapping for multilevel optimization. In: G. Ciuprina, D. Ioan (Eds.): *Scientific Computing in Electrical Engineering SCEE 2006*, Series Mathematics in Industry 11, Springer, pp. 325–330 (2007)
93. D.E. HOCEVAR, P. YANG, T.N. TRICK, B.D. EPLER: Transient sensitivity computation for MOSFET circuits. *IEEE Trans. on Comput.-Aided Design of Integr. Circuits and Systems (TCAD)*, 4-4, pp. 609–620 (1985)
94. C. HOMESCU, L.R. PETZOLD, R. SERBAN: Error estimation for reduced-order models of dynamical systems. *SIAM Review*, 49-2, pp. 277–299 (2007)
95. Z. ILIEVSKI: *Model order reduction and sensitivity analysis*. Ph.D.-Thesis, TU Eindhoven, <http://alexandria.tue.nl/extra2/201010770.pdf> (2010)
96. Z. ILIEVSKI, H. XU, A. VERHOEVEN, E.J.W. TER MATEN, W.H.A. SCHILDERS, R.M.M. MATTHEIJ: Adjoint transient sensitivity analysis in circuit simulation. In: G. Ciuprina, D. Ioan (eds.): *Scientific Computing in Electrical Engineering SCEE 2006*, Series Mathematics in Industry 11, Springer, pp. 183–189 (2007)

97. R. IONUTIU: *Model order reduction for multi-terminal Systems - with applications to circuit simulation*. Ph.D.-Thesis, TU Eindhoven, <http://alexandria.tue.nl/extra2/716352.pdf> (2011)
98. A. V. KNYAZEV, M.E. ARGENTATI: Principle angles between subspaces in an A-based scalar product: algorithms and perturbation estimates. *SIAM J. Sci. Comput.*, 23-6, pp. 2009-2041 (2002) [Algorithm available in Matlab, The Mathworks, <http://www.mathworks.com/>]
99. J. LANG, J.G. VERWER: On global error estimation and control for initial value problems. *SIAM J. Sci. Comput.*, 29-4, pp. 1460-1475 (2007)
100. O.P. LE MAÎTRE, O.M. KNIO: *Spectral methods for uncertainty quantification, with applications to computational fluid dynamics*. Springer, Science+Business Media B.V., Dordrecht (2010)
101. A. LUTOWSKA: *Model order reduction for coupled systems using low-rank approximations*. Ph.D.-Thesis, TU Eindhoven, <http://alexandria.tue.nl/extra2/729804.pdf> (2012)
102. E.J.W. TER MATEN, T.G.A. HEIJMEN, A. LIN, A. EL GUENNOUNI: Optimization of electronic circuits. In: V. Cutello, G. Fotia, L. Puccio (eds.): *Applied and Industrial Mathematics in Italy II, Selected contributions from the 8th SIMAI conference*. Series on Advances in Mathematics for Applied Sciences, Vol. 75, World Scientific Publishing Co. Pte. Ltd., Singapore, pp. 573–584 (2007)
103. M.L. PARKS, E. DE STURLER, G. MACKEY, D.D. JOHNSON, S. MAITI: Recycling Krylov subspaces for sequences of linear systems. *SIAM J. Sci. Comput.*, 28-5, pp. 1651–1674 (2006)
104. R. PINNAU: Model reduction via proper orthogonal decomposition. In: [108] pp. 95–109 (2008)
105. M. RATHINAM, L.R. PETZOLD: A new look at proper orthogonal decomposition. *SIAM J. Numer. Anal.*, 41-5, pp. 1893–1925 (2003)
106. W.H.A. SCHILDERS: Introduction to model order reduction. In [108], pp. 3–32 (2008)
107. W.H.A. SCHILDERS: The need for novel model order reduction techniques in the electronics industry. In [68], pp. 3–23 (2011)
108. W.H.A. SCHILDERS, H.A. VAN DER VORST, J. ROMMES (eds.): *Model order reduction: theory, research aspects and applications*. Mathematics in Industry Vol. 13. Springer-Verlag, Berlin, Heidelberg (2008)
109. K. STAVRAKAKIS, T. WITTIG, W. ACKERMANN, T. WEILAND: Linearization of parametric FIT-discretized systems for model order reduction. *IEEE Trans. on Magnetics*, 45-3, pp. 1380–1383 (2009)
110. K. STAVRAKAKIS, T. WITTIG, W. ACKERMANN, T. WEILAND: Threedimensional geometry variations of FIT systems for model order reduction. *2010 URSI Int. Symposium on Electromagnetic Theory*, pp. 788–791 (2010)
111. K. STAVRAKAKIS, T. WITTIG, W. ACKERMANN, T. WEILAND: Model order reduction methods for multivariate parameterized dynamical systems obtained by the Finite Integration Theory. *2011 URSI General Assembly and Scientific Symposium*, 4p (2011)
112. K. STAVRAKAKIS, T. WITTIG, W. ACKERMANN, T. WEILAND: Parametric model order reduction by neighbouring subspaces. In: B. Michielsen, J.-R. Poirier (eds.): *Scientific Computing in Electrical Engineering SCEE 2010*, Series Mathematics in Industry, Vol. 16, Springer, pp. 443–451 (2012)
113. K.K. STAVRAKAKIS: *Model order reduction methods for parameterized systems in electromagnetic field simulations*. Ph.D.-Thesis, TU-Darmstadt (2012)
114. A. STROUD: *Approximate Calculation of Multiple Integrals*. Prentice Hall (1971)
115. SUMO (SURrogate MOdeling) Lab. IBCN research group of the Dept. of Information Technology (INTEC), Ghent University, <http://www.sumo.intec.ugent.be/> (2012)
116. M.V. UGRYUMOVA: *Applications of Model Order Reduction for IC Modeling*. Ph.D.-Thesis, TU Eindhoven, <http://alexandria.tue.nl/extra2/711015.pdf> (2011)
117. S. VOLKWEIN: *Model reduction using proper orthogonal decomposition*. <http://www.uni-graz.at/imawww/volkwein/POD.pdf> (2008)

118. D. XIU: Numerical integration formulas of degree two. *Appl. Numer. Maths.*, Vol. 58, pp. 1515–1520 (2008)
119. D. XIU: Fast numerical methods for stochastic computations: a review. *Commun. Comput. Phys.*, Vol. 5, No. 2-4, pp. 242–272 (2009)
120. D. XIU: *Numerical methods for stochastic computations - A spectral method approach*. Princeton Univ. Press, Princeton, NJ, USA (2010)

#### 4.4 MOR for singularly perturbed systems

<sup>4</sup>For large systems of ordinary differential equations (ODEs), efficient MOR methods already exist in the linear case, see [121]. We want to generalize according techniques to the case of differential-algebraic equations (DAEs). On the one hand, a high-index DAE problem can be converted into a lower-index system by analytic differentiations, see [123]. A transformation to index zero yields an equivalent system of ODEs. On the other hand, a regularization is directly feasible in case of semi-explicit systems of DAEs. Thereby, we obtain a singularly perturbed problem of ODEs with an artificial parameter. Thus according MOR techniques can be applied to the ODE system. An MOR approach for DAEs is achieved by considering the limit to zero of the artificial parameter.

We consider a simplified, semi-explicit DAE system to illustrate some concepts only

$$\begin{aligned} \dot{\mathbf{y}}(t) &= \mathbf{f}(\mathbf{y}(t), \mathbf{z}(t)), & \mathbf{y} : \mathbb{R} &\rightarrow \mathbb{R}^k, \\ \mathbf{0} &= \mathbf{g}(\mathbf{y}(t), \mathbf{z}(t)), & \mathbf{z} : \mathbb{R} &\rightarrow \mathbb{R}^l, \end{aligned} \quad (4.184)$$

with differential and perturbation index 1 or 2. For the construction of numerical methods to solve initial value problems of (4.184), a direct as well as an indirect approach can be used. The direct approach applies an  $\varepsilon$ -embedding of the DAEs (4.184), i.e., the system changes into

$$\begin{aligned} \dot{\mathbf{y}}(t) &= \mathbf{f}(\mathbf{y}(t), \mathbf{z}(t)) & \Leftrightarrow & \dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), \mathbf{z}(t)) \\ \varepsilon \dot{\mathbf{z}}(t) &= \mathbf{g}(\mathbf{y}(t), \mathbf{z}(t)) & & \dot{\mathbf{z}}(t) = \frac{1}{\varepsilon} \mathbf{g}(\mathbf{y}(t), \mathbf{z}(t)) \end{aligned} \quad (4.185)$$

with a real parameter  $\varepsilon \neq 0$ . Techniques for ODEs can be employed for the singularly perturbed system (4.185). The limit  $\varepsilon \rightarrow 0$  yields an approach for solving the DAEs (4.184). The applicability and quality of the resulting method still has to be investigated.

Alternatively, the indirect approach is based on the *state space form* of the DAEs (4.184) with differential and perturbation index 1 or 2, for nonlinear cases see [135], i.e.,

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), \Phi(\mathbf{y}(t))) \quad (4.186)$$

with  $\mathbf{z}(t) = \Phi(\mathbf{y}(t))$ . To evaluate the function  $\Phi$ , the nonlinear system

---

<sup>4</sup> Section 4.4 has been written by: Kasra Mohaghegh, Roland Pulch and E. Jan W. ter Maten. For an extended version we refer to the Ph.D.-Thesis [132] of the first author and to the papers [131, 133].

$$\mathbf{g}(\mathbf{y}(t), \Phi(\mathbf{y}(t))) = 0 \quad (4.187)$$

is solved for given value  $\mathbf{y}(t)$ . Consequently, the system (4.186) represents ODEs for the differential variables  $\mathbf{y}$  and ODE methods can be applied. In each evaluation of the right-hand side in (4.186), a nonlinear system (4.187) has to be solved. More details on techniques based on the  $\varepsilon$ -embedding and the state space form can be found in [128].

Although some MOR methods for DAEs already exist, several techniques are restricted to ODEs or exhibit better properties in the ODE case in comparison to the DAE case. The direct or the indirect approach enables the usage of MOR schemes for ODEs (4.185) or (4.186), where an approximation with respect to the original DAEs (4.184) follows. The aim is to obtain suggestions for MOR schemes via these strategies, where the quality of the resulting approximations still has to be analyzed in each method.

In this section, we focus on the direct approach for semi-explicit system of DAEs, i.e., the  $\varepsilon$ -embedding (4.185) is considered. MOR methods are applied to the singularly perturbed system (4.185). Two scenarios exist to achieve an approximation of the behavior of the original DAEs (4.184) by MOR. Firstly, an MOR scheme can be applied to the system (4.185) using a constant  $\varepsilon \neq 0$ , which is chosen sufficiently small (on a case by case basis) such that a good approximation is obtained. Secondly, a parametric or parameterized Model Order Reduction (pMOR) method yields a reduced description of the system of ODEs, where the parameter  $\varepsilon$  still represents an independent variable. Hence the limit  $\varepsilon \rightarrow 0$  causes an approach for an approximation of the original DAEs.

We investigate the two approaches with respect to MOR methods based on an approximation of the transfer function, which describes the input-output behavior of the system in frequency domain.

#### 4.4.1 Model Order Reduction and $\varepsilon$ -Embedding

We restrict ourselves to semi-explicit DAE systems of the type (4.188)-(4.189) and introduce  $w(t)$  as an output instead of  $y(t)$  with exact the same condition. According to (4.184), after linearizing, we can write the system as

$$\mathbf{C}\dot{\mathbf{x}} = -\mathbf{G}\mathbf{x} + \mathbf{B}\mathbf{u}(t), \quad (4.188)$$

$$\mathbf{w}(t) = \mathbf{L}\mathbf{x}(t). \quad (4.189)$$

The solution  $\mathbf{x}$  and the matrix  $\mathbf{C}$  exhibit the partitioning:

$$\mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{I}_{k \times k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{l \times l} \end{pmatrix}.$$

$\mathbf{w}(t)$  is the output of the system. The order of the system is  $n = k + l$ , where  $k$  and  $l$  are the dimensions of the differential part and the algebraic part (constraints), re-

spectively, defined in the semi-explicit system (4.184).  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ;  $\mathbf{L} \in \mathbb{R}^{p \times n}$ . After taking the Laplace transform, the corresponding  $p \times m$  matrix-valued rational transfer function is

$$\mathbf{H}(s) = \mathbf{L} \cdot (\mathbf{G} + s\mathbf{C})^{-1} \cdot \mathbf{B} = \mathbf{L} \cdot \left( \mathbf{G} + s \begin{pmatrix} \mathbf{I}_{k \times k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{l \times l} \end{pmatrix} \right)^{-1} \cdot \mathbf{B},$$

provided that  $\det(\mathbf{G} + s\mathbf{C}) \neq 0$  and  $\mathbf{x}(0) = \mathbf{0}$  and  $\mathbf{u}(0) = \mathbf{0}$ . Following the direct approach [132], the  $\varepsilon$ -embedding changes the system (4.188)-(4.189) into:

$$\begin{cases} \mathbf{C}(\varepsilon) \frac{d\mathbf{x}(t)}{dt} = -\mathbf{G}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), & \mathbf{x}(0) = \mathbf{x}_0, \\ \mathbf{w}(t) = \mathbf{L}\mathbf{x}(t), \end{cases} \quad (4.190)$$

where

$$\mathbf{C}(\varepsilon) = \begin{pmatrix} \mathbf{I}_{k \times k} & \mathbf{0} \\ \mathbf{0} & \varepsilon \mathbf{I}_{l \times l} \end{pmatrix} \quad \text{for } \varepsilon \in \mathbb{R}$$

with the same inner state and input/output as before. For  $\varepsilon \neq 0$ , the matrix  $\mathbf{C}(\varepsilon)$  is regular in (4.190) and the transfer function reads:

$$\mathbf{H}_\varepsilon(s) = \mathbf{L} \cdot (\mathbf{G} + s \cdot \mathbf{C}(\varepsilon))^{-1} \cdot \mathbf{B}$$

provided that  $\det(\mathbf{G} + s\mathbf{C}(\varepsilon)) \neq 0$ . For convenience, we introduce the notation

$$\mathbf{M}(s, \varepsilon) := s\mathbf{C}(\varepsilon) = s \begin{pmatrix} \mathbf{I}_{k \times k} & \mathbf{0} \\ \mathbf{0} & \varepsilon \mathbf{I}_{l \times l} \end{pmatrix}.$$

It holds  $\mathbf{M}(s, 0) = s\mathbf{C}$  with  $\mathbf{C}$  from (4.188).

Concerning the relation between the original system (4.188)-(4.189) and the regularized system (4.190) with respect to the transfer function, we achieve the following statement. Without loss of generality, the induced matrix norm of the Euclidean vector norm is applied.

**Lemma 4.1.** *Let  $\mathbf{A}, \tilde{\mathbf{A}} \in \mathbb{R}^{n \times n}$ ,  $\det(\mathbf{A}) \neq 0$  and  $\|\mathbf{A} - \tilde{\mathbf{A}}\|_2 = \|\Delta\mathbf{A}\|_2$  where  $\Delta\mathbf{A}$  is small enough. Then it holds:*

$$\|\mathbf{A}^{-1} - \tilde{\mathbf{A}}^{-1}\|_2 \leq \frac{\|\mathbf{A}^{-1}\|_2^2 \cdot \|\Delta\mathbf{A}\|_2}{1 - \|\mathbf{A}^{-1}\|_2 \cdot \|\Delta\mathbf{A}\|_2}.$$

*Proof.* It holds

$$\|\mathbf{A}^{-1} - \tilde{\mathbf{A}}^{-1}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A}^{-1}\mathbf{x} - \tilde{\mathbf{A}}^{-1}\mathbf{x}\|_2.$$

Suppose  $\mathbf{y} := \mathbf{A}^{-1}\mathbf{x}$ ,  $\tilde{\mathbf{y}} := \tilde{\mathbf{A}}^{-1}\mathbf{x}$ , then the sensitivity analysis of linear systems yields



$$\frac{\|\Delta \mathbf{y}\|_2}{\|\mathbf{y}\|_2} \leq \frac{\kappa(\mathbf{A})}{1 - \kappa(\mathbf{A}) \frac{\|\Delta \mathbf{A}\|_2}{\|\mathbf{A}\|_2}} \left( \frac{\|\Delta \mathbf{A}\|_2}{\|\mathbf{A}\|_2} + \underbrace{\frac{\|\Delta \mathbf{x}\|_2}{\|\mathbf{x}\|_2}}_{=0} \right),$$

where the quantity

$$\kappa(\mathbf{A}) \equiv \|\mathbf{A}^{-1}\|_2 \|\mathbf{A}\|_2$$

is the relative *condition number*. So by substituting the value of  $\kappa(\mathbf{A})$  we have:

$$\|\mathbf{y} - \tilde{\mathbf{y}}\|_2 \leq \frac{\|\mathbf{A}^{-1}\|_2 \cdot \|\Delta \mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \|\mathbf{x}\|_2}{1 - \|\mathbf{A}^{-1}\|_2 \cdot \|\Delta \mathbf{A}\|_2}$$

then

$$\|\mathbf{A}^{-1} - \tilde{\mathbf{A}}^{-1}\|_2 \leq \frac{\|\mathbf{A}^{-1}\|_2^2 \cdot \|\Delta \mathbf{A}\|_2}{1 - \|\mathbf{A}^{-1}\|_2 \cdot \|\Delta \mathbf{A}\|_2}. \quad \square$$

We conclude from Lemma 4.1 that

$$\lim_{\Delta \mathbf{A} \rightarrow 0} \tilde{\mathbf{A}}^{-1} = \mathbf{A}^{-1},$$

for example.

**Theorem 4.1.** For fixed  $s \in \mathbb{C}$  with  $\det(\mathbf{G} + \mathbf{M}(s, 0)) \neq 0$  and  $\varepsilon \in \mathbb{R}$  satisfying

$$|s| \cdot |\varepsilon| \leq \frac{c}{\|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2} \quad (4.191)$$

for some  $c \in (0, 1)$ , the transfer functions  $\mathbf{H}(s)$  and  $\mathbf{H}_\varepsilon(s)$  of the systems (4.188)-(4.189) and (4.190) exist and it holds

$$\|\mathbf{H}(s) - \mathbf{H}_\varepsilon(s)\|_2 \leq \|\mathbf{L}\|_2 \cdot \|\mathbf{B}\|_2 \cdot K(s) \cdot |s| \cdot |\varepsilon|$$

with

$$K(s) = \frac{1}{1-c} \|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2^2.$$

*Proof.* Let  $\mathbf{A} = \mathbf{G} + \mathbf{M}(s, 0)$  and  $\tilde{\mathbf{A}} = \mathbf{G} + \mathbf{M}(s, \varepsilon)$ . The condition (4.191) guarantees that the matrices  $\tilde{\mathbf{A}}$  are regular. The definition of the transfer functions implies:

$$\|\mathbf{H}(s) - \mathbf{H}_\varepsilon(s)\|_2 \leq \|\mathbf{L}\|_2 \cdot \|\mathbf{A}^{-1} - \tilde{\mathbf{A}}^{-1}\|_2 \cdot \|\mathbf{B}\|_2.$$

We obtain:

$$\|\mathbf{A} - \tilde{\mathbf{A}}\|_2 = \|\mathbf{M}(s, 0) - \mathbf{M}(s, \varepsilon)\|_2 = |s| \cdot \left\| \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \varepsilon \mathbf{I}_{l \times l} \end{pmatrix} \right\|_2 = |s| \cdot |\varepsilon|.$$

Applying the Lemma 4.1, the term at the right-hand side of the expression above becomes:

$$\begin{aligned}
\|\mathbf{A}^{-1} - \tilde{\mathbf{A}}^{-1}\|_2 &\leq \frac{\|\mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{M}(s, 0) - \mathbf{M}(s, \varepsilon)\|_2}{1 - \|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{M}(s, 0) - \mathbf{M}(s, \varepsilon)\|_2} \\
&\leq \frac{1}{1-c} \|\mathbf{A}^{-1}\|_2^2 \cdot \|\mathbf{M}(s, 0) - \mathbf{M}(s, \varepsilon)\|_2 \\
&\leq K(s) \|\mathbf{M}(s, 0) - \mathbf{M}(s, \varepsilon)\|_2.
\end{aligned}$$

Thus the proof is completed.  $\square$

It is clear that for inequality (4.191) we have:

$$\begin{aligned}
s \neq 0 \in \mathbb{C} : |\varepsilon| &\leq \frac{c}{|s| \cdot \|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2} \\
s = 0 \in \mathbb{C} : &\quad \varepsilon \text{ arbitrary}
\end{aligned}$$

We conclude from Theorem 4.1 that

$$\lim_{\varepsilon \rightarrow 0} \mathbf{H}_\varepsilon(s) = \mathbf{H}(s)$$

for each  $s \in \mathbb{C}$  with  $\mathbf{G} + s\mathbf{C}$  regular. The relation (4.191) gives feasible domains of  $\varepsilon$

$$\begin{aligned}
|s| \leq 1 : |\varepsilon| &\leq \frac{c}{\|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2}, \\
|s| > 1 : |\varepsilon| &\leq \frac{c}{|s| \cdot \|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2}.
\end{aligned}$$

We also obtain the uniform convergence

$$\|\mathbf{H}(s) - \mathbf{H}_\varepsilon(s)\|_2 \leq \hat{K} |\varepsilon| \quad \text{for all } s \in S$$

in a compact domain  $S \subset \mathbb{C}$  and  $\varepsilon \leq \delta$  with:

$$\begin{aligned}
\delta &= c \cdot \min_{s \in S} \frac{1}{\|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2} \quad \text{for } \tilde{S} = \emptyset, \\
\delta &= c \cdot \left[ \min_{s \in S} \frac{1}{\|(\mathbf{G} + \mathbf{M}(s, 0))^{-1}\|_2} \right] \cdot \underbrace{\left[ \min_{s \in \tilde{S}} \frac{1}{|s|} \right]}_{\leq 1} \quad \text{for } \tilde{S} \neq \emptyset,
\end{aligned}$$

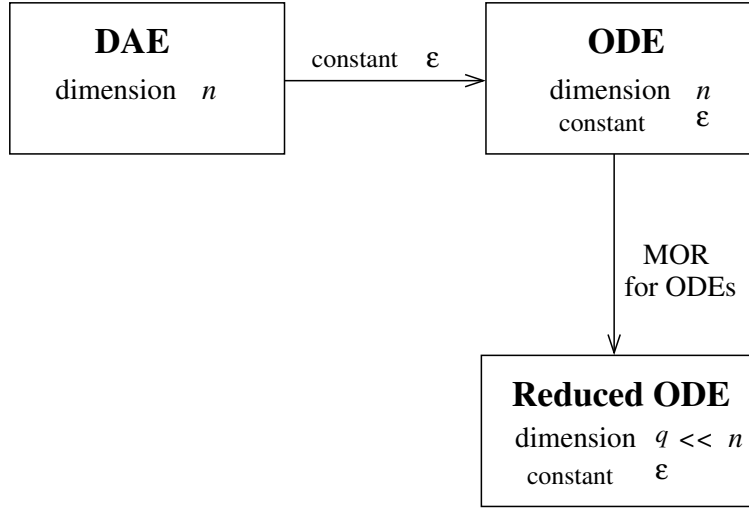
with  $\tilde{S} := \{z \in S : |z| \geq 1\}$ . Furthermore, Theorem 4.1 implies the property

$$\lim_{s \rightarrow 0} \mathbf{H}(s) - \mathbf{H}_\varepsilon(s) = 0$$

for fixed  $\varepsilon$  assuming  $\det \mathbf{G} \neq 0$ . However, we are not interested in the limit case of small variables  $s$ .

For reducing the DAE system (4.188)-(4.189), we have two ways to handle the artificial parameter  $\varepsilon$ , which results in two different scenarios. In the first scenario, we fix a small value of the parameter  $\varepsilon$ . Thus we use one of the standard techniques for the reduction of the corresponding ODE system. Finally, we achieve a reduced ODE (with small  $\varepsilon$  inside). The ODE system with small  $\varepsilon$  represents a regularized

DAE. Any reduction scheme for ODEs is feasible. Recent research shows that the Poor Man's TBR (PMTBR), see [134], can be applied efficiently to the ODE case. Figure 4.37 indicates the steps for the first scenario.



**Fig. 4.37** The approach of the  $\epsilon$ -embedding for MOR in the first scenario.

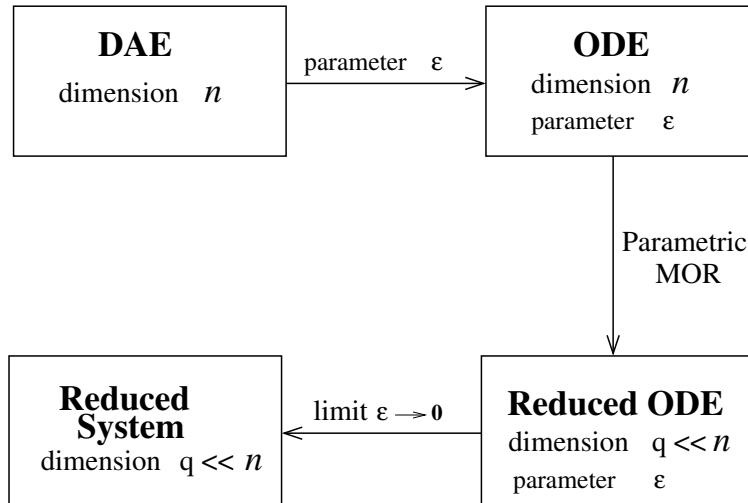
In the second scenario, the parameter  $\epsilon$  is considered as an independent variable (value not predetermined). We can use the parametric MOR for reducing the corresponding ODE system. The applied parametric MOR is based on [124, 125] in this case. The limit  $\epsilon \rightarrow 0$  yields the results in an approximation of original DAEs (4.188)-(4.189). The existence of the approximation in this limit still has to be analyzed. Figure 4.38 illustrates the strategy for the second scenario.

Theorem 4.1 provides the theoretical background for the both scenarios. We apply an MOR scheme based on an approximation of the transfer function to the system of ODEs (4.190). Let  $\tilde{\mathbf{H}}_\epsilon(s)$  be a corresponding approximation of  $\mathbf{H}_\epsilon(s)$ .

It follows

$$\|\mathbf{H}(s) - \tilde{\mathbf{H}}_\epsilon(s)\|_2 \leq \|\mathbf{H}(s) - \mathbf{H}_\epsilon(s)\|_2 + \|\mathbf{H}_\epsilon(s) - \tilde{\mathbf{H}}_\epsilon(s)\|_2 \quad (4.192)$$

for each  $s \in \mathbb{C}$  with  $\det(\mathbf{G} + s\mathbf{C}) \neq 0$ . Due to Theorem 4.1, the first term becomes small for sufficiently small parameter  $\epsilon$ . However,  $\epsilon$  should not be chosen smaller than the machine precision on a computer. The second term depends on the applicability of an efficient MOR method to the ODEs (4.190). Thus  $\tilde{\mathbf{H}}_\epsilon(s)$  can be seen as an approximation of the transfer function  $\mathbf{H}(s)$  belonging to the system of DAEs (4.188)-(4.189).



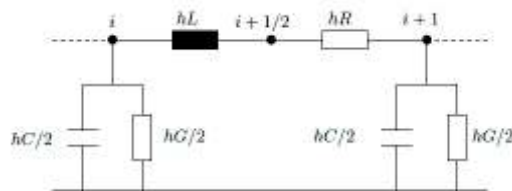
**Fig. 4.38** The approach of the  $\varepsilon$ -embedding for MOR in the second scenario.

#### 4.4.2 Test Example and Numerical Results

We consider a substitute model of a transmission line (TL), see [127], which consists of  $N$  cells. Each cell includes a capacitor, an inductor and two resistors, see Figure 4.39. This TL model represents a scalable benchmark problem (both in differential part and algebraic part but not separately), because we can select the number  $N$  of cells. The used physical parameters are

$$C = 10^{-14} \text{ F/m}, L = 10^{-8} \text{ H}, R = 0.1 \text{ } \Omega/\text{m}, G = 10 \text{ S/m}.$$

We apply modified nodal analysis, see [126], to the RLC circuit and then the state



**Fig. 4.39** One cell of the RLC transmission line.

variables  $\mathbf{x} \in \mathbb{R}^{3N+3}$  consist of the voltages at the nodes, the currents traversing the inductances  $L$  and the currents at the boundaries of the circuit:

$$\begin{aligned} & (V_0, V_1, \dots, V_N), & (I_{1/2}, I_{3/2}, \dots, I_{N-1/2}), \\ & (V_{1/2}, V_{3/2}, \dots, V_{N-1/2}), & (I_0, I_N). \end{aligned}$$

So far we have  $3N + 3$  unknowns and only  $3N + 1$  equations. Thus two boundary conditions are necessary. Equations for the main nodes and the intermediate nodes in each cell are

$$\begin{aligned} \frac{h}{2}C\dot{V}_0 + \frac{h}{2}GV_0 + I_{1/2} - I_0 &= 0, \\ hC\dot{V}_i + hGV_i + I_{i+1/2} - I_{i-1/2} &= 0, \quad i = 1, \dots, N-1, \\ \frac{h}{2}C\dot{V}_N + \frac{h}{2}GV_N + I_N - I_{N-1/2} &= 0, \\ -I_{i+1/2} + \frac{V_{i+1/2} - V_{i+1}}{hR} &= 0, \\ hL\dot{I}_{i+1/2} + (V_{i+1/2} - V_i) &= 0, \quad i = 0, 1, \dots, N-1, \end{aligned}$$

where the variable  $h > 0$  represents a discretization step size in space. We apply the boundary conditions

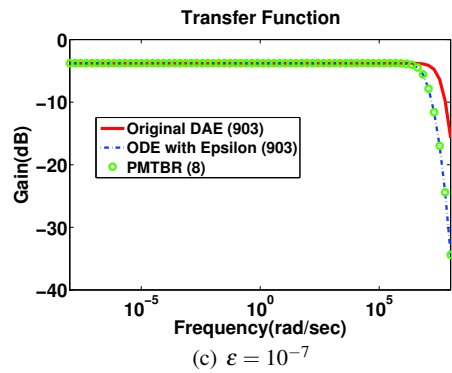
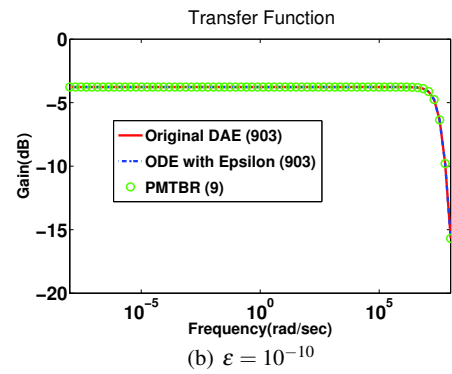
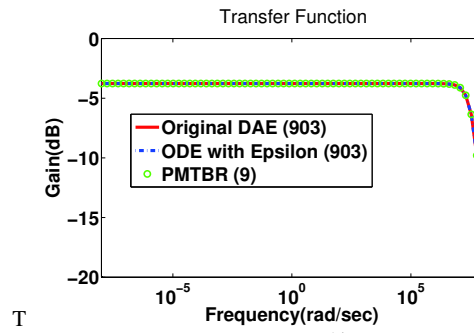
$$\begin{aligned} I_0 - u(t) &= 0, \\ L_1\dot{I}_N + V_N &= 0 \end{aligned}$$

with  $L_1 > 0$  and an independent current source  $u$ . Now a direct approach ( $\varepsilon$ -embedding) is used. For the first simulation the variable  $\varepsilon$  is fixed to  $10^{-14}$  and  $10^{-7}$ , respectively, and the PMTBR method [134] is used as a reduction scheme for the ODE system. For all runs we selected the number of cells to  $N = 300$ , which results in the order  $n = 903$  of the original system of DAEs (4.188)-(4.189). Figure 4.40 shows the transfer function both for the DAE and the ODE (including  $\varepsilon$ ) and the reduced ODE with fixed  $\varepsilon$  for frequencies  $s = i\omega$  with  $\omega \in \mathbb{R}$ . The number in parentheses shows the order of the systems.

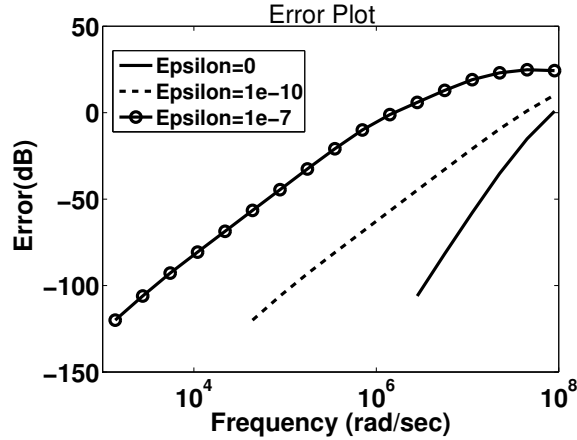
Finally the second scenario with parametric MOR is studied. We apply the PIMTAB parametric MOR following [129, 130]. The limit  $\varepsilon \rightarrow 0$  gives the result for the reduced DAE. The error plot for the parametric reduction scheme is shown in Figure 4.41. The error plot shows an overall nice match for the case of  $\varepsilon = 0, 10^{-10}$  and as the value for the parameter  $\varepsilon$  increases, the accuracy of the method and of the reduction algorithm decreases. It is also important to mention that the order of the reduced system in the second scenario is nearly half of the one in the first scenario. Table 4.4 shows which value for the parameter  $\varepsilon$  is acceptable for the both scenarios.

**Table 4.4** Acceptance of the method: different values for  $\varepsilon$  are mentioned. A dash indicates that the error is not calculated; A. and N.A indicate accepted and not accepted, respectively.

Value used for $\varepsilon$ :	0	$10^{-14}$	$10^{-10}$	$10^{-7}$
Scenario with fixed $\varepsilon$	Same as DAE	A.	A.	N.A.
Scenario with parametric $\varepsilon$	A.	-	A.	N.A.



**Fig. 4.40** Original transfer function for DAE and ODE and reduced transfer function of PMTBR in case of three different parameters  $\varepsilon$ . The frequency  $\omega$  ranges from  $10^{-8}$  to  $10^8$ .



**Fig. 4.41** Absolute error plot for the transfer function in the  $\varepsilon$ -embedding, reduction carried out by parametric MOR with PIMTAB,  $\varepsilon = 0, 10^{-10}, 10^{-7}$ .

#### 4.4.3 Conclusions

In this section we applied the  $\varepsilon$ -embedding to approximate a linear system of DAEs by a system of ODEs. We did consider the transfer function in the frequency domain as a function of  $\varepsilon$  and proved uniform convergence for frequencies  $s$  in a compact region  $S$  where the matrix  $\mathbf{G} + s\mathbf{C}$  is regular (and thus its inverse uniformly bounded). This motivated the usage of MOR methods for ODEs. Most of the reduction schemes are designed and adopted for linear ODEs. Well-known methods are PMTBR (Poor Man's Truncated Balanced Realization [134]) and the spectral zeros preservation MOR of Antoulas [122].

In the first scenario we applied a fixed  $\varepsilon$  and studied for a transmission line model the behavior of the transfer functions of the DAE, of the ODE and of the reduced model obtained with PMTBR for  $\varepsilon = 10^{-14}, 10^{-10}, 10^{-7}$ . Already for the last value the transfer functions between DAE and ODE differ significantly. If we choose bigger values for  $\varepsilon$ , the system is more friendly but the error is larger and the solution will be changed. On the other hand the transfer function obtained by PMTBR is able to approximate quite well the transfer function of the ODE.

In the second approach we applied the parametric MOR technique PIMTAB [129, 130] to the parameterized ODE. Here we do not need to predefine the value of the  $\varepsilon$ . We obtain a parameterized MOR that gives a reduced model for  $\varepsilon = 0$  for which the transfer function approximates well the one for the DAE [132, 133].

## References for Section 4.4

121. ANTOULAS, A.C.: *Approximation of large-scale Dynamical Systems, advances in design and control*. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2005)
122. ANTOULAS, A.C.: A new result on passivity preserving model reduction. *System & Control Letters*, vol. 54, pp. 361–374 (2005)
123. ASCHER, U.M., PETZOLD, L.R.: *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (1998)
124. DANIEL, L., SIONG, O.C., CHAY, L.S., LEE, K.H., WHITE, J.: A multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Trans. Computer-Aided Design*, 23(5), pp. 678–693 (2004)
125. FENG, L., BENNER, P.: A robust algorithm for parametric model order reduction based on implicit moment matching. *PAMM Proc. Appl. Math. Mech.*, 7, pp. 1021501–1021502 (2007)
126. GÜNTHER, M., FELDMANN, U.: CAD based electric circuit modeling in industry I: mathematical structure and index of network equations. *Surv. Math. Ind.*, 8, pp. 97–129 (1999)
127. GÜNTHER, M.: *Partielle differential-algebraische Systeme in der numerischen Zeitbereichsanalyse elektrischer Schaltungen*. Nr. 343 in Fortschritt-Berichte VDI Serie 20. VDI, Düsseldorf (2001)
128. HAIRER, E., WANNER, G.: *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. Springer, Berlin, 2nd edition (1996)
129. LI, Y., BAI, Z., SU, Y., ZENG, X.: Parameterized model order reduction via a two-directional Arnoldi process. In: *Proc. IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 868–873 (2007)
130. LI, Y.-T., BAI, Z., SU, Y., ZENG, X.: Model order reduction of parameterized interconnect networks via a two-directional Arnoldi process. *IEEE Trans. on Comput.-Aided Design of Integr. Circuits and Systems (TCAD)*, 27(9), pp. 1571–1582 (2008)
131. MOHAGHEGH, K., PULCH, R., STRIEBEL, M., TER MATEN, J.: Model order reduction for semi-explicit systems of differential algebraic systems. In: I. Troch, F. Breitenecker (eds): *Proceedings MATHMOD 09 Vienna - Full Papers CD Volume*, pp. 1256–1265 (2009)
132. MOHAGHEGH, K.: *Linear and nonlinear model order reduction for numerical simulation of electric circuits*. Ph.D.-Thesis, Bergische Universität Wuppertal, Germany. Available at Logos Verlag, Berlin, Germany (2010)
133. MOHAGHEGH, K., PULCH, R., TER MATEN, J.: *Model order reduction using singularly perturbed systems*. Submitted (2013)
134. PHILLIPS, J., SILVEIRA, L.M.: Poor’s man TBR: a simple model reduction scheme. In: *Proc. Design, Automation and Test in Europe Conference and Exhibition (DATE) 2004*, volume 2, pp. 938–943, 2004.
135. SCHWARZ, D.E., TISCHENDORF, C.: Structural analysis of electric circuits and consequences for MNA. *International Journal of Circuit Theory and Applications*, 28, pp. 131–162, 2000.



**PREVIOUS PUBLICATIONS IN THIS SERIES:**

Number	Author(s)	Title	Month
15-04	E.N.M. Cirillo M. Colangeli A. Muntean	Threshold effects in zero range processes	Febr. '15
15-05	A.Garroni P.J.P. van Meurs M.A. Peletier L. Scardia	Boundary-layer analysis of a pile-up of walls of edge dislocations at a lock	Febr. '15
15-06	B. Tasić J.J. Dohmen E.J.W. ter Maten T.G.J. Beelen H.H.J.M. Janssen W.H.A. Schilders M. Günther	Fast fault simulation to identify subcircuits involving faulty components	Febr. '15
15-07	A.C. Antoulas R. Ionutiu N. Martins E.J.W. ter Maten K. Mohaghegh R. Pulch J. Rommes M. Saadvandi M. Striebel	Model order reduction – Methods, concepts and properties	Febr. '15
15-08	G. Ciuprina J. Fernández Villena D. Ioan Z. Ilievski S. Kula E.J.W. ter Maten K. Mohaghegh R. Pulch W.H.A. Schilders L. Miguel Silveira A. Ștefănescu M. Striebel	Parameterized model order reduction	Febr. '15