Model order reduction for large systems in computational electromagnetics

Tilmann Wittig*,1, Rolf Schuhmann, Thomas Weiland

Technische Universität Darmstadt, Institut für Theorie Elektromagnetischer Felder (TEMF),
Schloßgartenstr. 8, 64289 Darmstadt, Germany

Received 31 March 2003; accepted 10 June 2004
Available online 18 August 2004
Submitted by P. Benner

Abstract

This paper examines classical Model Order Reduction (MOR) strategies in view of the particular properties and needs of computational electromagnetism. Hereby reduced models are mainly needed for two reasons: The fast calculation of certain characteristics such as the transfer behavior over a range of excitation frequencies—especially for highly resonant systems—and the generation of macromodels such as equivalent circuits simplifying coupled simulations. In the first case, the computational cost of the method is of main interest, while in the second one the size of the reduced model as well as the conservation of stability and passivity plays a major role. The considered methods—partial realization, moment matching and modal extraction—are well-known and have been investigated for about two decades now. However, their efficiency appears in a different light if the number of unknowns reaches hundreds of thousands or even millions.

This paper compares the suitability and efficiency of the mentioned methods for lossless or weakly lossy structures discretized by the Finite Integration Theory (FIT). Close relations and even transitions between the algorithms are shown. Finally, some specific properties of FIT enable the application of a method called Two-Step Lanczos (TSL): a successive application of partial realization and moment matching which is highly efficient in both computation time

* Corresponding author. Tel.: +49 6151 162161; fax: +49 6151 164611.
E-mail addresses: wittig@temf.tu-darmstadt.de (T. Wittig), schuhmann@temf.tu-darmstadt.de (R. Schuhmann), weiland@temf.tu-darmstadt.de (T. Weiland).

1 The work of T. Wittig was generously supported by GSI Darmstadt.

0024-3795/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2004.06.023
and model size, while preserving the passivity of the reduced models. TSL allows to compute the broadband transfer behavior of systems with hundreds of thousands of unknowns within minutes on a standard PC. Additionally, the resulting model can easily be implemented as a physical electric equivalent circuit.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Maxwell’s equations; Finite integration technique (FIT); Model order reduction (MOR); Two-step Lanczos (TSL); Equivalent circuit extraction

1. Introduction

For the analysis of microwave devices in computational electrodynamics, volume discretizing methods such as the Finite Difference Time Domain (FDTD) method, the Finite Integration Technique (FIT) or the Finite Element (FE) method are widely used and accepted tools. Hereby typically an electromagnetic component is considered that is excited by one or more so called ports—certain parts of the structure where external quantities like voltages, currents or waves are introduced to the structure. Although it is possible to calculate the electric or magnetic fields over the entire volume, for many engineering tasks only the frequency dependent transfer behavior—the ratio between the exciting input and the resulting output signals—is of interest. Depending on the system formulation, the above mentioned algorithms generate large linear systems of equations with orders of about three or even six times the number of volume cells that have to be solved to compute both the field distribution as well as the transfer function.

In case of FDTD and FIT explicit time domain simulations are very powerful tools to solve these problems efficiently, requiring only one matrix–vector multiplication per additional time-step. Frequency domain results such as transfer functions can be generated from the time signals via a Fast Fourier Transform (FFT) or similar techniques. Unfortunately, to ensure stability of the method the size of the time-step is limited by the Courant–Friedrich–Levi criterion, which in a consequence limits the method to “high frequency” problems, considering only wavelengths that are at least in the range of the spatial size of the component. For lower frequencies implicit time domain schemes can be applied, with the disadvantage of the need to solve a linear system per time-step. But even if the stability criterion is fulfilled, time domain methods suffer from long settling times to reach steady state for highly resonant structures like filters.

Standard frequency domain methods in contrast suffer from the computational effort of the solution of a very large linear system—often up to millions of unknowns—for each considered frequency point. Especially finding sharp resonances in the transfer behavior may require a high number of frequency values. The system at each frequency is typically solved using an iterative method building up a Krylov subspace $K(A, b)$ with the system matrix $A$ and the excitation vector $b$. Since
the system matrices at different frequencies in FIT only differ in a diagonal term and Krylov subspaces are known to be invariant to a constant diagonal shift of the matrix \( K(A - sI, b) = K(A, b) \), the same subspace has basically to be built up many times, causing the high computational cost (at the moment neglecting advanced preconditioners). Unfortunately, even for mid-size problems the subspace usually cannot be held in memory. But if only the transfer behavior is of interest, a reduced order model can be generated by projecting once the whole system matrix onto this Krylov subspace, which can be done implicitly by the Lanczos algorithm [1] without storing all vectors—again with only matrix–vector multiplications. The required subspace size can be assumed to be the maximum iteration number of all single solutions described above and remains at first the uncertainty of the method. Models of this type are called partial realizations. The basic knowledge of connections between the Lanczos algorithm and these reduced models as well as their moment matching properties around infinity frequency have already been known for about 20 years now [2] and were adopted to electromagnetic problems in [3,4] and to FIT in [5].

If the same number of iterations is applied to a Krylov based eigenvalue solver, typically all eigenvalues in the considered frequency range are also found to be well approximated. This is a further confirmation that the essential spectral information is contained in the subspace and offers a chance to find a stop criterion for the previously described projection based on eigenvalues [6]. Further on, it encourages the idea of utilizing directly the eigenvectors—belonging to eigenvalues in the considered frequency range—instead of the much larger Krylov space as projection matrix, a method known as modal expansion [7]. The smaller matrix can easily be kept in memory and therefore also allows to calculate full field solutions. However, the complete neglect of higher order modes may lead to an offset error, which has to be compensated by a suitable correction approach. With a combination of methods from eigenvector computation such as Chebychev acceleration polynomials and Krylov subspaces an almost smooth transition between both approaches can be found, combining the advantages of both.

An at first glance completely different approach of model order reduction is given by moment matching around one or more expansion points in the frequency domain, which can result in a Padé approximation of the transfer behavior. First explicit moment matching methods were used in control theory already about 30 years ago [8], in the electromagnetic community it become popular with the Asymptotic Waveform Evaluation (AWE) [9]. A much more efficient implicit implementation of Padé approximations again leads to Krylov subspaces, this time applied to the inverted system, a connection found in the late 1980s, made popular around 10 years ago as Padé Via Lanczos (PVL) [10,11], and applied to electromagnetics in [12]. This algorithm as well as modifications like the passivity preserving PRIMA [13] generate models of very low order and have proven to be highly reliable techniques, however, at the cost of inverting the system matrix or alternatively solving again a large linear system several times.
In combination with resonant electromagnetic devices the reduced order models of the discretized structure representations have two main applications:

- **Fast Frequency Sweep**: The transfer behavior of the structure shall be calculated quickly, besides accuracy the computational time is of major importance.
- **Equivalent circuit extraction**: It is desired to obtain a model as small as possible, preserving both stability and passivity, computational time is of secondary importance.

This paper will present a more detailed description and a comparison of the methods mentioned above in view of both of these application fields. Especially the successive application of partial realization as a pre-reduction and moment matching for the final reduction leads to a very efficient algorithm called Two Step Lanczos (TSL) [6]. Applied to lossless systems (in most cases resonant structures can be assumed to be lossless) it is advantageous to use the “curl–curl” formulation of the system—to be explained later—instead of the more common linear state-space. This further increases the efficiency of the TSL algorithm and keeps the system stable, passive, and real throughout the reduction, while matching the maximum number of moments. In presence of small losses caused either by a small conductivity in dielectric materials or by a high (but not infinite) conductivity in metals, the algorithm can easily be extended by also projecting the matrices representing the losses onto the Krylov subspaces generated from the lossless system. This extension does not match moments precisely anymore, however, due to the small influence of the losses the overall approximation error is still very small, while the method preserves TSL advantages like stability, passivity and realness.

With fulfillment of these conditions the reduced model can directly be used to generate an equivalent circuit, which only contains linear network elements and describes the port behavior of the device within the desired frequency range. This substantially simplifies the simulation of coupled circuit-electromagnetic problems. Compared to former approaches such as the parameter optimization of predefined network descriptions like e.g. coupled transmission line models [14] this technique is much more general and no knowledge on the structure of the electromagnetic device is needed in advance. For lossless devices there is a direct physical interpretation of the network; in case of lossy structures negative resistors may appear, but it will be shown that the overall passivity is still guaranteed.

The paper is organized as follows: In Section 2 the basic properties of the Finite Integration Technique (FIT) are derived and two related state-space formulations are proposed. It is shown that a specific property of FIT—namely the diagonal form of its material operators—is an important advantage for the efficient implementation of MOR methods. Section 3 reviews the various reduction strategies, collects known theory and highlights their close relations, whereby the efficiency of the Two-Step Lanczos algorithm is shown. Section 4 derives the direct usage of the reduced model as equivalent circuit also considering its physical interpretation. Finally, the algorithms
are compared by means of three filter structures in Section 5 before ending with some concluding words.

2. Finite integration technique

The model order reduction process is based on a state-space representation of the device obtained by discretization via the Finite Integration Technique (FIT) [15]. The FI-Technique supplies a consistent transformation of the integral form of Maxwell’s Equations into a set of matrix equations, that maintains the physical properties of the computed field and leads to a unique solution.

Maxwell’s equations are hereby discretized on a pair of spatially interlaced dual grids using integral degrees of freedom: electric voltages  are allocated on the edges between pairs of two grid points and the magnetic fluxes  through facets of the primary grid  . The magnetic grid voltages  are similarly defined on the edges and the electric grid fluxes  and currents  on the facets of the dual grid  . With the discrete curl-operators  and  for the primary and the dual grid respectively, Faraday’s and Ampère’s Grid Equations become:

\[
C \hat{e} = -\frac{d}{dt} \hat{b}, \quad \tilde{C} \hat{h} = \frac{d}{dt} \hat{d} + \hat{j}.
\]

(1a,b)

Hereby \( \tilde{C} = C^T \) holds, with both matrices sparse and only containing +1 or −1. This representation is still an exact representation on the grid and does not contain any approximation errors.

The unavoidable approximations of any numerical procedure are introduced by the discrete analogues of the coupling between voltages and fluxes, for the linear case represented by the material matrices:

\[
\hat{d} = M_\varepsilon \hat{e}, \quad \hat{b} = M_\mu \hat{h}, \quad \hat{j} = M_\varepsilon \hat{e} + M_\kappa \hat{e} + j_s.
\]

(2a,b,c)

For dual orthogonal grid systems these matrices are diagonal and therefore easy to invert,  and  are in addition positive definite,  is positive semidefinite. After application of the Laplace transformation Maxwell’s Grid Equations can be rewritten as

\[
C \hat{e} = -sM_\mu \hat{h}, \quad C^T \hat{h} = sM_\varepsilon \hat{e} + M_\kappa \hat{e} + j_s.
\]

(3a,b)

For Cartesian grids the number of unknowns can be approximated by  \( n_e, n_h \approx 3N_P \) (\( N_P = \) number of mesh points).

The excitation of the system is introduced by the source current vector  directly on the dual grid. For macromodeling of port-driven devices it is preferable to use the port currents  as input (with only one entry per port) and the port voltages  as output values. From  and  the transfer behavior in form of impedance and scattering matrices can easily be calculated [16]. The mapping of  on  and of  on  is defined as
\[- \mathbf{j}_x = \mathbf{R} \mathbf{i}, \quad \mathbf{u} = \mathbf{L} \bar{\mathbf{e}}. \quad (4a,b)\]

The matrices \( \mathbf{R} \) and \( \mathbf{L} \) contain the field patterns \( \mathbf{E}_{2D} \) and \( \mathbf{H}_{2D} \) of the port planes and are obtained by the solution of a computationally cheap 2D eigenvalue problem. \( \mathbf{R} \) imposes a current pattern on the port plane which corresponds to \( \mathbf{n} \times \mathbf{H}_{2D} \), and \( \mathbf{L} \) extracts the generalized voltage \( \mathbf{u} \) from the fields at the ports, based on the \( \mathbf{E}_{2D} \times \mathbf{H}_{2D} \) orthogonality of waveguide modes. It is obvious that, by proper normalization, the symmetry relation \( \mathbf{R} = \mathbf{L}^T \) can be achieved. In case of TEM ports (with transversal electric and magnetic field components) the matrices \( \mathbf{R} \) and \( \mathbf{L} \) are constant, but in general they are frequency dependent. In many cases, e.g., for microstrip ports, it is acceptable that \( \mathbf{R} \) and \( \mathbf{L} \) are calculated for the mid-frequency and—as an approximation—assumed to be constant for the considered frequency range. For a device with \( m \) external ports the matrix dimension of \( \mathbf{R} \) is \( n_e \times m \).

An alternative approach to define the matrices \( \mathbf{R} \) and \( \mathbf{L} \) is based on the theory of “electromagnetic circuit elements” and described in [5].

The combination of Eqs. (3) and (4) in one system containing all entries of \( \bar{\mathbf{e}} \) and \( \bar{\mathbf{h}} \) as unknown “states” results in a linear system of dimension \( n = n_e + n_h \) in “classical” state-space form, see also [6]:

\[
\begin{bmatrix}
\varepsilon_0 & 0 \\
0 & \mu_0
\end{bmatrix} \begin{bmatrix}
\bar{\mathbf{e}} \\
\bar{\mathbf{h}}
\end{bmatrix} = - \begin{bmatrix}
\mathbf{C}^T & 0 \\
0 & \mu_0
\end{bmatrix} \begin{bmatrix}
\bar{\mathbf{e}} \\
\bar{\mathbf{h}}
\end{bmatrix} + \begin{bmatrix}
\mathbf{R}_0 \\
0
\end{bmatrix} \mathbf{i},
\]

\[
\begin{bmatrix}
\mathbf{R}_0^T \\
0
\end{bmatrix} \begin{bmatrix}
\bar{\mathbf{e}} \\
\bar{\mathbf{h}}
\end{bmatrix} = \mathbf{u}.
\]

Eliminating \( \bar{\mathbf{h}} \) in this equations an alternative system can be obtained, which is not linear in \( s \), but only half as large:

\[
\begin{align*}
\varepsilon_0 \mu_0 \bar{\mathbf{e}} &= - \mathbf{C}^T \mathbf{M}_\mu^{-1} \mathbf{C} \bar{\mathbf{e}} - s \mathbf{M}_\varepsilon \bar{\mathbf{e}} + s \mathbf{R} \mathbf{i}, \\
\mathbf{u} &= \mathbf{R}_0^T \bar{\mathbf{e}}.
\end{align*}
\]

With a change of variables \( \mathbf{x} = \mathbf{M}_\varepsilon^{-1/2} \bar{\mathbf{e}} \) (with \( \mathbf{M}_\varepsilon^{1/2} \mathbf{M}_\varepsilon^{-1/2} = \mathbf{M}_\varepsilon \)) this system can be transformed to a symmetric one, the so-called curl–curl system, with \( n = n_e \) unknowns:

\[
\begin{align*}
\varepsilon_0 \mu_0 \mathbf{x} &= - \mathbf{A} \mathbf{x} - s \mathbf{K} \mathbf{x} + s \mathbf{B} \mathbf{i}, \\
\mathbf{u} &= \mathbf{B}^T \mathbf{x},
\end{align*}
\]

with

\[
\begin{align*}
\mathbf{A} &= \mathbf{M}_\varepsilon^{-1/2} \mathbf{C}^T \mathbf{M}_\mu^{-1} \mathbf{C} \mathbf{M}_\varepsilon^{-1/2}, \\
\mathbf{K} &= \mathbf{M}_\varepsilon^{-1} \mathbf{M}_\varepsilon \quad \text{and} \\
\mathbf{B} &= \mathbf{M}_\varepsilon^{-1/2} \mathbf{R} = (\mathbf{L} \mathbf{M}_\varepsilon^{-1/2})^T.
\end{align*}
\]

Note that this symmetrization is trivial in FIT—having diagonal material matrices with positive entries—but may be computationally expensive in standard FE schemes,
where the corresponding mass matrices are generally non-diagonal (but still symmet-
tric positive definite). This property of FIT is essential for the efficiency of the partial
realization which will be explained below.

In FIT the matrix \( A \) is sparse with—for standard Cartesian grids—a block banded
structure with only 13 non-zero entries per line. \( K \) is a diagonal matrix, both \( A \) and
\( K \) are positive semidefinite.

The system’s (normalized) impedance matrix \( Z(s) \), satisfying \( u = Z(s)i \), finally
results as

\[
Z(s) = sB^T(s^2I + sK + A)^{-1}B,
\]

where \( I \) is the identity matrix. Generally a matrix of the form \( (s^2E + sK + A) \) is
known to have its eigenvalues only in the left half of the Laplace plane if \( E \) is positive
definite and \( A \) and \( K \) are positive semidefinite [17]. In the time domain these eigen-
values with negative real parts are connected to oscillating signals with the magnitude
fading with time, representing a stable system. With the same prerequisite the passivity
of the impedance matrix can be proven [18]. Passivity is important for coupled simula-
tions, since only a connection of two passive systems is guaranteed to remain passive
and stable, while a connection of two ‘only’ stable systems can lead to an unstable
system. The system given in (7) or (9), respectively, therefore is—as expected—both
stable and passive.

For many resonant electrodynamic systems the values of \( K \) are caused by parasitic
losses in dielectric material and are relatively small. A measure for the losses is the
loss tangent \( \tan \delta = \kappa/\omega \varepsilon \), which for typical dielectric materials is in the range
of \( 10^{-2} \) to \( 10^{-4} \). For lossy metal the exact modeling would result in a very fine grid,
therefore, often analytical models are chosen modeling the skin depth as an impedance
wall. This additionally introduces a term to (9) proportional to \( \sqrt{s} \).

3. Model order reduction

The purpose of Model Order Reduction (MOR) is to generate a system of much
lower order compared to the original one, that approximates the transfer function (9)
over a predefined range of frequencies. Basically all modern MOR techniques can be
interpreted as projection of the original \( n \times n \) system onto two rectangular matrices
\( V \) and \( W \), both of dimension \( n \times p \) with \( p \ll n \), leading for the curl–curl formula-
tion to

\[
Z_{red}(s) = sB^T(s^2W^TV + sW^TKV + W^TAV)^{-1}W^TB.
\]

Generally the choice of the matrices \( V \) and \( W \) is free, but in order to get reliable
and low order reduced models the subspaces spanned by these matrices have to be
associated to certain properties of the frequency range of interest. In the simplest case
already a number of field solutions at different frequencies can lead to a satisfying
model, however, the approximation quality depends highly on the choice of these
solutions. The well-known balanced truncation [19] method utilizes Hankel singular vectors belonging to the largest singular values of the system, which requires the solution of the Lyapunov equation. Since the latter is generally of order $O(n^3)$ this method is unfortunately inapplicable for large scale systems.

The subspaces considered for $V$ and $W$ in this paper are a number of eigenvectors of the original system, leading to modal expansion, or Krylov-subspaces of the system matrix and its inverse, representing moment matching around zero and infinite frequency, respectively.

Though the projection (10) offers the largest flexibility in the choice of subspace attributes, it does generally not maintain the stability and passivity properties of the original system, but may lead to unstable models. This disadvantage can be overcome by restricting the projection to a symmetric and real one, however, taking a larger reduced model into account:

$$V = W, \quad V \in \mathbb{R}^{n \times p}. \quad (11)$$

For reasons of numerical stability of the algorithms it is additionally desirable to keep the subspaces orthogonal $V^T V = I$ or biorthogonal $W^T V = I$ respectively.

3.1. Partial realization

At first, only lossless systems ($M_\kappa = 0$) are considered and the block Krylov-subspace $K_p(A, B) = \text{span} \{B, AB, A^2B, \ldots\}$ connected to the symmetric system matrix $A$ and the excitation matrix $B$ is utilized for projection.

The subspace is created using the symmetric block Lanczos algorithm for a square matrix $A$ and a right hand side matrix $B$. In its general formulation it iteratively generates a banded matrix $T_p$ with $m$ ($m = \text{number of ports}$) bands on both sides next to the diagonal, where the dimension $p$ of $T_p$ increases by one in each iteration step. Additionally, a set of vectors $V_p = [v_1 \ldots v_p]$ is obtained building an orthonormal basis for the desired Krylov-subspace with the relations:

$$AV_p = V_p T_p + [0 \ldots 0, \hat{v}_1 \ldots \hat{v}_m], \quad V_p^T V_p = I, \quad V_p^T [\hat{v}_1 \ldots \hat{v}_m] = 0, \quad (12a,b,c)$$

where $\hat{v}_i$ are help vectors as in [11]. This results in the projection formulation:

$$V_p^T A V_p = T_p. \quad (13a)$$

A quasi-symmetric band formulation of the Block Lanczos algorithm (see Algorithm 1) was implemented that orthogonalizes the new generated vector against the previous $2m$ ones. Under full exploitation of the symmetry only $m$ orthogonalizing steps could be sufficient, however, for high iteration numbers $p > 500$ the quasi-symmetric version has proven to be numerically much more robust against round-off errors. A similar effect was already observed in [20], for further details to various Lanczos implementations see [1,10,11].
Algorithm 1:

Given: $A, B = (b_1, \ldots, b_m), p$

For $i = 1:p + m$ create Krylov and help vectors

If $i \leq m$

$v_i = b_i$

Else

$v_i = Av_{i-m}$

End

$j_0 = \max(1, i - 2m)$ orthonormalize

For $j = j_0 : i - 1$

$t_{j,i} = v_j^T v_i$

$v_i = v_i - t_{j,i}v_j$

End

$t_{i,i} = \|v_i\|_2$

$v_i = v_i / t_{i,i}$

End;

$B_p = [t_{1..p,1..m}]$ final Lanczos matrices

$T_p = [t_{1..p,m+1..m+p}]$

$V_p = [v_{1..p}]$, $(\tilde{V} = [v_{p+1..p+m}])$

The (quasi-)symmetric Lanczos algorithm is therewith the extension of the Arnoldi algorithm [1] to symmetric matrices with the important advantage that each generated vector $v_n$ only has to be orthogonalized with respect to the previous $2m$ vectors instead of all previous ones. This keeps the related numerical effort linear with the number of iterations, in contrast to the quadratic dependence in case of Arnoldi. Both algorithms span the same subspace and can basically be exchanged. Due to its computational advantages, in the following only the Lanczos algorithm is used.

As a result of the projection onto the Krylov subspace including $B_p = V_p^T B$, a reduced system is obtained from (7):

$$s^2z = -T_p z + sB_p i,$$

$$u = B_p^T z.$$  \hspace{1cm} (14)

The corresponding transfer matrix becomes

$$Z_p(s) = sB_p^T (s^2 I + T_p)^{-1} B_p.$$  \hspace{1cm} (15)

With $T_p = V_p^T A V_p$ the system remains both symmetric and positive definite and is therefore guaranteed passive and stable. It should be noted that only a very small
number \((2m)\) of the potentially large number of vectors in \(V_p\) actually has to be stored.

From (12a,c) it follows that:
\[
B^T A^k B = B^T V_p T_p^k V_p^T B = B_p^T T_p^k B_p \quad \text{for } k = 0, 1, \ldots, l, \ l = \text{floor}(p/m).
\]

The values \(B^T A^k B\) are the coefficients of the geometric series of \(Z(s)\) and are called moments around infinity or Markov parameters:

\[
Z(s) = \sum_{k=0}^{\infty} B \Lambda^k B \frac{1}{s^k}.
\]

The approximated transfer function \(Z_p(s)\) and \(Z(s)\) therewith share the same \(l\) moments around infinity. Or, in a time domain representation the Markov parameters represent Taylor coefficients around \(t = 0\) of the system’s impulse response. Therefore, also the impulse responses of original and approximated system are identical in \(l\) Taylor moments [23].

Additionally, the series of matrices \(T_p\) is known to generate increasingly good approximations of the subspace spanned by the eigenvectors of the original matrix \(A\), where the eigenvectors belonging to the largest and well “isolated” eigenvalues are “found” first. The number \(p\) of required iterations \((p\ is 500–5000 for typical problems)\) therefore can be determined during the computation by an eigenvalue criterion. It is based on the observation that the approximation in terms of transfer functions within a certain band of frequencies is sufficient as soon as the eigenvalues lying in the same band are good approximations of the eigenvalues of the initial system. Since the original eigenvalues \(\lambda_i\) are expensive to calculate, the eigenvalues \(\lambda_{i,p}\) of \(T_p\) are computed in pre-defined intervals. As soon as their number and numerical values coincide within a specified tolerance with the ones previously calculated, the iteration is terminated. To avoid a large computational overhead introduced by the eigenvalue calculation, the procedure can be optimized by not checking the eigenvalues up to a predefined number of iterations and by using decreasing iteration interval sizes with increasing accuracy.

The relation between the eigenvalue error \(\delta_E\), the eigenvalue difference \(\delta_D\) between two iteration intervals, both calculated in the frequency range of interest and the transfer matrix error \(\delta_Z\) is demonstrated in Fig. 1 for a small example with 2400 unknowns. All three error curves

\[
\delta_Z = \frac{\|Z_p - Z\|}{\|Z\|}, \quad \delta_E = \frac{\|\hat{\lambda}_{i,p} - \hat{\lambda}_i\|}{\|\hat{\lambda}_i\|}, \quad \delta_D = \frac{\|\hat{\lambda}_{i,p+\Delta p} - \hat{\lambda}_{i,p}\|}{\|\hat{\lambda}_{i,p}\|}
\]

show proportional behavior over the number of iterations \(p\). The distance \(\Delta p\) between two eigenvalue computations in the relative error calculation was chosen to be 10 iterations, the solver accuracy of the transfer function was chosen to be \(10^{-10}\), for the eigensolver \(10^{-12}\). For a relative eigenvalue deviation of \(10^{-7}\) the algorithm would have stopped after 330 iterations.
Of course, the curves in Fig. 1 have been recorded for a specific example, and there is no rigorous proof for the close relation between the eigenvalue error and the accuracy of the approximated transfer functions. However, numerical experiments with a large range of different kinds of applications have shown that the eigenvalue-based stop criterion is not only impressively robust, but also computationally efficient in the sense that the required number of iterations is accurately estimated with very low additional cost.

3.2. Partial realization for weakly lossy systems

If losses cannot be neglected completely for the investigated device the full equation (7) including the matrix $K$ has to be considered in the order reduction process. Since the Lanczos algorithm cannot be applied to more than one matrix simultaneously the described reduction algorithm in curl–curl formulation cannot be extended in a precise way to the matrices $A$ and $K$. But for this case the larger system (5) can always be used for reduction.

To maintain the advantages of the curl–curl formulation and the fast matrix–vector multiplications, an approximation can be introduced: The algorithm is hereby started with the matrix $(A + s_0K)$ followed by a projection of the system (7) on the obtained subspace $V_p$. However, with an optimally chosen (and therefore imaginary) $s_0$ in this case the reduced system becomes complex, too. For the calculation of a “fast frequency sweep” this approach is reasonable, but for the circuit extraction as proposed in the following section it is desirable to keep the system real and passive.

Nevertheless, most resonant systems only contain relatively small losses with minor influence on the field pattern inside the structure. Therefore, it is often sufficient to apply—as before—the partial realization algorithm only to the real matrix $A$ as described in Section 3.1. The matrix $K$ is also reduced by the projection on the matrices $V_p$. This approach is comparable to the classical Power-Loss-Method in continuous field theory. The reduced system becomes
\[
Z_{pt}(s) = sB_p^T(s^2I + sV_p^TKV_p + T_p)^{-1}B_p.
\] (19)

As can easily be seen the system remains real and preserves stability and passivity. It should be mentioned that the described method does not match moments in an exact way anymore. Experience shows, however, that it supplies good results for losses up to \(\tan \delta = 0.1\), which is much higher than the values of typical dielectric materials (see Fig. 4 for some results).

Unfortunately, one important disadvantage of this approach is the fact that the matrix \(V_p\) with \(n \times p\) elements has to be stored completely in the first step. If \(V_p\) gets too large for the memory, there are some options to still generate \(K_p = V_p^TKV_p\):

- The vectors \(V_p\) are not needed at once within the algorithm. They might be written periodically to hard disc.
- Often only some domains of the considered structure are lossy while others remain lossless. In this case the diagonal matrix \(K\) has a reduced number of entries and memory can be saved calculating the sparser \(K' = K^{1/2}V_p\) iteratively within the Lanczos computation and finally \(K_p = K'^TK'\).
- The matrix \(K_p\) shows some dominance for the entries close to the diagonal. If only a limited number of vectors \(V\) are kept in memory and the calculation of \(K_p\) is done simultaneously in the Lanczos iteration, only these inner entries are computed. The error is small for typical applications but it is difficult to estimate.

### 3.3. Convergence acceleration for partial realizations

Since the similarity between order reduction via partial realizations and iterative Krylov eigenvector computation is obvious, it is reasonable to also apply techniques known from eigenvalue computations to Model Order Reduction. Main purpose in this context is the further reduction of at the best the total computational cost or at least of the size of the Krylov subspace, to be able to keep it in memory. This is advantageous for the projection of weakly lossy systems—just described above, but also for the possibility to compute approximations of the full field solution of the original system at arbitrary frequencies in a post-processing step. The reduced state vector \(z\) (Eq. (14)) does not have any direct physical meaning, however, the electrical field described by the vector \(\bar{e}\) from Eq. (6), is connected to the solution vector \(z_s\) of the reduced model at a certain angular frequency \(s\) by the relation

\[
\bar{e}_s = M^{-1/2}Vz_s.
\]

One possibility is to use the Implicitly Restarted Arnoldi (IRA) or Implicitly Restarted Lanczos (IRL) algorithms to build up the projection matrix \(V_{p-IRA}\) with a predefined subspace size. This technique is described in detail in [3].

A different approach is given by application of Chebychev acceleration polynomials before starting the Krylov iteration. In a similar formulation this technique in known as explicit restart [21]. However, in connection with Model Order Reduction
the polynomial is applied only once to the excitation matrix \( B \) before starting the Lanczos algorithm. Hereby each excitation vector \( b_m \) can be expressed as a superposition of all \( n \) eigenvectors \( x_\eta \) (\( \eta = 1 \ldots n \)) of \( A \):

\[
b_m = \sum_{\eta=1}^{n} \alpha_\eta x_\eta.
\] (20)

Application of a matrix Chebychev polynomial \( P_K(A) \) with the roots \( r_k \) results in

\[
b_{\text{pol}} = \left\{ \prod_{k=1}^{K} (A - r_k I) \right\} b_m = \sum_{\eta=1}^{n} \alpha_\eta \left\{ \prod_{k=1}^{K} (\lambda_\eta - r_k) \right\} x_\eta = \sum_{\eta=1}^{n} \alpha_\eta P_K(\lambda_\eta) x_\eta.
\] (21)

With the right choice of the polynomial zeros \( r_k [22] \) this algorithm “accelerates” the spectral properties of the excitation that lie in the frequency range of interest, while suppressing all parts of higher modes. In a consequence, the Lanczos algorithm of the partial realization—started with \( b_{\text{pol}} \)—reaches the same accuracy of the stop criterion with a smaller number of iterations. The larger the polynomial size \( K \) is chosen, the less iteration steps are needed in the succeeding partial realization, forming a transition from partial realization to pure modal expansion as described in the following section.

Experience shows, that the overall computational cost remains almost constant for various polynomial sizes. The time saved in the partial realization has to be invested in applying the polynomials to the starting vectors. However, the projection matrix gets smaller and can be saved even for large systems. Unfortunately, with strongly suppressing the higher modes a new offset error occurs in the transfer function. It is due to the missing of their contribution even at low frequencies which is small for the single mode, but belongs to many of those. This error is well-known from modal expansion and in the following section an easy scheme for compensation is presented.

### 3.4. Corrected modal expansion

The modal expansion method can be considered as a variant of the partial realization approach above, where a set of previously calculated eigenvectors \( x_\eta \) of the system are explicitly used instead of the Krylov vectors. Concentrating again on the lossless case and the curl–curl-formulation, the complete set of all eigenvectors \( x_\eta \) (\( \eta = 1 \ldots n_e \)) build an orthonormal basis of the system matrix (or can at least be orthonormalized):

\[
A = XAX^T \quad \text{with} \quad X^{-1} = X^T \quad \text{and} \quad \Lambda = \text{diag}(\lambda_\eta) = \text{diag}(-s_\eta^2).
\] (22)

This leads to a simplified formulation for the impedance matrix, where the inverse of a diagonal matrix is trivial and a simple postprocessing step (after having provided the eigensolutions):
\[
Z(s) = sB^T(s^2I + XAX^T)^{-1}B \\
= s(X^TB)^T(s^2I + \Lambda)^{-1}(X^TB) \\
= (X^TB)^T \text{diag} \left( \frac{s}{s^2 - s^2_\eta} \right)(X^TB). \tag{23}
\]

Thus, one entry \(Z_{ij}(s)\) of \(Z(s)\) is given by the summation over the contributions of all \(n\) eigenvectors:

\[
Z_{ij}(s) = \sum_{\eta=1}^{n} (x^T_{\eta}b_i) \cdot (x^T_{\eta}b_j) \cdot \frac{s}{s^2 - s^2_\eta} = \sum_{\eta=1}^{n} Z^n_{ij}(s). \tag{24}
\]

Since the complete set of eigensolutions is not available for practical problems, the modal approach uses an approximation of this formula by a truncated sum including a number \(p\) of eigenvectors and a correction term

\[
\hat{Z}_{ij}(s) = \sum_{\eta=1}^{p} Z^n_{ij}(s) + Z^\text{corr}_{ij}(s) \quad \text{with} \quad Z^\text{corr}_{ij}(s) \approx \sum_{\eta=p+1}^{n} Z^n_{ij}(s). \tag{25}
\]

The quality of this approximation depends on the set of eigensolutions chosen and the correction term, and their relation to the frequency range of interest. From the frequency dependence of the single contributions \(Z_{ij}(s)\) in (24) it is reasonable to consider at least all modes with eigenfrequencies within this range, \(|s_{\min}| \leq |s_i| \leq |s_{\max}|\), which is in close analogy to the stop criterion for the partial realization proposed above.

Experience shows, however, that the projection on these eigenvectors is not sufficient to yield an accurate approximation of the impedance, and it is mandatory to find a numerically cheap, but efficient implementation of the correction term.

A quite sophisticated derivation of such a correction term was proposed in [24], where a complementary eigenproblem (with interchanged boundary conditions at the ports) is solved to obtain an approximation \(\hat{Y}(s)\) of the admittance matrix \(Y(s) = Z^{-1}(s)\). The relation between poles and zeros of \(\hat{Y}(s)\) and \(\hat{Z}(s)\) finally leads to a highly accurate approximation of \(Z^\text{corr}_{ij}(s)\).

A simpler approach has been introduced in [25] and will be used in this paper. Here, the correction term is assumed to have an approximately linear dependence in \(s\) which follows from (24) and \(\frac{s}{s^2 - s^2_\eta} \approx \frac{s}{s - s_\eta}\) for \(|s| < |s_p|\) (for the contribution of non-considered modes). Thus, the cumulative contribution of all higher modes can be estimated by a linear extrapolation of the results from one (or a small number of) exact solutions of (9). Here, a good start solution for iterative solvers can be provided by the truncated sum in (25), and thus the additional numerical effort is typically much less than the cost of the preceding solution of the eigenproblem.

Numerical experience shows that this simple correction approach yields highly accurate transfer functions, where, however, the numerical cost (strongly depending on the type and parameter set of the eigensolver) is typically higher than with the
partial realization. Additionally, a robust stop criterion for the number of modes and the choice of correction frequencies has not been established yet.

On the other hand, the method yields very small reduced models, in fact they can be considered to have the minimum size possible for a given frequency range. As a consequence, all projection vectors can typically be held in memory, and field solutions at arbitrary frequencies can easily be obtained. As another more practical advantage, the computed eigenmodes and eigenvalues—corresponding to electric field solutions and resonance frequencies, respectively—may give important physical insights in the analyzed structure, which is often an important benefit e.g. in the design of microwave filter devices.

Some ideas to extend the method to lossy problems have been sketched in [26], or follow directly from the considerations above.

3.5. Moment matching

Shifting the moment extraction point from infinity frequency into the frequency range of interest is certainly very advantageous concerning the resulting model size. Therefore, an expansion point $s_0^2$ can be introduced before the system has to be inverted:

$$Z(s) = sB^T(I + (s^2 + s_0^2)A')^{-1}B',$$

with $A' = (A - s_0^2I)^{-1}$, $B' = A'B$. (27a,b)

As experimental results show, in most cases the best choice for $s_0$ is just the mid-frequency of the interesting frequency band multiplied by $j2\pi$. This fact can be explained by the symmetric character of the moment matching on both sides of the extraction point. It should be noticed that though $s_0$ is an imaginary value, $s_0^2$ and therewith the whole reduced system remain real, which is an important advantage for the circuit extraction in a later stage.

Although the matrix $A'$ is still symmetric, the system becomes unsymmetric with $B \neq B'$. Generally, to obtain a $q$th order Padé approximation (matching at least $2q/m$ moments) of the transfer function of an unsymmetric system both matrices $V_q'$ and $W_q'$ related to the Krylov subspaces $K_q(A', B')$ and $K_q(A'^T, B)$, respectively, have to be considered. One possibility would be the usage of the unsymmetric version of the Lanczos algorithm with $W_q'^T A' V_q' = T_q$ as in [10]. After $q$ iterations (typically $q < 50$) the reduced model follows from a two-sided projection on $V_q'$ and $W_q'$:

$$Z_q(s) = sB_p^T V_q' (W_q'^T V_q' + (s^2 + s_0^2)T_q)^{-1} W_q'^T B_p'.$$

(28)

However, due to the symmetry of $A'$ and from the knowledge of $B' = A'B$ the relation $W_q = A'^{-1}V_q'$ holds for all $q$. Finally, Eq. (28) can be rewritten as

$$Z_q(s) = sB_p^T V_q' (s^2 V_q'^T V_q' + V_q'^T AV_q')^{-1} V_q'^T B_p.$$

(29)
From (29) it gets obvious that in this case the matrix $W'_q$ is actually dispensable in the projection and a Padé approximation of system (26) can be obtained by considering only $V'_q$ [11]. Applying a standard QR factorization $V'_q = V_q R$ a unitary matrix $V_q^T V_q = I$ can be found, that spans the same subspace as $V'_q$. This further simplifies (29):

$$Z_q(s) = s B_p^T V_q (s^2 I + V_q^T A V_q)^{-1} V_q^T B_p.$$  (30)

Much easier, the same matrix $V_q$ can be directly generated by the symmetric Lanczos algorithm from the matrices $A'$ and $B'_p$. This halves the computational effort compared to (28), while still the same number of moments are matched. Starting with a symmetric and positive semidefinite system, also the final model can be guaranteed to be stable and passive.

An additional advantage of formulation (30) is the simple implementation of a multipoint approximation (with multiple expansion frequencies) [27]. Instead of taking all vectors $V$ from only one expansion point, a matrix $V_q = [V_1 \ V_2 \ \cdots \ V_n]$ is used, where each block $V_1, V_2, \ldots$ contains a number $r < q$ of vectors obtained with different expansion frequencies $s_{01}, s_{02}, \ldots$. However, numerical experience shows that in most cases multipoint approximation does not significantly reduce the model size.

The projected model of the inverted system can be shown to represent a Padé approximation. For linear state-space systems and one expansion frequency this algorithm became famous under the name Padé Via Lanczos (PVL) [10]. A detailed proof of the moment matching properties of the formulations (28) and (30) around the expansion point $s_0 = 0$ is found in [11].

Again, there is also a close relation to eigenvalue computation. Since the dynamic poles with the lowest frequencies are usually the most interesting ones in technical applications, they are related to the largest eigenvalues in the shifted and inverted system $(A - s_0^2 I)^{-1}$. The Lanczos algorithm applied to the matrix $(A - s_0^2 I)^{-1}$ therefore generates a very good approximation of the original system within very few iterations.

Therefore, the stop criterion is very similar to the one in the partial realization approach. The relation of the (absolute) eigenvalue error and the transfer function error is shown in Fig. 2. It is obvious that the number of iterations $q$ to reach the error of $10^{-7}$ is drastically smaller compared to the partial realization. In this example the iteration would have been terminated after only eight iterations. The staircase character of the curve can be explained by the block formulation of the Lanczos algorithm, where $m = 2$ ports were used here.

However, the inversion or factorization of the matrix $(A - s_0^2 I)$, which in typical cases in electromagnetics can be of order up to millions of unknowns, requires a very high computational and memory storage cost. Alternatively, a very large linear system has to be solved for each iteration.

In the lossy case, the unsymmetric variant (28) of the moment matching approach in connection with the linear system can be employed. To maintain the advantages of the
curl–curl formulation, an alternative MOR algorithm exists with the Well Conditioned Asymptotic Waveform Evaluation (WCAWE) [28], that matches moments of general matrix equations that are polynomials in $s$, unfortunately again based on the need of iterative solutions. If the losses are weak, again a projection can be used similar to Section 3.2, Eq. (19).

3.6. Two-step Lanczos

Although the small size and the accuracy of reduced models via moment matching techniques is quite impressive, the numerical cost for large original systems is often too high for a really “fast” frequency sweep. On the other side, partial realizations of FIT systems are often computed within seconds or at least minutes, but the resulting models are not really what is called “low” order but rather mid-size. Even for a frequency sweep, but especially for the usage of the model as a macromodel or equivalent circuit for coupled simulations, the minimum model size is very desirable.

This motivates the idea of combining both algorithms in a successive manner. In a first step a mid-size model is generated by the application of a partial realization on the original matrix. The resulting system usually has a size were it easily can be LU-decomposed, which enables a fast implementation of a PVL-like moment matching algorithm in a second step. The total reduction time is almost identical to the pure partial realization step, while the final model size is identical to the moment matching model. Since the Lanczos algorithm is the central method in both steps, the combined technique is called Two-Step Lanczos (TSL) algorithm. By the employment of eigenvalue based stop-criteria in both steps the approximation errors are balanced while the total algorithm can be run completely automatically, see Algorithm 2. Note again, that the high efficiency of the first step of TSL only holds for FIT-discretized systems (with diagonal material matrices).
Algorithm 2: \( \text{(Two-Step Lanczos)} \)

\[
Z(s) = sB^T \left( s^2 I + sK + A \right)^{-1} B.
\]

\( s_0, \text{tol}, p_0, k_0, \text{eig}(T_0) = 0 \)

**STEP 1:**
- If lossy or field solution:
  - Find \( k_0 \) Chebychev polynomial roots \( r_k \)
  - Apply polynomial to \( b_1, \ldots, b_n \)
- Compute \( p_0 \) iterations: \( \{T_p, V_p\} = \text{Algorithm 1} \{A, B\} \)
- If \( \text{eig}(T_p) - \text{eig}(T_{p-k_0}) < \text{tol} \)
  - then
    - If lossy: \( K_p = V_p^T K V_p \)
    - If field solution: \( \tilde{c} = V z \) with \( z = (T_p + s^2 I)^{-1} B_p \)
      \( B_p = V_p^T B \)
  - else
    - then
      - Compute inverted system (LU factorization):
        \( A_p = (T_p - s_0^2 I)^{-1}, \quad B_p = A_p B_p \)
      - Compute one iteration: \( \{T_q, V_q\} = \text{Algorithm 1} \{A_p, B_p\} \)
      - If \( \text{eig}(T_q) - \text{eig}(T_p) < \text{tol} \)
        - then
          - Projection: \( Z_q(s) = sB_p^T V_q \left( s^2 I + V_q^T T_q V_q \right)^{-1} V_q^T B_p \)
          - If lossy: compute correction: \( Z_{\text{corr}} \)
            \[ Z_q(s) = sB_p^T V_q \left( s^2 I + V_q^T K_p V_q + V_q^T T_p V_q \right)^{-1} V_q^T B_p + Z_{\text{corr}} \]

Originally proposed for linear state-space systems [6] there are numerous advantages to apply TSL to the curl–curl formulation of resonant systems. Hereby it should be noticed that the final reduced model is of the same order for both linear and curl–curl systems.

The first advantage concerns the total computation time, since the starting system is only half the size of the linear one. Although each matrix–vector multiplication of the first step takes the same computation time in both cases (the curl–curl matrix has the double amount of entries per line), the number of eigenvalues in the initial system
is also halved and therefore only about half the number of iterations are required to obtain a good approximation. This effect also reduces the computation time for the system’s inversion in the beginning of the second step and the subsequent iterations.

To evaluate the complexity of the TSL algorithm it was applied to several discretizations of a test problem with up to 7.5 mio. degrees of freedom (DOF). The results in Fig. 3 show that the number of iterations required in the first step is of order $O(n^{0.45})$, and the total computation time for both steps is of order $O(n^{4/3})$. The curl–curl formulation shows the same complexity as the linear case, but the total numbers are reduced by a factor of about three. The complexity of TSL therewith has the same order as the time domain implementation of FIT (or FDTD).

Further advantages of the curl–curl formulation concern the usage of the reduced model for circuit realization. Even with an optimal (imaginary) choice of the extraction frequency the reduced system remains real. Due to the symmetry the system is and can be maintained stable and passive throughout the whole reduction process. In comparison: to obtain the symmetry of the poles for the linear case (which also enables passivity), the model (and therewith also the number of circuit elements) has to be doubled in a post-processing step.

The algorithm can easily be extended to weakly lossy systems by projecting the matrix $K$ successively to the both subspaces $V_p$ and $V_q$ with $A_q = V_q^T T_p V_q$ and $B_q = V_q^T B_p$:

$$Z_{TSL}(s) = sB_q^T (s^2 I + sV_q^T V_p KV_p V_q + A_q)^{-1} B_q.$$  

To calculate the matrix $K_q = V_q^T V_p^T K V_p V_q$ the variants suggested in Section 3.1 or 3.2 can be used.
Fig. 4. Typical approximation errors in the transfer function of a filter structure with dielectric losses. The dashed lines mark the errors in the lossless case (tan δ = 0), which are near rounding-off. The same holds for the Padé approximation of the lossy structure. The approximate projection approaches lead to errors in the range of 10^{-8} to 10^{-6}—which is still below most practical needs. Note that typical dielectric losses in filter applications are much lower than tan δ = 0.1.

Typical approximation errors of the transfer function of a lossy filter structure are shown in Fig. 4 for varying dielectric loss angles. For a rather large tan δ = 0.1 the errors of the projection approaches are below 10^{-6} and thus below most practical needs, whereas the Padé approximation still produces highly accurate results in the range of numerical rounding off.

A possible bottleneck of MOR algorithms if applied to large systems of millions of unknowns is their sensitivity to rounding-off errors. It has e.g. been reported in [29] that small deviations of zero eigenvalues can lead to qualitative changes in the system behavior and even to instabilities. In the TSL algorithms such problems have to be expected—if at all—in the first step dealing with the original full-size system. Here, an additional difficulty may arise from the large number of zero eigenvalues in this systems due to so-called static solutions from the kernel of the discrete curl-operator (the major part of them, however, do not contribute to the transfer function).

In the practical implementation of TSL these questions can rather easily be tested, since the lowest eigenvalues of the reduced system are computed to evaluate the stop criterion. In all experiments (linear and curl–curl-formulation) no critical shifts of zero eigenvalues could be observed. Of course, the static eigenvalues will not be exactly zero on the numerical level, but typical deviations are found to be orders of magnitudes below the spectral distance to the first dynamic mode. In the final system after the second step these eigenvalues can be easily detected and set to exactly zero if necessary, without any impact on the accuracy. The numerical applications of TSL in Section 5 also include a low-pass filter structure, where highly accurate transfer functions near zero frequency are obtained without any difficulties with rounding-off.
4. Circuit extraction from reduced models

The network generation from reduced order models is based on the interpretation of the state vector of unknowns \( z \) (after the second step of TSL) as node voltages in a nodal analysis formulation.

4.1. Circuit extraction for lossless systems

Again the extraction method is first stated for lossless systems. With \( A_q = V_q^T T_p V_q \) and \( B_q = V_q^T B_p \) the final reduced model (30) becomes

\[
\begin{align*}
    s^2 z &= -A_q z + s B_q i, \\
    u &= B_q^T z.
\end{align*}
\]

With a computationally cheap eigenvalue decomposition \( L = X^{-1} A_q X \) a positive semidefinite (real) diagonal matrix \( L \) and a unitary matrix \( X^T X = I \) are obtained. After division by \( s \) and setting \( D = X^T B_q \) and \( C = I \) (not to be mixed up with the discrete curl-matrix in (1)) this system has the typical form of a lossless nodal analysis formulation:

\[
\begin{align*}
    s C z' &= -\frac{1}{s} L z' + D i, \\
    u &= D^T z'.
\end{align*}
\]

\( D \) represents a set of \( m \) full vectors, \( C \) is in the easiest case just the identity matrix, but might be extended to use some scaling factors for a better balancing of the values in \( L \) and \( C \). Each entry \( C_{ii} \) and \( L_{ii} \) represents a capacitor or (inverse) inductor, respectively, between node \( i \) and the ground. Additionally each node is excited by \( m \) parallel current sources \( D_{i1} \ldots D_{im} \), controlled by the port currents \( i_1 \ldots i_m \). Finally the port voltage is coupled out by a series of controlled voltage sources, also described by the matrix entries \( D_{v1} \ldots D_{vm} \). The resulting network is shown in Fig. 5 for the case with \( m = 2 \) ports.

The netlist contains \((2m + 2)q\) elements. By using nonlinear controlled sources (not available in all circuit simulators), the parallel current sources at each node and the serial voltage sources at each node can be combined, ending up with a netlist of \( 3q + m \) elements.

Fig. 5. Equivalent circuit generation for a system with \( m = 2 \) ports.
The physical meaning of the generated network gets clearer by considering the equivalent circuit of an ideal transformer in Fig. 6.

Therewith each node of the circuit in Fig. 5, contains an LC resonance representing one pole of the system that is connected via an ideal transformer network with the ports. Since all values of L and C are positive also the passivity of the network gets obvious. However, for the use of the netlist in a circuit simulator such as SPICE the controlled sources are to be preferred to ideal transformers.

4.2. Circuit extraction for lossy systems

The treatment of lossy systems is quite similar to the previously described method, with the difference that the systems (33) additionally contains a positive semidefinite but full matrix $G = X^T V^T q T K P V q T X$:

$$s C z' = -\frac{1}{s} L z' + G z' + D i,$$
$$u = D^T z'.$$

The entries $G_{ij}$ represent the admittance of a resistor between the nodes $i$ and $j$, the sum $G_i = \sum_j G_{ij}$ the admittance between node $i$ and ground, see also Fig. 7.

The number of elements in the circuit increases to $(2m + 2.5)q + 0.5q^2$. Since some of the elements in $G$ are negative, the circuit now also contains negative resistors, but, due to the positive definiteness of $G$, the overall passivity of the network is still guaranteed. To obtain a smaller circuit that consists of $(2m + 3)q$ positive elements, only the diagonal entries of $G$ can be considered while the $G_{ij}$ ($i \neq j$) are neglected. Due to some diagonal dominance in $G$ the introduced error is small for many practical applications.
5. Numerical results

5.1. Dielectric filter

The proposed algorithm is applied to a dielectric filter shown in Fig. 8. It consists of a perfect electric conducting box containing two dielectric rings ($\varepsilon_r = 38$) excited by two coaxial antennas. The structure discretization with at least 10 lines per wavelength results in 14,632 mesh points and an problem dimension of 43,896 unknowns. The bandwidth of interest is 4–8 GHz, the extraction frequency is chosen to be the mid-frequency $s_0 = j2\pi 6$ GHz. The transfer function $Z_q$ obtained by reduction is then used to compute the scattering-parameters—reflection $S_{11}(f)$ and transmission $S_{21}(f)$—of the filter.

The algorithm is implemented in Matlab and runs on a 731 MHz PC. The computation time for this problem with TSL curl–curl including the fast frequency sweep is only 42 s, 99% of which are used for the first step (1000 iterations) and less than 1% for the second step (22 iterations). The automatic stop criteria is started for the first time after 600 first step iterations and then successively repeated every 40 steps, leading to an computational overhead of only 7%. The numerical cost of the stop criterion in the second step is negligible.

To point out the superiority of this algorithm concerning the computation time for this resonant example, Table 1 shows the comparison to other methods [16], which are: a FIT time domain simulation (equivalent to FDTD), without and with application of an autoregressive (AR) filter as advanced signal processing technique, and direct PVL, using the iterative solvers BiConjugate Gradient method (BiCG) [1] and Conjugate Orthogonal Conjugate Gradient method (COCG). In both cases the solver accuracy is chosen to be $10^{-6}$, a value found to be sufficient to avoid follow-up errors in the iteration.

Additionally, the numerical cost of accelerated TSL and corrected modal expansion are listed in the table and will be described below. Besides computation time
Table 1
Various methods for the calculation of the filter’s $S$-parameters

<table>
<thead>
<tr>
<th>Method</th>
<th>Computation time [s]</th>
<th>Matrix–vector multiplications</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIT time domain/FDTD</td>
<td>$\sim 9000$</td>
<td>$\sim 1,500,000$</td>
</tr>
<tr>
<td>FIT time domain + AR-filter</td>
<td>1047</td>
<td>59,803</td>
</tr>
<tr>
<td>Corr. modal approach, 16 modes</td>
<td>344</td>
<td>10,049</td>
</tr>
<tr>
<td>Corr. modal approach, 10 modes</td>
<td>181</td>
<td>7090</td>
</tr>
<tr>
<td>PVL linear, BiCG, order 22</td>
<td>6901</td>
<td>69,270</td>
</tr>
<tr>
<td>PVL curl–curl, COCG, order 22</td>
<td>613</td>
<td>18,146</td>
</tr>
<tr>
<td>TSL linear, order 2020/22</td>
<td>137</td>
<td>2020</td>
</tr>
<tr>
<td><strong>TSL curl–curl, order 1000/22</strong></td>
<td><strong>42</strong></td>
<td><strong>1000</strong></td>
</tr>
<tr>
<td>TSL curl–curl, 330/22, accel. (450)</td>
<td>37</td>
<td>1230</td>
</tr>
<tr>
<td>TSL curl–curl, 330/22, accel.+corr.</td>
<td>65</td>
<td>1924</td>
</tr>
</tbody>
</table>

A second possibility for a comparison is given by counting the number of matrix–vector multiplications (original-sized system matrix only). However, the algorithms also differ slightly in the number of orthogonalization steps.

The reduced model can be transformed to a sparse symmetric system and realized as a SPICE netlist containing 132 linear elements. The SPICE simulation time based on this circuit model is only 0.6 s. Fig. 9 shows the absolute value of $|S_{21}|$ for the initial system (reference) and for reduced systems obtained by various methods: non-inverted Lanczos (order 600, stop criterion not yet fulfilled), TSL curl–curl (order 1000, 22) and the SPICE simulation result.

It can be noticed that if the eigenvalue criterion is not fulfilled for all eigenvalues in the first step and the iteration is stopped too early, the results are not satisfactory in the whole frequency range of interest. In contrast, the full TSL of order $p = 1000$ in the first and order $q = 22$ in the second step (as found by the automatic stop criterion) and the SPICE simulation results are indistinguishable from the initial system, the averaged error is below $10^{-5}$ for both of them.

![Fig. 9. Frequency variation of the $S_{21}$ parameter.](image-url)
If a standard frequency domain solver is employed to calculate the transfer behavior at single frequency points, about 600 iterations are sufficient between 4 GHz and 7 GHz, and around 1000 are needed above. This again points out the strong relation between the first step iterations and iterative solvers.

Additionally, the equivalent network is also compared to FIT in a time domain simulation, which is an important application e.g. for verifying signal integrity in digital circuits. In time domain FIT (with the original system size), the filter is excited by a digital wave pulse \( a_1(t) \) with rise and fall times of 0.1 ns and a hold time of 0.3 ns, and the reflected signal \( b_1(t) \) is measured. The extracted circuit netlist is excited in a SPICE simulator with a voltage equivalent to the FIT wave signals (corresponding to the relation \( u = Z_l^{1/2}(a + b) \) with the reference line impedance matrix \( Z_l \)). From the resulting SPICE currents again wave quantities are calculated according to \( i = Z_l^{-1/2}(a - b) \). Fig. 10 shows the comparison between FIT and the equivalent circuit wave signals.

The simulation times of SPICE is 5.6 s. Both input and reflected signal show very good agreement. Obviously the reconstructed SPICE signals are overlaid by a high frequency ripple that is due to a pole outside the frequency range of validity of the model. It can be removed by filtering the signal or simply by neglecting the relevant pole in the circuit extraction process (with some consequence on the transfer function).

Finally the extension for lossy materials is investigated. The dielectric rings of the filter are provided with a conductivity of 0.2 S/m, representing a loss angle of \( \tan \delta \approx 0.016 \). The results of the TSLlossy algorithm with some variants as described above are shown in Fig. 11.

By keeping the whole matrix \( V_p \) in memory, the curves of the FIT reference and of TSLlossy are indistinguishable, and the computation time is slightly increased to 48 s. If maximal 50 out of 1000 vectors are considered both curves slightly differ, but the memory requirements are drastically reduced. The equivalent circuits are based

![Fig. 10. Time domain simulation: The introduced wave pulse and the reflected signal. FIT simulations are dashed lines, SPICE simulations of the extracted circuit are solid lines. The high frequency ripples in the SPICE signals are due to a pole outside the model’s frequency range of validity.](image-url)
on the complete $V_p$ model, one under consideration of the full matrix $G$ having 385 elements, one with only the diagonal elements of $G$ with only 154 elements in total. The first SPICE simulation takes 6.6 s and again covers the reference curve, while the second case with a simulation time of only 1.4 s varies insignificantly from the reference.

Applying a Chebychev polynomial of order 450 to both starting vectors before employing TSL reduces the number of first step iterations to only 330 instead of 1000 at TSL. Although the total number of matrix–vector multiplications is increased by the polynomials, the overall computation time is slightly reduced due to less orthonormalization operations and stop criterion evaluations in TSL. However, as shown in Fig. 12, for the curves to fit as well as in pure TSL a correction step is needed which takes additionally 28 s of computation time. The 330 vectors can easily be held in memory ($\sim 110$ MB using double precision numbers) allowing the calculation of full field solutions additionally to the transfer behavior. The resulting numerical cost—represented by a total number of 1924 matrix–vector multiplications and 65 s of CPU-time—is higher than with pure TSL, but still outperforms classical approaches by an order of magnitude.

For a modal approach also shown in Fig. 12, ten eigenvectors were employed for projection, followed by a correction step. Since the two upper eigenvalues are not considered in this computation, the $S_{21}$-parameter differs for frequencies above 7 GHz, but shows good agreement below.

5.2. Waveguide filter

The second structure investigated with various MOR techniques is a waveguide filter shown in Fig. 13. At both ends of the structure it is excited by a waveguide port, modeling the infinite continuation of the waveguide by exciting its 2D mode pattern. The filter consists of six coupled waveguide cavities and performs as a sharp bandpass between 7.1 and 7.3 GHz. Further explanations on the principles of such filters can
Fig. 12. $S_{21}$ parameter for an accelerated first step in TSL. The uncorrected implementation show quite some differences, the corrected curve and the reference are almost indistinguishable. The corrected modal approach shows with only 10 modes good agreement up to 7 GHz.

be found in [30]. To represent the slopes in the passband with high accuracy, the filter has to be discretized with a very fine mesh using 40 gridlines per smallest wavelength. This results in a model with 159,720 mesh nodes, yielding 479,160 unknowns in the case of the curl–curl formulation and 958,320 unknowns for a linear state space.

The application of the TSL algorithm to the curl–curl formulation is again found to be superior to existing approaches. The transfer behavior is calculated at 1000 frequency points within only 5 min on a 731 MHz PC, compared to about 70 min. needed by a standard time domain solver including an AR-filter. As shown in Fig. 14, the curves for the $S$-parameters are almost indistinguishable. TSL needs 760 iterations in the first step and 10 in the second. The stop criterion in the first step is first employed after 500 iterations, repeating the check every 20 iterations. The criterion hereby only consumes 1% of the total computational cost.

In case of the modal expansion, eight modes are used to cover all six poles within the frequency range of interest and two below this range. The pure modal expansion shows only poor agreement with the reference curve (Fig. 15). After the correction step the curves are matching quite well, emphasizing the importance of the correction

Fig. 13. Analyzed waveguide filter structure.
Fig. 14. Frequency variation of the S parameters in the waveguide filter. The curves of the reference and TSL with $p = 760$ and $q = 10$ are almost indistinguishable.

Fig. 15. The same frequency variation with modal expansion considering 8 modes. It is obvious, that correction is crucial for the method. The corrected curves show good agreement with the reference.

The computation time is 1027 s, 27% of which are used for the correction step. An overview of the considered algorithms is given in Table 2.

Table 2
Various methods for the calculation of the waveguide filter’s $S$-parameters

<table>
<thead>
<tr>
<th>Method</th>
<th>Computation time [s]</th>
<th>Matrix–vector multiplications</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIT time domain + AR-filter</td>
<td>4657</td>
<td>59,801</td>
</tr>
<tr>
<td>TSL linear, order 1560/10</td>
<td>881</td>
<td>1560</td>
</tr>
<tr>
<td>TSL curl–curl, order 760/10</td>
<td><strong>287</strong></td>
<td><strong>760</strong></td>
</tr>
<tr>
<td>Corr. modal expansion, 8 modes</td>
<td>1027</td>
<td>3245</td>
</tr>
</tbody>
</table>
5.3. Lowpass filter

Finally, the lowpass microstrip filter shown in Fig. 16 is analyzed. The mode patterns at the ports of this filter (the excitation matrix $B$ in the state-space formulation) describe quasi-TEM modes which are only approximately constant over frequency. For broadband simulations within the lower frequency range this effect is typically neglected. The resulting small systematic error can be further reduced if the frequency range of interest is subdivided in smaller intervals and the mode pattern are recalculated for each of these intervals. Since this error is introduced into the approaches compared here (TSL, time domain and frequency domain) in different ways, small deviations in the transfer functions have to be expected.

The discrete model consists of 34,560 unknowns, and the computation times are 40.2 s for TSL with $p = 1260$, $q = 9$, and 168 s for the time domain reference solution.

![Fig. 16. Microstrip lowpass filter with quasi-TEM mode excitation.](image)

![Fig. 17. Transmission coefficient of lowpass filter: TSL solution and references from time and frequency domain. The inlet shows the high accuracy of the reduced model even in the near of zero frequency.](image)
The results in Fig. 17 show that in spite of the discussion above the curves from the different approaches are again almost indistinguishable. The inlet in Fig. 17 zooms the frequency range near zero, where of course this microstrip structure behaves like a DC line with $S_{21}(0) = 1$. This is reproduced with high accuracy by the TSL approach which obviously has no rounding-off problems even in this nearly singular range. Only for extremely small frequencies the state-space formulation degenerates, however as demonstrated here without any impact on practical simulation results.

6. Conclusions

Various projection based Model Order Reduction techniques have been reviewed in the context of large state-space systems arising from electromagnetic field simulations using FIT. Strong relations between Krylov subspaces—applied to the system matrix or its shifted inverse, respectively—and modal expansion techniques have been found. Depending on the need for either pure transfer functions and/or also field solutions, the application of eigenvalue acceleration techniques allow a ‘smooth’ transition between partial realizations and the classical modal approach. To compensate for non-considered higher modes a simple but efficient correction scheme is proposed.

In the so-called Two-Step Lanczos (TSL) algorithm a Krylov-based partial realization step is applied to the non-inverted system first, followed by a PVL-like algorithm on the pre-reduced system. This allows to benefit from the superior properties of moment matching techniques—such as their high approximation accuracy around the expansion point and the small resulting model sizes—without the need to invert or factorize the original system matrix. Thus, a highly accurate and efficient MOR technique is available even for very large systems with millions of unknowns.

Originally developed for lossless structures and a linear formulation, TSL is applied to the curl–curl system matrix. Some extensions to weakly lossy structures are proposed, representing a trade-off between exact moment matching (like in the lossless case) and the numerical cost of alternative methods such as linear TSL or WCAWE. TSL curl–curl in combination with a fully automatic stop criterion has been tested for more than 2,500,000 mesh nodes ($\sim$7,500,000 unknowns) and appears to be very robust, with an overall complexity of about $O(n^{4/3})$ referring to the problem size. It additionally provides a netlist for an equivalent circuit, which is guaranteed to be stable, passive, and real. This is a prerequisite for a direct usage in general circuit simulators, opening a new field of applications such as electromagnetic field-circuit coupled problems.

Two filter applications demonstrate the reliability and high accuracy of TSL, outperforming advanced time domain methods by a factor of around 20. Thus, it seems to be the most efficient simulation method for resonant multiport applications discretized by FIT.
References


