Improved Numerical Solutions for the Simulation of Monolithic Microwave Integrated Circuits

G. Hebermehl, * R. Schlundt †

Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, D-10117 Berlin, Germany

H. Zscheile, ‡ W. Heinrich §

Ferdinand-Braun-Institut für Höchstfrequenztechnik, Rudower Chaussee 5, D-12489 Berlin, Germany

April 22, 1996

1991 Mathematics Subject Classification.
35Q60, 35L20, 65N22, 65F10, 65F15.
Keywords.
Microwave device, three-dimensional simulation,
scattering matrix, orthogonality relation, Maxwellian equations,
finite-volume method, finite-difference method,
eigenvalue problem, system of simultaneous linear equations.

*e-mail: hebermehl@wias-berlin.de, URL: http://hyperg.wias-berlin.de
†e-mail: schlundt@wias-berlin.de, URL: http://hyperg.wias-berlin.de
‡e-mail: zscheile@fbh.fta-berlin.de, URL: http://www.fta-berlin.de
§e-mail: heinrich@fbh.fta-berlin.de, URL: http://www.fta-berlin.de

Abstract

The electrical properties of the circuits are described in terms of their scattering matrix using Maxwellian equations. Using a finite-volume scheme a three-dimensional boundary value problem for the Maxwellian equations in the frequency domain can be solved. This results in a two-step procedure: a time and memory consuming eigenvalue problem for nonsymmetric matrices and the solution of a large-scale system of linear equations with indefinite symmetric matrices. Improved numerical solutions for these two linear algebraic problems, the computation of the scattering matrix and of the used orthogonality relation are treated in this paper. The numerical effort could be reduced considerably.

Contents

List of Figures

1 Introduction

The model for the simulation of monolithic microwave integrated circuits, and the finite-volume method for the solution of the corresponding three-dimensional boundary value problem for the Maxwellian equations [4], [5], [2] are in detail treated in [10]. We refer in this paper to [10].
The essential points of the method are :

- Microwave devices can be described by an interconnection of infinitely long homogeneous transmission lines which are attached to a structure. The waveguides and the structure are shielded by electric or magnetic walls. The waveguides and the enclosures are cut at cross-sectional planes. The tangential electric or the tangential magnetic field is known on the whole surface. Figure 1 illustrates the principal structure under investigation.

- The electrical properties of the circuits are described in terms of their scattering matrix using Maxwellian equations.

- The scattering matrix can be computed if an orthogonal decomposition of the electric field can be calculated at a pair of two neighboring cross-sectional planes on each transmission line for a number of linear independent excitations of the transmission lines.

- The boundary region is divided into elementary rectangular parallelepipeds by using a three-dimensional nonequidistant Cartesian grid. The application of the finite-volume method to the three-dimensional boundary value problem for the Maxwellian equations in the frequency domain results in an eigenvalue problem for nonsymmetric matrices and the solution of a system of linear equations with indefinite symmetric matrices.

We introduce in this paper improved numerical methods for the time and memory consuming eigenvalue problem and the solution of the system of linear equations of the program package F3D (Finite Differenzen dreidimensional) [5], [2]. We represent the computation of the scattering matrix and of the orthogonality relation used in F3D.
The program package F3D allows to simulate the electromagnetic field of nearly arbitrary shaped structures.

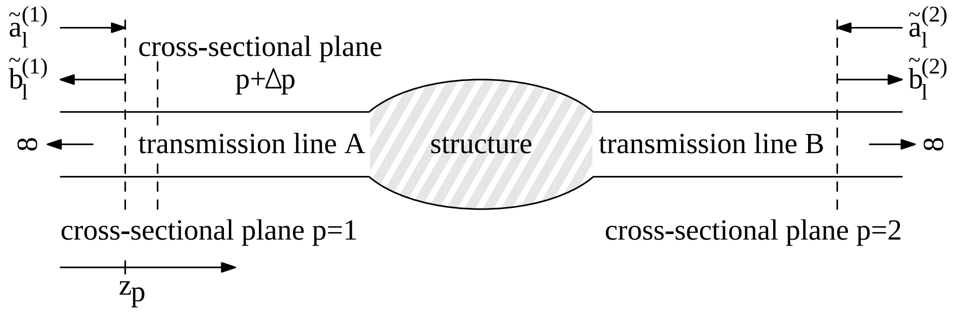

Figure 1: Structure under investigation

Improving the eigenvalue procedure increased efficiency by factors of 25 until 40 in typical medium-sized examples compared to the original version of this part of the package F3D. The storage requirements could be reduced by a factor of 20 in the example used for demonstration in this paper, since the sparse storage technique is applied.
Furthermore, the execution time for the solution of the linear algebraic equations was reduced by a factor of 5.
The improvements in the computing times and the storage requirements increase essentially with the dimension of the problem. Especially the reduction of the storage requirements permits the solution of problems of higher dimensions because the size of the memory of the workstations is restricted.

2 The Numerical Solution of the Eigenvalue Problem

(1) and (2) form an eigenvalue problem for the transverse electric field on the transmission line region (see Figure 1). The cross-sectional plane is located on the (x, y)-plane of the enclosure.
If the cross-sectional plane is located on the (x, z)-plane or on the (y, z)-plane of the enclosure, one can derive similar equations which correspond to (1) and (2).
The eigenvalue problem is derived in [10].
$\gamma(h)$ are the eigenvalues. $E_{x_{i,j,k}}$, $E_{y_{i,j,k}}$, $k = const$, are the components of the eigenfunctions (transverse electric field).

$$
\begin{aligned}
&-\frac{\tilde{c}^{z,y}_{i,j-1,k}}{\tilde{c}^{y,z}_{i,j,k}} E_{x_{i,j-1,k}} - \frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \frac{\tilde{g}^{y,z}_{i-1,j,k}}{\tilde{g}^{x,y}_{i,j,k}} E_{x_{i-1,j,k}} + \\
&\left(\frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \left(\frac{\tilde{g}^{y,z}_{i,j,k}}{\tilde{g}^{x,y}_{i,j,k}} + \frac{\tilde{g}^{y,z}_{i,j,k}}{\tilde{g}^{x,y}_{i+1,j,k}} \right) + \frac{\tilde{c}^{z,y}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} + \frac{\tilde{c}^{z,y}_{i,j-1,k}}{\tilde{c}^{y,z}_{i,j,k}} - 2\varkappa_0^2 \frac{\tilde{g}^{y,z}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \right) E_{x_{i,j,k}} - \\
&\frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \frac{\tilde{g}^{y,z}_{i+1,j,k}}{\tilde{g}^{x,y}_{i+1,j,k}} E_{x_{i+1,j,k}} - \frac{\tilde{c}^{z,y}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} E_{x_{i,j+1,k}} - \\
&\left(\frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \frac{\tilde{g}^{x,z}_{i,j-1,k}}{\tilde{g}^{x,y}_{i,j,k}} - \frac{\tilde{c}^{z,x}_{i,j-1,k}}{\tilde{c}^{y,z}_{i,j,k}} \right) E_{y_{i,j-1,k}} + \left(\frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \frac{\tilde{g}^{x,z}_{i+1,j-1,k}}{\tilde{g}^{x,y}_{i+1,j,k}} - \frac{\tilde{c}^{z,x}_{i,j-1,k}}{\tilde{c}^{y,z}_{i,j,k}} \right) E_{y_{i+1,j-1,k}} + \\
&\left(\frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \frac{\tilde{g}^{x,z}_{i,j,k}}{\tilde{g}^{x,y}_{i,j,k}} - \frac{\tilde{c}^{z,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \right) E_{y_{i,j,k}} + \left(-\frac{\tilde{c}^{y,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \frac{\tilde{g}^{x,z}_{i+1,j,k}}{\tilde{g}^{x,y}_{i+1,j,k}} + \frac{\tilde{c}^{z,x}_{i,j,k}}{\tilde{c}^{y,z}_{i,j,k}} \right) E_{y_{i+1,j,k}} \\
&\qquad\qquad = \gamma(h) E_{x_{i,j,k}} \quad ,
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
&\left(-\frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \frac{\tilde{g}^{y,z}_{i-1,j,k}}{\tilde{g}^{x,y}_{i,j,k}} + \frac{\tilde{c}^{z,y}_{i-1,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \right) E_{x_{i-1,j,k}} + \left(\frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \frac{\tilde{g}^{y,z}_{i,j,k}}{\tilde{g}^{x,y}_{i,j,k}} - \frac{\tilde{c}^{z,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \right) E_{x_{i,j,k}} + \\
&\left(\frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \frac{\tilde{g}^{y,z}_{i-1,j+1,k}}{\tilde{g}^{x,y}_{i,j+1,k}} - \frac{\tilde{c}^{z,y}_{i-1,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \right) E_{x_{i-1,j+1,k}} + \left(-\frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \frac{\tilde{g}^{y,z}_{i,j+1,k}}{\tilde{g}^{x,y}_{i,j+1,k}} + \frac{\tilde{c}^{z,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \right) E_{x_{i,j+1,k}} - \\
&\frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \frac{\tilde{g}^{x,z}_{i,j-1,k}}{\tilde{g}^{x,y}_{i,j,k}} E_{y_{i,j-1,k}} - \frac{\tilde{c}^{z,x}_{i-1,j,k}}{\tilde{c}^{x,z}_{i,j,k}} E_{y_{i-1,j,k}} + \\
&\left(\frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \left(\frac{\tilde{g}^{x,z}_{i,j,k}}{\tilde{g}^{x,y}_{i,j,k}} + \frac{\tilde{g}^{x,z}_{i,j,k}}{\tilde{g}^{x,y}_{i,j+1,k}} \right) + \frac{\tilde{c}^{z,x}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} + \frac{\tilde{c}^{z,x}_{i-1,j,k}}{\tilde{c}^{x,z}_{i,j,k}} - 2\varkappa_0^2 \frac{\tilde{g}^{x,z}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \right) E_{y_{i,j,k}} - \\
&\frac{\tilde{c}^{z,x}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} E_{y_{i+1,j,k}} - \frac{\tilde{c}^{x,y}_{i,j,k}}{\tilde{c}^{x,z}_{i,j,k}} \frac{\tilde{g}^{x,z}_{i,j+1,k}}{\tilde{g}^{x,y}_{i,j+1,k}} E_{y_{i,j+1,k}} \\
&\qquad\qquad = \gamma(h) E_{y_{i,j,k}}
\end{aligned}
\tag{2}
$$

with (k_z is the propagation constant.)

$$
\gamma(h) = e^{-jk_z 2h} + e^{+jk_z 2h} - 2 = -4\sin^2(k_z h) \tag{3}
$$

and with

$$
\begin{aligned}
\tilde{c}^{z,t}_{i,j,k} &= \tfrac{4h}{\tilde{\mu}_{i,j,k} t_{i,j,k}}, \quad t \in \{x,y\}, \\
\tilde{c}^{s,z}_{i,j,k} &= \left(\tfrac{s_{i,j,k}}{\tilde{\mu}_{i,j,k}} + \tfrac{s_{i',j',k'}}{\tilde{\mu}_{i',j',k'}}\right) \tfrac{1}{2h}, \quad s \in \{x,y\}, \\
\tilde{c}^{s,t}_{i,j,k} &= \left(\tfrac{s_{i,j,k}}{\tilde{\mu}_{i,j,k}} + \tfrac{s_{i',j',k'}}{\tilde{\mu}_{i',j',k'}}\right) \tfrac{1}{t_{i,j,k}}, \quad s,t \in \{x,y\},
\end{aligned}
\tag{4}
$$

and with

$$
\begin{aligned}
\tilde{g}^{y,z}_{i,j,k} &= h\big(y_{i,j,k}\tilde{\epsilon}_{i,j,k} + y_{i,j-1,k}\tilde{\epsilon}_{i,j-1,k}\big), \\
\tilde{g}^{x,z}_{i,j,k} &= h\big(x_{i,j,k}\tilde{\epsilon}_{i,j,k} + x_{i-1,j,k}\tilde{\epsilon}_{i-1,j,k}\big), \\
\tilde{g}^{x,y}_{i,j,k} &= \big(\tfrac{x_{i,j,k}y_{i,j,k}}{4}\tilde{\epsilon}_{i,j,k} + \tfrac{x_{i-1,j,k}y_{i-1,j,k}}{4}\tilde{\epsilon}_{i-1,j,k} + \\
&\quad \tfrac{x_{i-1,j-1,k}y_{i-1,j-1,k}}{4}\tilde{\epsilon}_{i-1,j-1,k} + \tfrac{x_{i,j-1,k}y_{i,j-1,k}}{4}\tilde{\epsilon}_{i,j-1,k}\big) \ .
\end{aligned}
\tag{5}
$$

(i', j', k') are the indices of the elementary cell which is located in s-direction in front of the cell (i, j, k):

$$
\begin{aligned}
s = x: &\quad i' = i-1, \quad j' = j, \quad k' = k \quad , \\
s = y: &\quad i' = i, \quad j' = j-1, \quad k' = k \quad .
\end{aligned}
$$

Because we use a Cartesian grid, we have

$$
\begin{aligned}
x_{i,j,k} &= x_{i,j-1,k} = x_{i,j+1,k} \ , \\
y_{i,j,k} &= y_{i+1,j,k} = y_{i-1,j,k} \ .
\end{aligned}
$$

There are the following relations between the quantities (complex permittivity $\underline{\epsilon}$, permeability μ, conductivity κ, circular frequency ω, wavenumber in vacuo $\varkappa_0$) (see [10])

$$
\underline{\epsilon} = \epsilon + \frac{\kappa}{j\omega}, \quad \underline{\epsilon} = \tilde{\epsilon}\epsilon_0, \quad \mu = \tilde{\mu}\mu_0, \quad \varkappa_0 = \omega\sqrt{\epsilon_0\mu_0} \ . \tag{6}
$$

The quantities $\tilde{\epsilon}$ and $\tilde{\mu}$ can be different from cell to cell.
The cross-sectional planes can be located on the 6 different planes of the rectangular parallelepiped.
Let be

$$
\begin{aligned}
\underline{\vec{e}} = (\underline{\vec{e}}_x, \underline{\vec{e}}_y)^T \quad , \quad \underline{\vec{e}}_x &= (\underline{e}_{x_1}, \underline{e}_{x_2}, \ldots, \underline{e}_{x_{n_{xy}}}) \quad , \quad \underline{e}_{x_\ell} = E_{x_{i,j,k}}, \\
\underline{\vec{e}}_y &= (\underline{e}_{y_1}, \underline{e}_{y_2}, \ldots, \underline{e}_{y_{n_{xy}}}) \quad , \quad \underline{e}_{y_\ell} = E_{y_{i,j,k}}
\end{aligned}
\tag{7}
$$

with

$$
\begin{aligned}
&\ell = (j-1)n_x + i, \quad i = 1(1)n_x, \quad j = 1(1)n_y, \\
&n_{xy} = n_x n_y, \quad \text{and} \quad k = 1 \quad \text{or} \quad k = n_z \ .
\end{aligned}
\tag{8}
$$

The assumption $k = 1$ corresponds to the case, in which the cross-sectional plane (see Figure 1) is located on the left-handed (x, y)-plane of the enclosure.
On the transmission line wall, the tangential component $\vec{E}_{tang}$ or the tangential component $\vec{H}_{tang}$ must vanish. Because we reduced the wave propagation on the transmission line to a two-dimensional problem (see [10]), we have to take into account boundary conditions on the port of the structure only (see Figure 2). We consider the boundary condition

$$
\vec{E}_{tang} = 0 \ , \tag{9}
$$

on the left-hand side and at the bottom of the port as an example, that is,

$$
\underline{e}_{x_l} = E_{x_{i,1,1}} = 0, \quad l = i, \quad i = 1(1)n_x, \tag{10}
$$

and

$$
\underline{e}_{y_l} = E_{y_{1,j,1}} = 0, \quad \ell = (j-1)n_x + 1, \quad j = 1(1)n_y \ . \tag{11}
$$

Let be $M = (m_{p,q}), \quad p, q = 1(1)2n_{xy}$, the matrix of the eigenvalue problem (1, 2). Because of the boundary conditions (10) and (11) we get from the eigenvalue problem (1, 2) for $i = 1(1)n_x$ and for $j = 1(1)n_y$:

$$
\begin{aligned}
&m_{1,2}E_{x_{2,1,1}} + m_{1,n_x+1}E_{x_{1,2,1}} + \\
&m_{1,n_{xy}+1}E_{y_{1,1,1}} + m_{1,n_{xy}+2}E_{y_{2,1,1}} = 0, \\
&m_{i,i-1}E_{x_{i-1,1,1}} + m_{i,i+1}E_{x_{i+1,1,1}} + m_{i,n_x+i}E_{x_{i,2,1}} + \\
&m_{i,n_{xy}+i}E_{y_{i,1,1}} + m_{i,n_{xy}+i+1}E_{y_{i+1,1,1}} = 0, \\
&i = 2(1)n_x - 1, \\
&m_{n_x,n_x-1}E_{x_{n_x-1,1,1}} + m_{n_x,n_x+n_x}E_{x_{n_x,2,1}} + \\
&m_{n_x,n_{xy}+n_x}E_{y_{n_x,1,1}} = 0
\end{aligned}
\tag{12}
$$

and

$$
\begin{aligned}
&m_{n_{xy}+1,1}E_{x_{1,1,1}} + m_{n_{xy}+1,n_x+1}E_{x_{1,2,1}} + \\
&m_{n_{xy}+1,n_{xy}+2}E_{y_{2,1,1}} + m_{n_{xy}+1,n_{xy}+n_x+1}E_{y_{1,2,1}} = 0, \\
&m_{n_{xy}+(j-1)n_x+1,(j-1)n_x+1}E_{x_{1,j,1}} + \\
&m_{n_{xy}+(j-1)n_x+1,jn_x+1}E_{x_{1,j+1,1}} + \\
&m_{n_{xy}+(j-1)n_x+1,n_{xy}+(j-2)n_x+1}E_{y_{1,j-1,1}} + \\
&m_{n_{xy}+(j-1)n_x+1,n_{xy}+(j-1)n_x+2}E_{y_{2,j,1}} + \\
&m_{n_{xy}+(j-1)n_x+1,n_{xy}+jn_x+1}E_{y_{1,j+1,1}} = 0, \quad j = 2(1)n_y-1, \\
&m_{n_{xy}+(n_y-1)n_x+1,(n_y-1)n_x+1}E_{x_{1,n_y,1}} + \\
&m_{n_{xy}+(n_y-1)n_x+1,n_{xy}+(n_y-2)n_x+1}E_{y_{1,n_y-1,1}} + \\
&m_{n_{xy}+(n_y-1)n_x+1,n_{xy}+(n_y-1)n_x+2}E_{y_{2,n_y,1}} = 0 \ .
\end{aligned}
\tag{13}
$$

That means, the Equations (12) and (13) have not to be taken into account. The corresponding components of the eigenvectors are known (see (10) and (11)). Thus, the dimension of the eigenvalue problem is reduced to $2n_{xy} - n_b$, $n_b = n_x + n_y$, in this case. We denote the matrix of the reduced eigenvalue problem with $\bar{M}$.

We have to remark that we generally have to take into account also boundary conditions at interior boundaries, which are to be treated in the same manner. The n_{xy} components of the vector $\underline{\vec{e}}_x$ are the first n_{xy} components of $\vec{e}_x$, and the n_{xy} components of the vector $\underline{\vec{e}}_y$ are the first n_{xy} components of $\vec{e}_y$, both defined in ([10], Equation (36)). After solving the eigenvalue problem these $2n_{xy}$ components are used as boundary values for the three-dimensional boundary value problem described in [10] which results in the large scale system of linear equations with the $3n_{xy}n_z$ unknowns $\vec{e}_x, \vec{e}_y, \vec{e}_z$.

If the cross-sectional plane is located on the right-handed (x, y)-plane of the enclosure, we have $k = n_z$. In this case, the n_{xy} components of $\underline{\vec{e}}_x$ are the last n_{xy} components of $\vec{e}_x$, and the n_{xy} components of $\underline{\vec{e}}_y$ are the last n_{xy} components of $\vec{e}_y$.

If the cross-sectional planes are located on the (x, z)- or (y, z)- planes of the enclosure, we can find easily in a similar way the appropriate components of $\underline{\vec{e}}$ and $\vec{e}$.

If the cross-sectional plane is located on the (x, z)-plane, the y-direction is the longitudinal direction of the corresponding transmission line, and we have

$$
\underline{\vec{e}} = (\underline{\vec{e}}_x, \underline{\vec{e}}_z)^T \quad , \quad
\begin{aligned}
\underline{\vec{e}}_x &= (\underline{e}_{x_1}, \underline{e}_{x_2}, \ldots, \underline{e}_{x_{n_{xz}}}) \quad , \quad \underline{e}_{x_\ell} = E_{x_{i,j,k}}, \\
\underline{\vec{e}}_z &= (\underline{e}_{z_1}, \underline{e}_{z_2}, \ldots, \underline{e}_{z_{n_{xz}}}) \quad , \quad \underline{e}_{z_\ell} = E_{z_{i,j,k}}
\end{aligned}
\tag{14}
$$

with

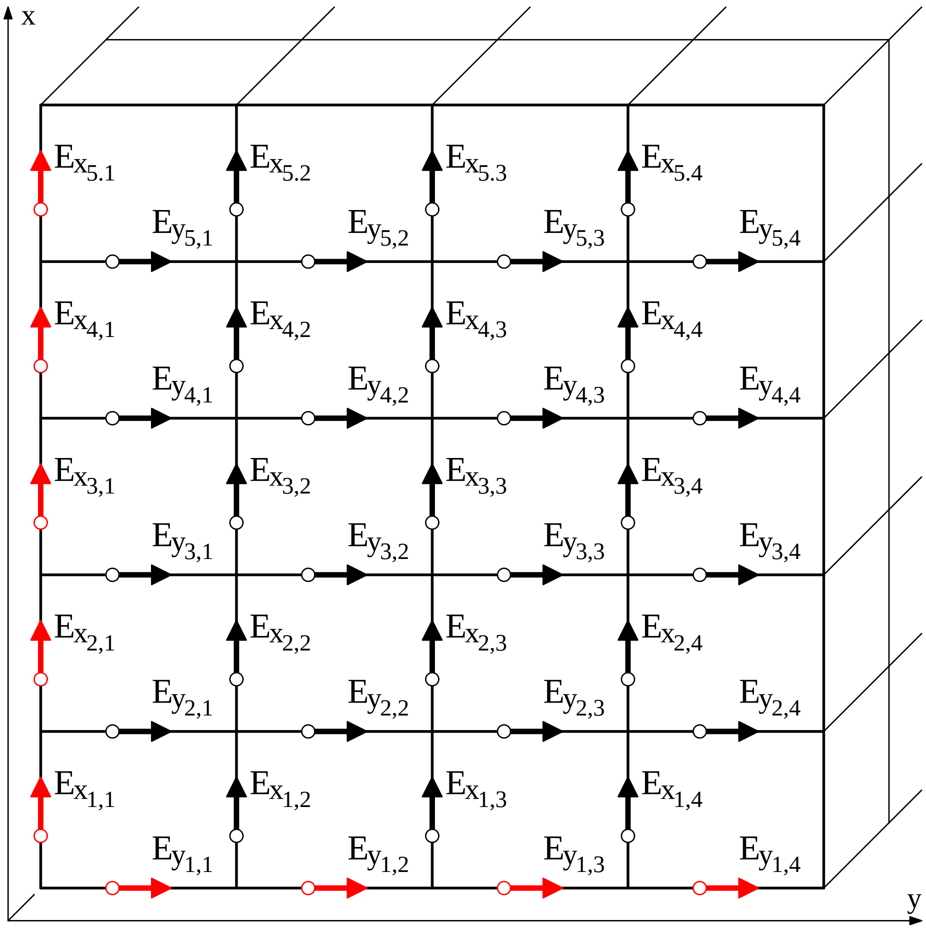

Figure 2: Boundary conditions on the port of the structure

$$
\begin{aligned}
&\ell = (k-1)n_x + i, \quad i = 1(1)n_x, \quad k = 1(1)n_z, \\
&n_{xz} = n_x n_z, \quad \text{and} \quad j = 1 \quad \text{or} \quad j = n_y \ .
\end{aligned} \tag{15}
$$

If the cross-sectional plane is located on the (y,z)-plane, the x-direction is the longitudinal direction of the corresponding transmission line, and we have

$$
\underline{\vec{e}} = (\underline{\vec{e}}_y, \underline{\vec{e}}_z)^T \quad , \quad
\begin{aligned}
\underline{\vec{e}}_y &= (\underline{e}_{y_1}, \underline{e}_{y_2}, \ldots, \underline{e}_{y_{n_{yz}}}) \quad , \quad \underline{e}_{y_\ell} = E_{y_{i,j,k}} , \\
\underline{\vec{e}}_z &= (\underline{e}_{z_1}, \underline{e}_{z_2}, \ldots, \underline{e}_{z_{n_{yz}}}) \quad , \quad \underline{e}_{z_\ell} = E_{z_{i,j,k}}
\end{aligned} \tag{16}
$$

with

$$\begin{aligned} &\ell=(k-1)n_y+j, \quad j=1(1)n_y, \quad k=1(1)n_z, \\ &n_{yz}=n_y n_z, \quad \text{and} \quad i=1 \quad \text{or} \quad i=n_x \ . \end{aligned} \tag{17}$$

We refer to the case (7) in the discussion to follow.
The matrix $\bar{M}$ of the reduced eigenvalue problem (1), (2), (12), (13) depends on h^2 (see (4), (5)):

$$\bar{M}(h^2)\underline{\vec{e}}=\gamma(h)\underline{\vec{e}} \ , \quad \dim(\bar{M})=(2n_{xy}-n_b, 2n_{xy}-n_b). \tag{18}$$

The size of n_b depends on the boundary conditions.
The dependence of $\bar{M}$ on h can be used in numerical solution methods to scale the set of eigenvalues γ of (18).
The sparse matrix $\bar{M}$ is nonsymmetric. There are $2n_{xy}$ eigenvalues $\gamma=u+\jmath v$ and corresponding propagation constants $k_z=\beta-\jmath\alpha$:

$$\gamma_\iota=-4\sin^2(hk_{z_\iota})=u_\iota+\jmath v_\iota, \quad k_{z_\iota}=\beta_\iota-\jmath\alpha_\iota, \quad \iota=1(1)2n_{xy} \ . \tag{19}$$

The propagation constants k_z can be computed from γ after the solution of the eigenvalue problem (1), (2), (12), (13).
We get from (3)

$$k_{z_\iota}=\frac{1}{h}\arcsin\left(\frac{\jmath}{2}\sqrt{\gamma_\iota}\right)=\frac{\jmath}{2h}\ln\left(\frac{\gamma_\iota}{2}+1+\sqrt{\frac{\gamma_\iota}{2}\left(\frac{\gamma_\iota}{2}+2\right)}\right). \tag{20}$$

A propagation constant k_z and its corresponding eigenfunction are called a mode.
In the discussion to follow we denote sets of propagation constants k_z with an indicated $\underline{\wp}$ and the corresponding eigenvalues in accordance with (19) with indicated $\wp$.
Le be $\underline{\wp}_0$ the set of all propagation constants k_z and $\wp_0$ the set of the corresponding eigenvalues γ of $\bar{M}$.
The energy of the complex and evanescent modes decreases exponentially with the distance from the structure. Thus, most of the modes can be neglected within the limit of accuracy. Only some modes (see [10], section 2) are taken into account. The other modes are assumed to vanish within a given distance d between the cross-sectional plane and the discontinuity. Generally speaking, the larger the magnitude of the imaginary part of k_z the stronger the decay. Therefore, the propagation constants k_{z_ι}, $\iota=1(1)2n_{xy}$,

of the set $\underline{\wp}_0$ are sorted in ascending order of $\mid \alpha_\iota \mid$. In the case if some $\mid \alpha_\iota \mid$ have the same value the constants k_{z_ι} are sorted in descending order of $\mid \beta_\iota \mid$. Let be $m_{\overline{s}}$ the number of propagation constants which are able to propagate. These $m_{\overline{s}}$ propagation constants are then the first of the sorted list $\underline{\wp}_0$ and have to be taken into account.
In the former version of the program package F3D the set $\underline{\wp}_0$ of all propagation constants was computed and sorted in order to select the set $\underline{\wp}_f$ of the first $m_{\overline{s}}$ propagation constants of $\underline{\wp}_0$. This way is very time-consuming. The sparse matrix is stored as a dense matrix.
In order to avoid the time-consuming computation of all eigenvalues γ we use an iterative method now, which is carried out twice to find the $m_{\overline{s}}$ propagation constants.
Using the iterative method the computation of the needed propagation constants in a typical example is 40-fold faster than in the old version. The reduction of the memory consumption amounts the 20-fold in the represented example, since the sparse storage technique is applied.
We use the nonsymmetric version of the implicitly restarted Arnoldi iteration [17], [15], [13], [18], [11].
The Arnoldi iteration produces a partial orthogonal factorization of a matrix A of order m into an upper Hessenberg matrix H_r of order r, $r \leq m$. Using the eigenvalues of this small matrix an approximation of a subset of the eigenvalues of the matrix A can be obtained. The approximation of the eigenvalues of H_r to those of A is improved if r increases. But for higher r the algorithm needs more time and storage. To avoid the higher cost and storage requirements the Arnoldi iteration is used with an implicitly re-starting technique. The implicitly re-started Arnoldi iteration can be considered as a truncation of the implicitly shifted QR algorithm.
We can compute a set of eigenvalues of largest or smallest magnitude, real part or imaginary part in different modes with the help of the Arnoldi iteration. The standard eigenvalue problem $Ax = \lambda x$ can be solved using the regular mode $Ax = \lambda x$ or the inverse mode $A^{-1}x = \frac{1}{\lambda}x$ [18]. The Arnoldi algorithm is called iteratively to solve one of these problems generating Arnoldi vectors. Using the regular mode most of the cost in generating each Arnoldi vector is in the matrix-vector product. Using the inverse mode we have to solve a system of linear algebraic equations on each iteration.
The method does not converge in the regular mode for our eigenvalue problem. Thus, the Arnoldi algorithm is called iteratively to solve the standard

eigenvalue problem using the inverse mode $A^{-1}x = \frac{1}{\lambda}x$ with the more time-consuming solution of linear algebraic equations.
The process of solving the systems of linear algebraic equations consists of the factorization of the coefficient matrix and the subsequent forward and backward substitution. Because the matrix A does not change using the Arnoldi iteration, we have to factorize A only once at the beginning of the method. The forward and backward substitution is done on each iteration step.
The coefficient matrix of the system of linear algebraic equations is sparse. We use a direct method, a Gaussian elimination for sparse matrices, [16], [3], [9] to solve the ill-conditioned system of linear equations. The Gaussian elimination is carried out with a special pivoting (Markowitz criterion, [12]) which permits a compromise between a minimum fill-in (increase of the number of non-zero elements during elimination) and numerical stability (connection of topological and numerical pivoting). We use the maximal numerical stability in this criterion.
Thus, using iterative methods we are able to find a subset of the eigenvalues γ of the high-dimensional problem. That is, first we need an estimation $n_1 \geq m_{\overline{s}}$ of the number of propagation constants $m_{\overline{s}}$ which have to be taken into account. The engineer is able to give a good estimation of n_1. Then we compute a subset $\wp_1$ of n_1 eigenvalues γ of smallest magnitude.
The real propagation constants are of course the propagation constants with the smallest magnitude of imaginary part. In most applications one has at least one propagation constant with zero imaginary part. These propagation constants have to be taken into account anyway. It is important, however, to know also some propagation constants of the sorted list which have the property $\mid \alpha \mid > 0$ (see (19)), in order to decide whether the eigenfunctions decrease strong enough in a given distance d. If we find real propagation constants only, we have to choice a greater n_1 for a new computation. If we also find propagation constants with imaginary parts which do not vanish, we can decide whether the eigenfunctions decrease strong enough in a given distance d between the cross-sectional plane and the discontinuity. If necessary, we have to change the distance d in our model.
In generally, we have $\underline{\wp}_f \cap \underline{\wp}_1 \neq \emptyset$, but $\underline{\wp}_f \not\subseteq \underline{\wp}_1$. The relation between γ and k_z is nonlinear. Because of (3) and (19) we have

$$
\begin{aligned}
\gamma(h) &= e^{+\jmath k_z 2h} + e^{-\jmath k_z 2h} - 2 = (e^{+\jmath k_z h} - e^{-\jmath k_z h})^2 \\
&= (e^{+\alpha 2h} + e^{-\alpha 2h})(1 - 2\sin^2(\beta h)) - 2 + \\
&\quad\; 2\jmath(e^{+\alpha 2h} - e^{-\alpha 2h})\sin(\beta h)\cos(\beta h)\ .
\end{aligned} \tag{21}
$$

We consider special values of $\gamma = u + \jmath v$ and $k_z = \beta - \jmath\alpha$.
Because of (21) we have

$$
\begin{array}{lcll}
\alpha &=& 0: & \gamma = -4\sin^2(\beta h), \quad u \le 0, \quad v = 0, \\
\beta &=& 0: & \gamma = e^{+\alpha 2h} + e^{-\alpha 2h} - 2, \quad u \ge 0, \quad v = 0, \\
\beta h &=& \frac{\pi}{2}: & \gamma = -(e^{+\alpha 2h} + e^{-\alpha 2h}) - 2, \quad u \le 0, \quad v = 0.
\end{array} \tag{22}
$$

We will come back to this special values in the discussion to follow.
How we can find the first $m_{\overline{s}}$ propagation constants of the sorted list of $\underline{\wp}_0$ and the appropriate value of the distance d?
We compute the set $\wp_1$ of eigenvalues γ of smallest magnitude using the Arnoldi method in inverse mode looking for eigenvalues of largest magnitude. If we choose h small enough, we have (see (3))

$$
|\,\gamma\,| \;=\; |\,-4\sin^2(hk_z)\,| \;\le\; |\,-4(hk_z)^2\,|\,. \tag{23}
$$

$y = \sin^2 x$ and $y = x$ are monotonically increasing functions in the intervals considered. Thus, the corresponding set of propagation constants $\underline{\wp_1}$ is also a set of propagation constants of smallest magnitude.
However, we have to find a set of propagation constants with the smallest possible magnitude of the imaginary part.
We consider separately the subset $\underline{\wp}_{e_1}$ of imaginary, the subset $\underline{\wp}_{r_1}$ of real and the subset $\underline{\wp}_{c_1}$ of complex propagation constants of the set $\underline{\wp}_1$:

$$
\underline{\wp}_{e_1} \cup \underline{\wp}_{r_1} \cup \underline{\wp}_{c_1} = \underline{\wp}_1, \quad \underline{\wp}_{e_1} \cap \underline{\wp}_{r_1} \cap \underline{\wp}_{c_1} = \emptyset\ . \tag{24}
$$

Imaginary propagation constants ($\beta = 0$)
Because of (23) there cannot exist imaginary propagation constants in $\underline{\wp}_{e_1}$ with a smaller magnitude than we find in the set $\underline{\wp}_1$.
Because of (22) the corresponding eigenvalues in $\wp_1$ are positive real numbers.
Complex propagation constants
Let be $\underline{\wp}_{c_1}$ the subset of complex propagation constants which we find in $\underline{\wp}_1$. Let be k_{z_λ} the complex propagation constant of $\underline{\wp}_{c_1}$ with the smallest magnitude of the imaginary part. The question in this case is:

Is there a subset $\underline{\wp}_{c_2}$, $\underline{\wp}_{c_2} \cap \underline{\wp}_{c_1} = \emptyset$ and $\underline{\wp}_{c_2} \subset \underline{\wp}_0$, of other complex propagation constants with

$$k_{z_\varrho} = \beta_\varrho - \jmath\alpha_\varrho \quad \text{with} \quad |\, \alpha_\varrho \,|\leq|\, \alpha_\lambda \,|, \quad \varrho = 1(1)n_{c_2} \;\; ? \tag{25}$$

But, we have not to decide this question in the lossless case (μ, $\underline{\epsilon}$ real, $\kappa = 0, \infty$), since always evanescent modes exist in this case, and it is sufficient to use the evanescent modes to judge the distance d between the cross-sectional plane and the discontinuity.

Real propagation constants ($\alpha = 0$)

The set $\underline{\wp}_{r_1}$ of real propagation constants of smallest magnitude belong to $\underline{\wp}_1$ (see (24)). The corresponding set $\wp_{r_1}$ of eigenvalues consists of negative real numbers (see (22)):

Now let be $\wp_{r_2}$ the subset

$$\gamma_\varrho = u_\varrho + \jmath v_\varrho \quad \text{with} \quad u_\varrho \leq 0, \quad v_\varrho = 0, \quad \varrho = 1(1)n_{r_2}, \tag{26}$$

of negative eigenvalues of the set $\wp_1$, which is computed.

We note that also complex propagation constants can correspond to negative eigenvalues (see third row of (22)). We have $\wp_{r_1} \subseteq \wp_{r_2} \subseteq \wp_1$. We separate $\wp_{r_2}$ from $\wp_1$ and compute the corresponding set $\underline{\wp}_{r_2}$ of propagation constants

$$k_{z_\varrho} = \frac{1}{h} \arcsin\left(\frac{\jmath}{2}\sqrt{u_\varrho}\right), \quad u_\varrho \leq 0, \quad \varrho = 1(1)n_{r_2} \;. \tag{27}$$

We are interested in the subset $\underline{\wp}^*_{r_2}$ of real propagation constants of the set $\underline{\wp}_{r_2}$ only (see remark after (26)).

Let be $k_{z_{\underline{\varrho}}} = \beta_{\underline{\varrho}}$ the propagation constant of smallest magnitude of $\underline{\wp}^*_{r_2}$.

The question is:

Is there another set $\underline{\wp}_{r_3}$ of propagation constants with $\underline{\wp}_{r_3} \cap \underline{\wp}^*_{r_2} = \emptyset$ and $\underline{\wp}_{r_3} \cap \underline{\wp}_1 = \emptyset$ with

$$k_{z_\varsigma} = \beta_\varsigma - \jmath\alpha_\varsigma, \quad \text{with} \quad \alpha_\varsigma = 0 \quad \text{and} \quad |\, \beta_\varsigma \,|>|\, \beta_{\underline{\varrho}} \,|, \quad \varsigma = 1(1)n_{r_3}? \tag{28}$$

We use the matrix $(-\bar{M})$ rather than $\bar{M}$ to decide this question.

The corresponding set $\wp_{r_3}$ of n_{r_3} eigenvalues consists of negative real numbers (see (22)). The set $\wp_{r_2} \cup \wp_{r_3}$ of negative eigenvalues of $\bar{M}$ corresponds one to one with a set of positive eigenvalues of the same modulus of $(-\bar{M})$.

The matrix $(-\bar{M})$ is extended by some, for example a set $\hbar_1$ of positive elements $\gamma_\tau^{(\bar{M}^*)}$, $\tau = 1(1)n_{\hbar_1}$, adding $n_{\hbar_1}$ rows and $n_{\hbar_1}$ columns such that the elements $\gamma_\tau^{(\bar{M}^*)}$ are diagonal elements of the extended matrix $\bar{M}^*$. The other elements of the new rows and columns are chosen to be zero:

$$\bar{M}^* = \begin{pmatrix} \gamma_1^{(\bar{M}^*)} & & & \\ & \ddots & & \\ & & \gamma_{n_{\hbar_1}}^{(\bar{M}^*)} & \\ & & & \begin{pmatrix} & & \\ & -\bar{M} & \\ & & \end{pmatrix} \end{pmatrix} \tag{29}$$

The spectrum of $\bar{M}^*$ consists of the spectrum of $-\bar{M}$ extended by the added $n_{\hbar_1}$ eigenvalues $\gamma_\tau^{(\bar{M}^*)}$.
How should one choose the $n_{\hbar_1}$ additional eigenvalues?
We can give an upper bound $\mid \overline{\gamma}^{(\bar{M}^*)} \mid$ for the smallest element of the set $\hbar_1$ of the $n_{\hbar_1}$ additional eigenvalues $\gamma_\tau^{(\bar{M}^*)}$ such that all interesting positive eigenvalues of $\bar{M}^*$ are smaller than the bound $\mid \overline{\gamma}^{(\bar{M}^*)} \mid$.
Because of (6) we have

$$k_z \leq \omega\sqrt{\epsilon\mu} = \sqrt{\tilde{\epsilon}\tilde{\mu}}\varkappa_0. \tag{30}$$

The quantities $\tilde{\epsilon}$ and $\tilde{\mu}$ can be different from cell to cell. We select the maxima of this quantities (see Eqn. (16) in [10] for the indices):

$$\tilde{\epsilon}^{(max)} = \max_{i,j,k}\{\tilde{\epsilon}_{i,j,k}\}, \quad \tilde{\mu}^{(max)} = \max_{i,j,k}\{\tilde{\mu}_{i,j,k}\} \ . \tag{31}$$

From (23), (6), (31) and (30) and with

$$c_0 = \frac{1}{\sqrt{\epsilon_0\mu_0}}, \quad c_0 = \lambda_0 f, \quad \omega = 2\pi f \tag{32}$$

we obtain the following estimation

$$\mid \overline{\gamma}^{(\bar{M}^*)} \mid \leq 4\tilde{\epsilon}^{(max)}\tilde{\mu}^{(max)}(\varkappa_0 h)^2 = 4\tilde{\epsilon}^{(max)}\tilde{\mu}^{(max)}\left(\frac{2\pi}{\lambda_0}h\right)^2, \tag{33}$$

with

- c_0 $[\frac{m}{s}]$ velocity of light in vacuo,

- λ_0 $[m]$ wavelength of light in vacuo,

- f $[\frac{1}{s}]$ frequency.

We use the following set $\hbar_1$ of additional eigenvalues

$$\gamma_\tau^{(\bar{M}^*)} = \mid \overline{\gamma}^{(\bar{M}^*)} \mid (1 + \frac{\tau}{10}), \quad \tau = 1(1)n_{\hbar_1} \tag{34}$$

to build $\bar{M}^*$ and compute a set $\hbar_2$ of $n_{\hbar_2} = n_{\hbar_1} + n_{r_2}$ eigenvalues of largest real part of $\bar{M}^*$ using the Arnoldi method in inverse mode. Eigenvalues which fulfill the condition (28) belong to $\hbar_2$ rather than the eigenvalues of $\hbar_1$, and we can separate the $n_{\hbar_3}$ new eigenvalues from the set $\hbar_2$. Now we have $n_{r_2} + n_{\hbar_3}$ negative eigenvalues of $\bar{M}$, and we have to select the corresponding propagation constants which are real.
If we find $n_{\hbar_3} < n_{\hbar_1}$ new eigenvalues, we need not change $n_{\hbar_1}$. Otherwise we have to increase $n_{\hbar_1}$ for a new computation because we do not know whether more than $n_{\hbar_1}$ new negative eigenvalues of $\bar{M}$ exist. If we demand more eigenvalues of largest real part of $\bar{M}^*$ than positive eigenvalues of $\bar{M}^*$ exist, the Arnoldi method does not converge for our problem.

<u>Factorization of $\bar{M}^*$</u>
The factorization of the sparse matrix $\bar{M}^*$ (see (29)) is obviously a modification of the factorization of $\bar{M}$, i.e., we can avoid a second matrix factorization. We have to change the chain lists representing the sparse structure and to reorder the non-zeros of the matrix only.

3 The Numerical Solution of the System of Linear Algebraic Equations

In [10] we explained that the application of the finite-volume method to the three-dimensional boundary value problem for the Maxwellian equations in the frequency domain results in a system of linear equations with indefinite symmetric matrices. We remember the essential points.
The matrix representations of the first and of the second Maxwellian equation are:

$$\oint_{\partial\Omega} \frac{\vec{B}}{\tilde{\mu}} \cdot d\vec{s} = \int_{\Omega} (\jmath\omega\tilde{\epsilon}\epsilon_0\mu_0\vec{E}) \cdot d\vec{\Omega} \quad \Rightarrow \quad A^T D_{s/\tilde{\mu}}\vec{b} = \jmath\omega\epsilon_0\mu_0 D_{A\tilde{\epsilon}}\vec{e} \ , \tag{35}$$

$$\oint_{\partial\Omega} \vec{E} \cdot d\vec{s} = \int_{\Omega} (-\jmath\omega\vec{B}) \cdot d\vec{\Omega} \quad \Rightarrow \quad AD_s\vec{e} = -\jmath\omega D_A\vec{b} \ . \tag{36}$$

$\vec{e}$ and $\vec{b}$ are the vectors of the electric and the magnetic field components, respectively. The matrices are defined in [10], section 6.
Using the representation of $\varkappa_0$ (see (6)) we get from (35) and (36) the matrix representation of the system of linear equations (see [10], Eqns. (30)-(34))

$$Q_1\vec{e} = 0, \quad Q_1 = A^T D_{s/\tilde{\mu}} D_A^{-1} A D_s - \varkappa_0^2 D_{A\tilde{\epsilon}} \ . \tag{37}$$

Taking into account the boundary conditions we get from (37) a partitioning of the matrix Q_1 into the sum of two matrices:

$$Q_1\vec{e} = (Q_{1,A} + Q_{1,r})\vec{e} = 0 \ , \tag{38}$$

where $Q_{1,r}\vec{e}$ is known.
That means we have to solve the following system of linear algebraic equations

$$\tilde{Q}_{1,A}\vec{\tilde{e}} = -\tilde{Q}_{1,r}\vec{\tilde{e}} = -D_s^{\frac{1}{2}} Q_{1,r} D_s^{-\frac{1}{2}} D_s^{\frac{1}{2}}\vec{e} = \vec{\tilde{r}} \tag{39}$$

with

$$\vec{\tilde{r}} = D_s^{\frac{1}{2}}\vec{r} \quad , \quad \vec{r} = -Q_{1,r}\vec{e} \ . \tag{40}$$

As mentioned in [10], we do not solve the linear algebraic system of equations (39), but we substract the gradient of the electric-field divergence (see [10], section 6), which vanishes in our problems,

$$\oint_{\Omega} \tilde{\epsilon}\epsilon_0\vec{E} \cdot d\vec{\Omega} = 0 \quad \Rightarrow \quad BD_{A\tilde{\epsilon}}\vec{e} = 0 \tag{41}$$

from (39). Thus, we get the system of linear equations (see (39), (40) and [10], Eqns. (63) and (64))

$$D_s^{-\frac{1}{2}}\tilde{Q}_A D_s^{\frac{1}{2}}\vec{e} = \vec{r} \ . \tag{42}$$

The motivation for the addition of the gradient of the electric-field divergence is given in the discussion to follow (see also [2], [19], [20]).
Let for simplicity be $\underline{\epsilon}$ and μ constant in contrast to our assumption mentioned after (6). Applying the curl operator to the differential form of the second Maxwellian equation (see [10], Eqn. (13)) gives:

$$\nabla \times (\nabla \times \vec{E}) = -\jmath\omega \nabla \times \vec{B}. \tag{43}$$

Substituting the first Maxwellian equation (see [10], Eqn. (13)) into (43) yields

$$\nabla \times (\nabla \times \vec{E}) - \varkappa^2 \vec{E} = 0, \quad \varkappa = \omega\sqrt{\tilde{\epsilon}\tilde{\mu}\epsilon_0\mu_0}. \tag{44}$$

Using the vector identity

$$\nabla \times (\nabla \times \vec{E}) = \nabla(\nabla \cdot \vec{E}) - (\nabla \cdot \nabla)\vec{E} = \nabla(\nabla \cdot \vec{E}) - \Delta\vec{E} \tag{45}$$

we get from (44) the following form of the wave equation

$$\nabla(\nabla \cdot \vec{E}) - \Delta\vec{E} - \varkappa^2 \vec{E} = 0. \tag{46}$$

Combining the two Equations (44) and (46) yields

$$\nabla(\nabla \cdot \vec{E}) + (\nabla \times (\nabla \times \vec{E}) - \varkappa^2 \vec{E}) = \Delta\vec{E} + \varkappa^2 \vec{E}. \tag{47}$$

The solution of (44) will also fulfill (47).
The first term of the left-hand side of this equation contains the derivative of the third equation $\nabla \cdot \vec{E} = 0$ ($\underline{\epsilon} = const.$!) in ([10], Eqn. (13)), that is, this term vanishes. The second term of the left-hand side of (47) can be interpreted as an analogy of the system of linear algebraic equations (37) if we neglect that $\underline{\epsilon}$ and μ may differ from cell to cell in (37).
The Maxwellian equations also could be solved using a discretization of the wave equation (47) with the restriction that $\underline{\epsilon}$ and μ have to be constant. Thus, an addition of the derivative of the equation $\nabla \cdot \underline{\epsilon}\vec{E} = 0$ to the grid equations (37) refers to the wave equation (47).
We use the equation $\nabla \cdot \underline{\epsilon}\vec{E} = 0$ in the form

$$\underline{\epsilon}\nabla(\frac{1}{\underline{\epsilon}^2}\nabla \cdot \underline{\epsilon}\vec{E}) = 0 \ . \tag{48}$$

The linear system of equations (42) can be solved numerically much faster than the original system (39) [2].

Another physical interpretation of this method is described in [2], [20] taking into consideration the eigenvalues of the characteristic equation

$$\det(F - \lambda I) = 0, \quad F = D_{A\tilde{\epsilon}}^{-1} A^T D_{s/\tilde{\mu}} D_A^{-1} A D_s \ , \tag{49}$$

which is derived multiplying equation (37) with $D_{A\tilde{\epsilon}}^{-1}$,

$$D_{A\tilde{\epsilon}}^{-1} A^T D_{s/\tilde{\mu}} D_A^{-1} A D_s - \varkappa_0^2 I = 0 \ , \tag{50}$$

and taking into account that the term $-\varkappa_0^2 I$, I identity matrix, causes only a shift.
Let us consider how to compute $\underline{\epsilon}\nabla(\frac{1}{\underline{\epsilon}^2}\nabla \cdot \underline{\epsilon}\vec{E}) = 0$.
We denote the divergence of $\underline{\epsilon}\vec{E}$ with U:

$$U(E_x, E_y, E_z) = \frac{1}{\underline{\epsilon}^2}\nabla \cdot \underline{\epsilon}\vec{E}, \tag{51}$$

that is, we have to compute the gradient of U:

$$\underline{\epsilon}\nabla U = \underline{\epsilon}\nabla(\frac{1}{\underline{\epsilon}^2}\nabla \cdot \underline{\epsilon}\vec{E}). \tag{52}$$

The gradient of a scalar can be defined as

$$\nabla U = \lim_{V \to 0} \frac{\oint\limits_{\Omega} U \, d\vec{\Omega}}{V} \quad , \tag{53}$$

where Ω is the closed surface of the volume V.
Using Cartesian coordinates the surface integral from (53) can be computed as follow:

$$\int\limits_{\Omega} U \, d\vec{\Omega} = \int\limits_{\Omega_{yz}} U \, dy \, dz \, \vec{i}_x + \int\limits_{\Omega_{zx}} U \, dz \, dx \, \vec{i}_y + \int\limits_{\Omega_{xy}} U \, dx \, dy \, \vec{i}_z \ . \tag{54}$$

$\vec{i}_x$, $\vec{i}_y$ and $\vec{i}_z$ are the unit vectors in x-, y- and z-direction, respectively.
We consider the surface integral at the point (i, j, k) (see Figure 3):

$$\left[\oint_{\Omega} U\, d\vec{\Omega}\right]_{(i,j,k)} = \left(\left[\oint_{\Omega_{yz}} U\, dy\, dz\right]_{(i,j,k)} - \left[\oint_{\Omega_{yz}} U\, dy\, dz\right]_{(i+1,j,k)}\right)\vec{\imath}_x + \left(\left[\oint_{\Omega_{zx}} U\, dz\, dx\right]_{(i,j,k)} - \left[\oint_{\Omega_{zx}} U\, dz\, dx\right]_{(i,j+1,k)}\right)\vec{\imath}_y + \left(\left[\oint_{\Omega_{xy}} U\, dx\, dy\right]_{(i,j,k)} - \left[\oint_{\Omega_{xy}} U\, dx\, dy\right]_{(i,j,k+1)}\right)\vec{\imath}_z \ . \tag{55}$$

The directional derivative of a scalar function W with respect to the distance L along an unit tangent $\vec{\tau}$ of a curve is

$$\frac{dW}{dL} = (\nabla W) \cdot \vec{\tau} \ . \tag{56}$$

Thus, we get with $W = U$ and $\vec{\tau} = (dx_i, dy_j, dz_k)^T$ from (53) and (55)

$$(\nabla U)_{(i,j,k)} = \frac{1}{dx_i}\left(\left[\frac{\oint_{\Omega_{yz}} U\, dy\, dz}{V}\right]_{(i,j,k)} - \left[\frac{\oint_{\Omega_{yz}} U\, dy\, dz}{V}\right]_{(i+1,j,k)}\right)\vec{\imath}_x + \frac{1}{dy_j}\left(\left[\frac{\oint_{\Omega_{zx}} U\, dz\, dx}{V}\right]_{(i,j,k)} - \left[\frac{\oint_{\Omega_{zx}} U\, dz\, dx}{V}\right]_{(i,j+1,k)}\right)\vec{\imath}_y + \frac{1}{dz_k}\left(\left[\frac{\oint_{\Omega_{xy}} U\, dx\, dy}{V}\right]_{(i,j,k)} - \left[\frac{\oint_{\Omega_{xy}} U\, dx\, dy}{V}\right]_{(i,j,k+1)}\right)\vec{\imath}_z \ . \tag{57}$$

The volume $V_{i,j,k}$ consists of 8 partial volumes (see Figure 3), and we have

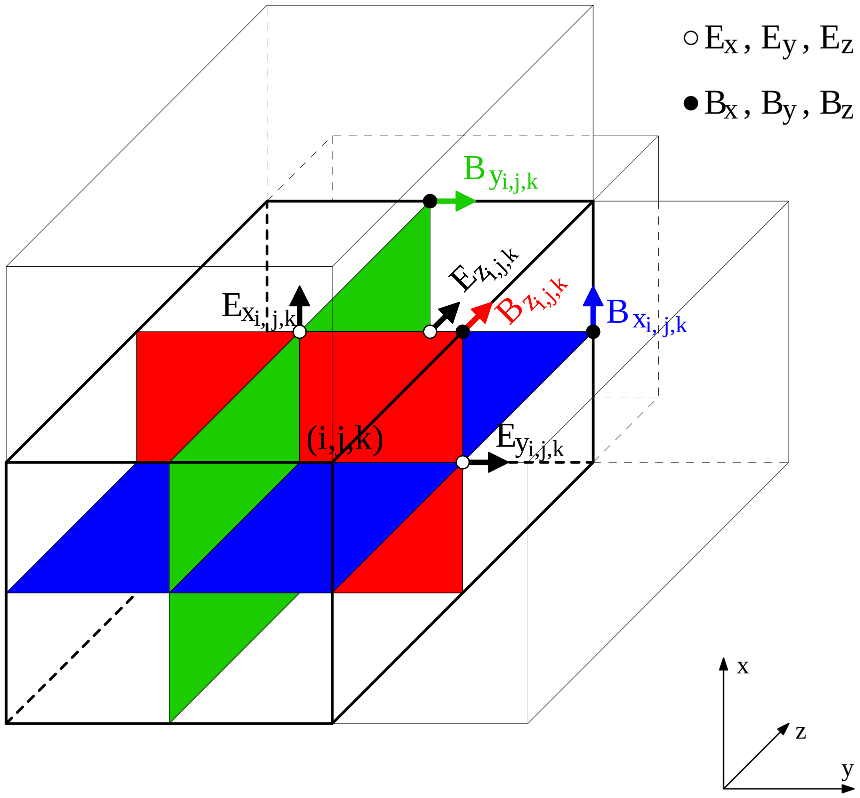

Figure 3: The dual elementary cell used for the computation of $\nabla(\nabla \cdot \epsilon \vec{E})$

$$
\begin{aligned}
V_{i,j,k} = \tfrac{1}{8}(\quad & x_{i-1,j-1,k-1} y_{i-1,j-1,k-1} z_{i-1,j-1,k-1} \epsilon^2_{i-1,j-1,k-1} + \\
& x_{i,j-1,k-1} y_{i,j-1,k-1} z_{i,j-1,k-1} \epsilon^2_{i,j-1,k-1} + \\
& x_{i-1,j,k-1} y_{i-1,j,k-1} z_{i-1,j,k-1} \epsilon^2_{i-1,j,k-1} + \\
& x_{i,j,k-1} y_{i,j,k-1} z_{i,j,k-1} \epsilon^2_{i,j,k-1} + \\
& x_{i-1,j-1,k} y_{i-1,j-1,k} z_{i-1,j-1,k} \epsilon^2_{i-1,j-1,k} + \\
& x_{i,j-1,k} y_{i,j-1,k} z_{i,j-1,k} \epsilon^2_{i,j-1,k} + \\
& x_{i-1,j,k} y_{i-1,j,k} z_{i-1,j,k} \epsilon^2_{i-1,j,k} + x_{i,j,k} y_{i,j,k} z_{i,j,k} \epsilon^2_{i,j,k}) \ .
\end{aligned} \tag{58}
$$

Using the denotations (26), (27) and (28) from [10] we obtain for U

$$\begin{aligned} U_{i,j,k} \;=\;& g^{y,z}_{i,j,k} E_{x_{i,j,k}} - g^{y,z}_{i-1,j,k} E_{x_{i-1,j,k}} + g^{x,z}_{i,j,k} E_{y_{i,j,k}} - \\ & g^{x,z}_{i,j-1,k} E_{y_{i,j-1,k}} + g^{x,y}_{i,j,k} E_{z_{i,j,k}} - g^{x,y}_{i,j,k-1} E_{z_{i,j,k-1}} \quad . \end{aligned} \tag{59}$$

Thus, for the components of ∇U along the unit vectors i_x, i_y and i_y one has

$$\begin{aligned} (\tilde{\epsilon}\nabla U)_{x_{i,j,k}} =& \frac{1}{x_{i,j,k}} \left(\frac{U_{i,j,k} g^{y,z}_{i,j,k}}{V_{i,j,k}} - \frac{U_{i+1,j,k} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} \right) \\ =& \frac{1}{x_{i,j,k}} \left(-\frac{g^{y,z}_{i-1,j,k} g^{y,z}_{i,j,k}}{V_{i,j,k}} E_{x_{i-1,j,k}} + \left(\frac{g^{y,z}_{i,j,k} g^{y,z}_{i,j,k}}{V_{i,j,k}} + \frac{g^{y,z}_{i,j,k} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} \right) E_{x_{i,j,k}} - \right. \\ & \frac{g^{y,z}_{i+1,j,k} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} E_{x_{i+1,j,k}} - \frac{g^{x,z}_{i,j-1,k} g^{y,z}_{i,j,k}}{V_{i,j,k}} E_{y_{i,j-1,k}} + \frac{g^{x,z}_{i+1,j-1,k} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} E_{y_{i+1,j-1,k}} + \\ & \frac{g^{x,z}_{i,j,k} g^{y,z}_{i,j,k}}{V_{i,j,k}} E_{y_{i,j,k}} - \frac{g^{x,z}_{i+1,j,k} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} E_{y_{i+1,j,k}} - \frac{g^{x,y}_{i,j,k-1} g^{y,z}_{i,j,k}}{V_{i,j,k}} E_{z_{i,j,k-1}} + \\ & \left. \frac{g^{x,y}_{i+1,j,k-1} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} E_{z_{i+1,j,k-1}} + \frac{g^{x,y}_{i,j,k} g^{y,z}_{i,j,k}}{V_{i,j,k}} E_{z_{i,j,k}} - \frac{g^{x,y}_{i+1,j,k} g^{y,z}_{i,j,k}}{V_{i+1,j,k}} E_{z_{i+1,j,k}} \right) \quad , \end{aligned} \tag{60}$$

$$\begin{aligned} (\tilde{\epsilon}\nabla U)_{y_{i,j,k}} =& \frac{1}{y_{i,j,k}} \left(\frac{U_{i,j,k} g^{x,z}_{i,j,k}}{V_{i,j,k}} - \frac{U_{i,j+1,k} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} \right) \\ =& \frac{1}{y_{i,j,k}} \left(-\frac{g^{y,z}_{i-1,j,k} g^{x,z}_{i,j,k}}{V_{i,j,k}} E_{x_{i-1,j,k}} + \frac{g^{y,z}_{i,j,k} g^{x,z}_{i,j,k}}{V_{i,j,k}} E_{x_{i,j,k}} + \right. \\ & \frac{g^{y,z}_{i-1,j+1,k} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} E_{x_{i-1,j+1,k}} - \frac{g^{y,z}_{i,j+1,k} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} E_{x_{i,j+1,k}} - \\ & \frac{g^{x,z}_{i,j-1,k} g^{x,z}_{i,j,k}}{V_{i,j,k}} E_{y_{i,j-1,k}} + \left(\frac{g^{x,z}_{i,j,k} g^{x,z}_{i,j,k}}{V_{i,j,k}} + \frac{g^{x,z}_{i,j,k} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} \right) E_{y_{i,j,k}} - \\ & \frac{g^{x,z}_{i,j+1,k} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} E_{y_{i,j+1,k}} - \frac{g^{x,y}_{i,j,k-1} g^{x,z}_{i,j,k}}{V_{i,j,k}} E_{z_{i,j,k-1}} + \\ & \left. \frac{g^{x,y}_{i,j+1,k-1} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} E_{z_{i,j+1,k-1}} + \frac{g^{x,y}_{i,j,k} g^{x,z}_{i,j,k}}{V_{i,j,k}} E_{z_{i,j,k}} - \frac{g^{x,y}_{i,j+1,k} g^{x,z}_{i,j,k}}{V_{i,j+1,k}} E_{z_{i,j+1,k}} \right) \quad , \end{aligned} \tag{61}$$

$$
\begin{aligned}
(\tilde{\epsilon}\nabla U)_{z_{i,j,k}} &= \frac{1}{z_{i,j,k}}\left(\frac{U_{i,j,k}g^{x,y}_{i,j,k}}{V_{i,j,k}} - \frac{U_{i,j,k+1}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}\right) \\
&= \frac{1}{z_{i,j,k}}\left(-\frac{g^{y,z}_{i-1,j,k}g^{x,y}_{i,j,k}}{V_{i,j,k}}E_{x_{i-1,j,k}} + \frac{g^{y,z}_{i,j,k}g^{x,y}_{i,j,k}}{V_{i,j,k}}E_{x_{i,j,k}} + \right. \\
&\quad \frac{g^{y,z}_{i-1,j,k+1}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}E_{x_{i-1,j,k+1}} - \frac{g^{y,z}_{i,j,k+1}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}E_{x_{i,j,k+1}} - \\
&\quad \frac{g^{x,z}_{i,j-1,k}g^{x,y}_{i,j,k}}{V_{i,j,k}}E_{y_{i,j-1,k}} + \frac{g^{x,z}_{i,j,k}g^{x,y}_{i,j,k}}{V_{i,j,k}}E_{y_{i,j,k}} + \frac{g^{x,z}_{i,j-1,k+1}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}E_{y_{i,j-1,k+1}} - \qquad (62) \\
&\quad \frac{g^{x,z}_{i,j,k+1}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}E_{y_{i,j,k+1}} - \frac{g^{x,y}_{i,j,k-1}g^{x,y}_{i,j,k}}{V_{i,j,k}}E_{z_{i,j,k-1}} + \\
&\quad \left.\left(\frac{g^{x,y}_{i,j,k}g^{x,y}_{i,j,k}}{V_{i,j,k}} + \frac{g^{x,y}_{i,j,k}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}\right)E_{z_{i,j,k}} - \frac{g^{x,y}_{i,j,k+1}g^{x,y}_{i,j,k}}{V_{i,j,k+1}}E_{z_{i,j,k+1}}\right) \quad .
\end{aligned}
$$

The Equations (60, 61, 62) describe the matrix Q_2 (see [10], Eqn. (56)). We give now a short summary of the solution of the system of linear equations (42).

Some symbols:

$A(N,N)$	matrix containing the entries of the first and second Maxwellian equation (see [10], Eqn. (13)), $A = \tilde{Q}_{1,A}$ (39).
$B(N,N)$	matrix containing the entries of the third equation in (see [10], Eqn. (13)), $B = \tilde{Q}_{2,A}$ (see [10], Eqn. (63)).
$M(N,N)$	preconditioner for the matrix A.
$x(N)$	vector containing the electric field of the elementary cells, $x \equiv \vec{\tilde{e}} = D_s^{\frac{1}{2}}\vec{e}$ (see [10], Eqn. (49)).
$b(N)$	right-hand side, $b \equiv \vec{\tilde{r}} = D_s^{\frac{1}{2}}\vec{r}$, $\vec{r} = -Q_{1,r}\vec{e}$ (see [10], Eqn. (51)).
N	number of equations, $N = 3n_{xyz}$.

We consider the solution of the non-singular system of N linear algebraic equations

$$Ax = b \quad . \tag{63}$$

Iterative methods for solving system (63) are attractive because their complexity is $\mathcal{O}(N^2)$, provided the number of iterations required for convergence

is small compared to N, whilst direct methods like LU and Cholesky decompositions are $\mathcal{O}(N^3)$, which is prohibitive for very large N.
The convergence rate of iterative methods depends on spectral properties of the coefficient matrix A of (63). Hence one may attempt to transform the linear system (63) into one that is equivalent in the sense that it has the same solution but more favorable spectral properties. A preconditioner is a matrix that performs such a transformation.
For instance, if a matrix M approximates the coefficient matrix A (63) in some way, the transformed system

$$M^{-1}Ax = M^{-1}b \tag{64}$$

has the same solution as the original system (63), but the spectral properties of its coefficient matrix $M^{-1}A$ may be more favorable.
In devising a preconditioner, we are faced with the choice between finding a matrix M that approximates A, and for which solving a system is easier than solving one with A, or finding a matrix M that approximates A^{-1}, so that only multiplication by M is needed. The majority of preconditioners falls in the first category.
The above transformation of the linear system (63) to (64) is not what is used in practice. A more correct way of introducing the preconditioner would be to split the preconditioner as $M = M_1 M_2$ and to transform the system as

$$M_1^{-1}AM_2^{-1}(M_2x) = M_1^{-1}b \ . \tag{65}$$

The matrices M_1 and M_2 are called the left- and right-hand preconditioners, respectively.
Using the linear system $Bx = 0$ ($B = \tilde{Q}_{2,A}$, see [10], Eqn. (63)) we construct a preconditioner M for the original system (63):

$$M^{-1} = I + BA^{-1} \quad , \quad M = (I + BA^{-1})^{-1} \ . \tag{66}$$

M^{-1} is a symmetric matrix. Equations (64) and (66) can be combined to (67) by substituting the matrix M^{-1}:

$$\begin{aligned} (I + BA^{-1})Ax &= (I + BA^{-1})b, \\ Ax + BA^{-1}Ax &= b + BA^{-1}b, \\ (A + B)x &= b + Bx = b \ . \end{aligned} \tag{67}$$

We get the system of linear equations:

$$\tilde{A}\tilde{x} = \tilde{b} \tag{68}$$

with

$$\tilde{A} = A + B, \quad \tilde{x} = x, \quad \tilde{b} = b \ . \tag{69}$$

The Equation (69) has more favorable spectral properties as the original system (63).
We transform Equation (69) into Equation (80) and into Equation (83), respectively.

1. Let be

$$\tilde{D}_{\tilde{A}} = diag(\tilde{A}) \tag{70}$$

the diagonal matrix of $\tilde{A}$ (69), then we set

$$\tilde{M}_1 = \tilde{D}_{\tilde{A}}^{\frac{1}{2}} \quad , \quad \tilde{M}_2 = \tilde{D}_{\tilde{A}}^{\frac{1}{2}} \tag{71}$$

and combine (71) and (70) with the Equations (65) and (69):

$$\tilde{D}_{\tilde{A}}^{-\frac{1}{2}} \tilde{A} \tilde{D}_{\tilde{A}}^{-\frac{1}{2}} \tilde{D}_{\tilde{A}}^{\frac{1}{2}} \tilde{x} = \tilde{D}_{\tilde{A}}^{-\frac{1}{2}} \tilde{b} \ . \tag{72}$$

We obtain the system of linear equations:

$$\hat{A}\hat{x} = \hat{b} \tag{73}$$

with

$$\hat{A} = \tilde{D}_{\tilde{A}}^{-\frac{1}{2}} \tilde{A} \tilde{D}_{\tilde{A}}^{-\frac{1}{2}} \quad , \quad \hat{x} = \tilde{D}_{\tilde{A}}^{\frac{1}{2}} \tilde{x} \quad , \quad \hat{b} = \tilde{D}_{\tilde{A}}^{-\frac{1}{2}} \tilde{b} \ . \tag{74}$$

We construct a preconditioner for the matrix $\hat{A}$ (73) with (74). The SSOR preconditioner can be derived from the coefficient matrix without any additional effort. If the original, symmetric, matrix (73) with (74) is decomposed as

$$\hat{A} = I + L + L^T \tag{75}$$

in its diagonal, strict lower, and strict upper triangular part, the SSOR matrix [1] is defined as

$$\hat{M} = (I + L)(I + L^T) \quad , \tag{76}$$

or, parameterized by ω

$$\hat{M}_\omega = (I + \omega L)(I + \omega L^T) \tag{77}$$

with $0 < \omega < 2$.
The optimal value of the ω parameter will reduce the number of iterations to a lower order.
The SSOR matrix is given in factorized form, so this preconditioner shares many properties of other factorization-based methods. For instance, its suitability for vector processors or parallel architectures depends strongly on the ordering of the variables. On the other hand, since this factorization is given a priori, there is no possibility of breakdown as in the construction phase of incomplete factorization methods.
The preconditioner $\hat{M}_\omega$ (77) is a symmetric matrix:

$$\hat{M}_\omega^T = ((I + \omega L)(I + \omega L^T))^T = (I + \omega L)(I + \omega L^T) = \hat{M}_\omega \ , \qquad (78)$$

and positive definite:

$$\begin{array}{rcl} (\hat{M}_\omega y, y) & = & ((I + \omega L)(I + \omega L^T)y, y) \\ & = & ((I + \omega L^T)y, (I + \omega L^T)y) \\ & = & (z, z) \\ & \geq & 0 \end{array} \qquad (79)$$

with $z = (I + \omega L^T)y$.
We have to solve the system of linear equations

$$\hat{M}_\omega^{-1} \hat{A} \hat{x} = \hat{M}_\omega^{-1} \hat{b} \ . \qquad (80)$$

2. A commonly used approach for solving large sparse linear systems is to resort to the general multi-color orderings. The simplest form of the general multi-color orderings is the red-black ordering. One may use a coloring scheme for reordering the unknowns, so that unknowns with the same color are not explicitly coupled.
   We first multi-color the matrix $\tilde{A}$ (68), and then we permute the matrix according to the multicolor ordering to get a matrix in the form

$$\tilde{P} \tilde{A} \tilde{P}^T = \begin{pmatrix} \tilde{D} & \tilde{F} \\ \tilde{E} & \tilde{H} \end{pmatrix} \ , \qquad (81)$$

   where $\tilde{D}$ is a diagonal matrix and $\tilde{F} = \tilde{E}^T$. It is clear that the numbering of the colors is arbitrary. We can select the color yielding the largest color set to be first color. Finally, we reduce the system by eliminating the unknowns of the first color, to get the reduced matrix via the formula.

$$\breve{A} = \tilde{H} - \tilde{E}\tilde{D}^{-1}\tilde{E}^T \quad . \tag{82}$$

The transformation used in the elimination process, i.e., the $\tilde{E}$, need not be saved but the upper $\tilde{E}^T$ matrix must be since it will be used to solve for the remaining equation. Regarding the reordering, we can either permute the matrix or simply keep a permutation array. The process can be continued recursively a few times.

One of the advantages of this approach comes from the fact that the last matrix for which we must solve a linear system (82) is usually much smaller than the original matrix (68). This, supported by the fact that it also usually takes fewer steps for the higher-level iterations to converge makes the scheme quite attractive. The drawbacks are its complexity and the fact that the reduced matrices may become quite dense.

We permute the system (68) according to the multicolor ordering (81) into the form

$$\tilde{P}\tilde{A}\tilde{P}^T\breve{x} = \begin{pmatrix} \tilde{D} & \tilde{E}^T \\ \tilde{E} & \tilde{H} \end{pmatrix} \begin{pmatrix} \breve{x}_1 \\ \breve{x}_2 \end{pmatrix} = \begin{pmatrix} \breve{b}_1 \\ \breve{b}_2 \end{pmatrix} \tag{83}$$

with

$$\breve{x} = \tilde{P}\tilde{x} = \begin{pmatrix} \breve{x}_1 \\ \breve{x}_2 \end{pmatrix} \quad , \quad \breve{b} = \tilde{P}\tilde{b} = \begin{pmatrix} \breve{b}_1 \\ \breve{b}_2 \end{pmatrix} \quad . \tag{84}$$

Using (82) we have to solve the system of linear equations

$$\breve{A}\breve{x}_2 = (\tilde{H} - \tilde{E}\tilde{D}^{-1}\tilde{E}^T)\breve{x}_2 = \breve{b}_2 - \tilde{E}\tilde{D}^{-1}\breve{b}_1 \tag{85}$$

for $\breve{x}_2$ (84).

Thus, we get

$$\breve{x}_1 = \tilde{D}^{-1}(\breve{b}_1 - \tilde{E}^T\breve{x}_2) \quad . \tag{86}$$

Then we have to permute the solution vector $\breve{x}$ (85, 86) back to the original ordering $\tilde{x}$ (68).

The Equation (85) are to be transformed (see (80)) into equation

$$\breve{M}_\omega^{-1}\breve{A}\breve{x}_2 = \breve{M}_\omega^{-1}(\breve{b}_2 - \tilde{E}\tilde{D}^{-1}\breve{b}_1) \quad . \tag{87}$$

The Equations (80) and (87) are to be solved with algorithms described in [14], [8], [7].

4 The Computation of the Scattering Matrix

The transverse mode fields $\vec{E}_{t,l} = \vec{E}_{t,l}(z)$ satisfy the orthogonality relation [6]

$$\int\limits_{\Omega} (\vec{E}_{t,l} \times \vec{H}_{t,m}) \cdot d\vec{\Omega} = \eta_m \delta_{l,m}, \tag{88}$$

where $\delta_{l,m}$ is the Kronecker symbol. The mode fields $\vec{E}_{t,l}$ and $\vec{H}_{t,m}$ are generally complex. The amplitudes of the transverse electric fields are normalized such that

$$|\ \eta_m \ |= 1[W], \quad m = 1(1)m^{(p)}, \quad p = 1(1)\overline{p} \ . \tag{89}$$

We will see that we can calculate the scattering matrix by means of the orthogonality relation (88).
Let be

$$\vec{\tilde{E}}_t(z) = a_l \vec{E}_{t,l} e^{-\jmath k_{z_l} z} \tag{90}$$

any excited mode which propagates in the positive z-direction. a_l is the amplitude at $z = 0$. k_{z_l} is the propagation constant (see [10], section 8, and section 2 of this paper).
In the discussion to follow we consider all modes (exciting modes with amplitudes a_l in positive z-direction and outgoing modes with amplitudes b_l in negative z-direction). The transverse electric field for a fixed cross-sectional plane at z is

$$\vec{E}_t(z) = \sum_{l=1}^{\bar{m}} a_l \vec{E}_{t,l} e^{-\jmath k_{z_l} z} + \sum_{l=1}^{\bar{m}} b_l \vec{E}_{t,l} e^{+\jmath k_{z_l} z} = \sum_{l=1}^{\bar{m}} w_l(z) \vec{E}_{t,l} \tag{91}$$

with

$$w_l(z) = a_l e^{-\jmath k_{z_l} z} + b_l e^{+\jmath k_{z_l} z} = \tilde{a}_l(z) + \tilde{b}_l(z). \tag{92}$$

For $z = z_p$ we get from (91) with $\vec{E}_t(z_p) = \vec{E}_t^{(p)}$, $\bar{m} = m^{(p)}$, $k_{z_l} = k_{z_l}^{(p)}$, $\vec{E}_{t,l} = \vec{E}_{t,l}^{(p)}$ and $w_l(z_p) = w_l^{(p)}$ the boundary condition ([10], Eqn. (14)).
Using the orthogonality relation (88) for (91) gives

$$\int\limits_{\Omega}(\vec{E}_t(z)\times\vec{H}_{t,m})\cdot d\vec{\Omega}=\sum_{l=1}^{\bar{m}}w_l(z)\int\limits_{\Omega}(\vec{E}_{t,l}\times\vec{H}_{t,m})\cdot d\vec{\Omega}=w_m(z)\eta_m. \tag{93}$$

The application of (93) with (92) at a pair of two neighboring cross-sectional planes p and $p+\Delta p$ which cut the axis at $z=z_p$ and at $z=z_p+\Delta z_p=z_{p+\Delta p}$, respectively, gives

$$\begin{aligned}
\frac{1}{\eta_m}\int\limits_{\Omega}(\vec{E}_t^{(p)}\times\vec{H}_{t,m}^{(p)})\cdot d\vec{\Omega} &= \tilde{a}_m^{(p)}+\tilde{b}_m^{(p)} &&= w_m^{(p)},\\
\frac{1}{\eta_m}\int\limits_{\Omega}(\vec{E}_t^{(p+\Delta p)}\times\vec{H}_{t,m}^{(p+\Delta p)})\cdot d\vec{\Omega} &= \tilde{a}_m^{(p+\Delta p)}+\tilde{b}_m^{(p+\Delta p)} &&= w_m^{(p+\Delta p)}
\end{aligned} \tag{94}$$

with

$$\begin{aligned}
\vec{M}^{(q)} &= \vec{M}(z_q), && \vec{M}\in\{\vec{E}_t,\vec{H}_{t,m}\},\\
d^{(q)} &= d(z_q), && d\in\{\tilde{a}_m,\tilde{b}_m,\tilde{w}_m\}, \qquad q\in\{p,p+\Delta p\}.
\end{aligned} \tag{95}$$

Because the transmission lines are assumed to be longitudinally homogeneous, we have

$$\vec{H}_{t,m}^{(p+\Delta p)}=\vec{H}_{t,m}^{(p)}. \tag{96}$$

Thus, from the second equation of (94) one obtains

$$\frac{1}{\eta_m}\int\limits_{\Omega}(\vec{E}_t^{(p+\Delta p)}\times\vec{H}_{t,m}^{(p)})\cdot d\vec{\Omega}=\tilde{a}_m^{(p+\Delta p)}+\tilde{b}_m^{(p+\Delta p)}=w_m^{(p+\Delta p)}\ . \tag{97}$$

What is known in the first equation of (94)?
The values of the weighted mode amplitude sums $w_m^{(p)}$ are given (see the discussion to follow). $\vec{E}_t^{(p)}$ is known solving the eigenvalue problem. $H_{t,m}^{(p)}$ can be computed from the known electric field of mode m (see section 5). The numerical calculation of the orthogonality relation is treated in section 5.
Thus, the normalization constant η_m can be computed by using the first equation of (94).

The weighted mode amplitude sums $w_m^{(p+\Delta p)}$ are computed solving the boundary value problem ([10], Eqns. (13), (14) and (15)) evaluating the orthogonality relation (97) (see section 5).
The scattering matrix is defined with $\tilde{a}_m^{(p)}$ and $\tilde{b}_m^{(p)}$ (see [10], section 2).
Because the weighted mode amplitude sums $w_m^{(p)}$ and $w_m^{(p+\Delta p)}$ are known now, the amplitudes $\tilde{a}_m^{(p)}$ and $\tilde{b}_m^{(p)}$ can be determined by means of Equations (94) and (97). Because of (92) we have

$$\begin{aligned} \tilde{a}_m^{(p+\Delta p)} &= \tilde{a}_m^{(p)} e^{-\jmath k_{z_l}^{(p)} \Delta z_p}, \\ \tilde{b}_m^{(p+\Delta p)} &= \tilde{b}_m^{(p)} e^{+\jmath k_{z_l}^{(p)} \Delta z_p}. \end{aligned} \tag{98}$$

Using (98) we eliminate $\tilde{a}_m^{(p+\Delta p)}$ and $\tilde{b}_m^{(p+\Delta p)}$ in (94) and obtain

$$\begin{aligned} \tilde{a}_m^{(p)} &= \frac{w_m^{(p)} e^{+\jmath k_{zm}^{(p)} \Delta z_p} - w_m^{(p+\Delta p)}}{e^{+\jmath k_{zm}^{(p)} \Delta z_p} - e^{-\jmath k_{zm}^{(p)} \Delta z_p}} \quad , \\ \tilde{b}_m^{(p)} &= \frac{w_m^{(p+\Delta p)} - w_m^{(p)} e^{-\jmath k_{zm}^{(p)} \Delta z_p}}{e^{+\jmath k_{zm}^{(p)} \Delta z_p} - e^{-\jmath k_{zm}^{(p)} \Delta z_p}} \quad . \end{aligned} \tag{99}$$

With $\tilde{a}_m^{(p)}$ and $\tilde{b}_m^{(p)}$ reflection coefficients are defined:

$$r_m^{(p)} = \frac{\tilde{b}_m^{(p)}}{\tilde{a}_m^{(p)}} \quad . \tag{100}$$

Using (99) the reflection coefficients are written as

$$r_m^{(p)} = \frac{e^{-\jmath k_{zm}^{(p)} \Delta z_p} - \frac{w_m^{(p+\Delta p)}}{w_m^{(p)}}}{\frac{w_m^{(p+\Delta p)}}{w_m^{(p)}} - e^{+\jmath k_{zm}^{(p)} \Delta z_p}} \quad . \tag{101}$$

The reflection coefficients are computed for all modes $\rho = 1(1)m_{\overline{s}}$ (see [10], Eqn. (3)) and all excitations $\nu = 1(1)m_{\overline{s}}$. The linear independent excitations are given and can be described by the vectors

$$\vec{\bar{w}}_\nu = (\bar{w}_{1,\nu}, \ldots, \bar{w}_{\rho,\nu}, \ldots, \bar{w}_{m_{\overline{s}},\nu})^T, \quad \nu = 1(1)m_{\overline{s}}, \tag{102}$$

with, for example

$$w_m^{(p)} = 1.0, \quad m = 1(1)m^{(p)}, \quad p = 1(1)\overline{p}, \tag{103}$$

$$\bar{w}_{\rho,\nu} = \begin{cases} |w_m^{(p)}| & \text{for} \quad 1 \le \rho \le m_{\overline{s}} + 1 - \nu \\ -|w_m^{(p)}| & \text{for} \quad m_{\overline{s}} + 2 - \nu \le \rho \le m_{\overline{s}} \end{cases}, \quad \rho = m + \sum_{q=1}^{p-1} m^{(q)}. \quad (104)$$

This choice of $\vec{\bar{w}}_\nu$ guarantees that the excitations are linearly independent. With this choice of $\bar{w}_{\rho,\nu}$ the vectors $\vec{\bar{r}}_\nu$, $\vec{\bar{a}}_\nu$ and $\vec{\bar{b}}_\nu$ are built up analogously (see (99), (100) and (101)):

$$\begin{array}{rclcrcl} \vec{\bar{r}}_\nu & = & (\bar{r}_{1,\nu}, \ldots, \bar{r}_{\rho,\nu}, \ldots , \bar{r}_{m_{\overline{s}},\nu})^T, & & \bar{r}_{\rho,\nu} & = & r_m^{(p)}, \\ \vec{\bar{a}}_\nu & = & (\bar{a}_{1,\nu}, \ldots, \bar{a}_{\rho,\nu}, \ldots , \bar{a}_{m_{\overline{s}},\nu})^T, & & \bar{a}_{\rho,\nu} & = & \tilde{a}_m^{(p)}, \\ \vec{\bar{b}}_\nu & = & (\bar{b}_{1,\nu}, \ldots, \bar{b}_{\rho,\nu}, \ldots , \bar{b}_{m_{\overline{s}},\nu})^T, & & \bar{b}_{\rho,\nu} & = & \tilde{b}_m^{(p)}, \end{array} \quad (105)$$

with

$$\rho = m + \sum_{q=1}^{p-1} m^{(q)}. \quad (106)$$

That means, we have to solve $m_{\overline{s}}$ boundary value problems ([10], Eqns. (13), (15)) with the boundary condition (see [10], Eqn. (14))

$$\vec{E}_{t,\nu} = \sum_{\rho=1}^{m_{\overline{s}}} \bar{w}_{\rho,\nu} \vec{E}_{t,l}^{(p)}, \quad \rho = l + \sum_{q=1}^{p-1} m^{(q)}, \quad p = 1(1)\overline{p}, \quad \nu = 1(1)m_{\overline{s}}, \quad (107)$$

in order to compute $w_m^{(p+\Delta p)}$ and $\vec{\bar{r}}_\nu$.
The scattering matrix S (see [10], Eqn. (1)) is defined by

$$\vec{\bar{b}}_\nu = S\vec{\bar{a}}_\nu, \quad \nu = 1(1)m_{\overline{s}}, \quad (108)$$

or

$$\bar{b}_{\rho,\nu} = \sum_{\sigma=1}^{m_{\overline{s}}} S_{\rho,\sigma} \cdot \bar{a}_{\sigma,\nu}, \quad \rho,\nu = 1(1)m_{\overline{s}}. \quad (109)$$

Because of (94), (100) we have for $\rho, \nu = 1(1)m_{\overline{s}}$

$$\begin{array}{rcl} \bar{w}_{\rho,\nu} & = & \bar{a}_{\rho,\nu} + \bar{b}_{\rho,\nu} , \\ 0 & = & \bar{r}_{\rho,\nu}\bar{a}_{\rho,\nu} - \bar{b}_{\rho,\nu} \end{array} \quad (110)$$

or

$$\begin{aligned} \bar{a}_{\rho,\nu}(1+\bar{r}_{\rho,\nu}) &= \bar{w}_{\rho,\nu}, \\ \bar{b}_{\rho,\nu}(1+\bar{r}_{\rho,\nu}) &= \bar{r}_{\rho,\nu}\bar{w}_{\rho,\nu}. \end{aligned} \tag{111}$$

We multiply Equation (109) with the product $\prod_{\mu=1}^{m_{\bar{s}}}(1+\bar{r}_{\mu,\nu})$ and get

$$\bar{b}_{\rho,\nu}\prod_{\mu=1}^{m_{\bar{s}}}(1+\bar{r}_{\mu,\nu}) = \sum_{\sigma=1}^{m_{\bar{s}}} S_{\rho,\sigma}\bar{a}_{\sigma,\nu}\prod_{\mu=1}^{m_{\bar{s}}}(1+\bar{r}_{\mu,\nu}), \quad \rho,\nu = 1(1)m_{\bar{s}}. \tag{112}$$

Substitution of (111) into the relation (112) gives

$$\bar{r}_{\rho,\nu}\bar{w}_{\rho,\nu}\prod_{\substack{\mu=1\\ \mu\neq\rho}}^{m_{\bar{s}}}(1+\bar{r}_{\mu,\nu}) = \sum_{\sigma=1}^{m_{\bar{s}}} S_{\rho,\sigma}\bar{w}_{\sigma,\nu}\prod_{\substack{\mu=1\\ \mu\neq\rho}}^{m_{\bar{s}}}(1+\bar{r}_{\mu,\nu}), \quad \rho,\nu = 1(1)m_{\bar{s}}, \tag{113}$$

or

$$R_{\rho,\nu} = \sum_{\sigma=1}^{m_{\bar{s}}} S_{\rho,\sigma}W_{\sigma,\nu} \tag{114}$$

with

$$W_{\rho,\nu} = \bar{w}_{\rho,\nu}\prod_{\substack{\mu=1\\ \mu\neq\rho}}^{m_{\bar{s}}}(1+\bar{r}_{\mu,\nu}), \quad R_{\rho,\nu} = \bar{r}_{\rho,\nu}W_{\rho,\nu}. \tag{115}$$

We can write (115) as matrix equation:

$$R = SW. \tag{116}$$

To compute the coefficients of the scattering matrix S we have to solve a system of complex linear algebraic equations for each row of S:

$$W^T(S_{\rho,1},\ldots,S_{\rho,m_{\bar{s}}})^T = (R_{\rho,1},\ldots,R_{\rho,m_{\bar{s}}})^T, \quad \rho = 1(1)m_{\bar{s}}. \tag{117}$$

That means, we have to solve only $m_{\bar{s}}$ linear algebraic equations rather than $(m_{\bar{s}})^2$ in order to compute the $(m_{\bar{s}})^2$ coefficients of S.

5 The Orthogonality Relation

In order to normalize the transverse electric mode fields we have to evaluate numerically the integral

$$\Im = \int\limits_{\Omega} (\vec{E}_t \times \vec{H}_{t,m}) \cdot d\vec{\Omega} = \int\limits_{\Omega} (E_x H_{y,m} - E_y H_{x,m}) d\Omega \tag{118}$$

for the mode m at a cross-sectional plane p. Ω is the area of the cross-sectional plane p. Because of ([10], Eqn. (7) and (11)) it follows from (118)

$$\Im = \int\limits_{\Omega} (E_x H_{y,m} - E_y H_{x,m}) d\Omega = \int\limits_{\Omega} (E_x \frac{1}{\tilde{\mu}\mu_0} B_{y,m} - E_y \frac{1}{\tilde{\mu}\mu_0} B_{x,m}) d\Omega \ . \tag{119}$$

Because the components of the electric field $\vec{E}_t$ are defined on the centers of the edges of the elementary cells and the components of the magnetic field $\vec{H}_{t,m}$ are normal to the face centers (see [10], section 4), both are located on different grid planes.
Let be $\overline{\Omega}$ the grid plane which corresponds to the cross-sectional plane p.
Let be $E_{x_{i,j,k}}$ and $E_{y_{i,j,k}}$, $i_1 \leq i \leq i_2, \quad j_1 \leq j \leq j_2, \quad k = const.$ the transverse electric field components on the grid plane $\overline{\Omega}$.
Let be $\tilde{B}_{x_{i,j,k},m}$ and $\tilde{B}_{y_{i,j,k},m}$, $i_1 \leq i \leq i_2, \quad j_1 \leq j \leq j_2, \quad k = const.$ the transverse magnetic flux density on the same grid plane $\overline{\Omega}$ (see Figure 4).
We use the lowest order integration formula ([10], Eqn. (18)) to approximate the integral $\Im$ (see (119)). Because the material constants can be different between two different elementary cells of the primary grid, we have to divide the integration domain (see also [10], section 4, and Figure 4 in this paper). Thus, the grid plane $\overline{\Omega}$ consists of $(i_2 - i_1 + 1)(j_2 - j_1 + 1)$ partial planes, and we get the following approximation of the integral:

$$\begin{aligned} \Im \; &= \sum_{\substack{i_1 \leq i \leq i_2 \\ j_1 \leq j \leq j_2}} E_{x_{i,j,k}} \big(x_{i,j,k} \tfrac{y_{i,j-1,k}}{2} \tfrac{1}{\mu_0 \tilde{\mu}_{i,j-1,k}} \tilde{B}_{y_{i,j,k},m} + x_{i,j,k} \tfrac{y_{i,j,k}}{2} \tfrac{1}{\mu_0 \tilde{\mu}_{i,j,k}} \tilde{B}_{y_{i,j,k},m} \big) \\ &- \sum_{\substack{i_1 \leq i \leq i_2 \\ j_1 \leq j \leq j_2}} E_{y_{i,j,k}} \big(\tfrac{x_{i-1,j,k}}{2} y_{i,j,k} \tfrac{1}{\mu_0 \tilde{\mu}_{i-1,j,k}} \tilde{B}_{x_{i,j,k},m} + \tfrac{x_{i,j,k}}{2} y_{i,j,k} \tfrac{1}{\mu_0 \tilde{\mu}_{i,j,k}} \tilde{B}_{x_{i,j,k},m} \big) \end{aligned} \tag{120}$$

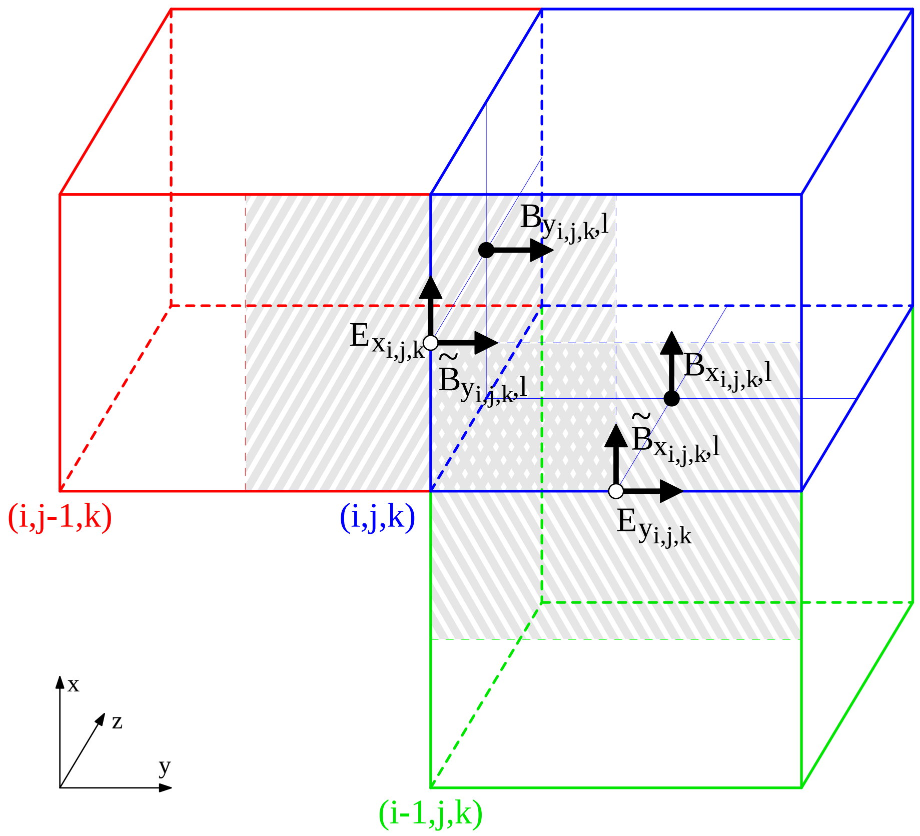

Figure 4: Integration domain for the orthogonality relation

or

$$\Im = \sum_{\substack{i_1 \leq i \leq i_2 \\ j_1 \leq j \leq j_2}} E_{x_{i,j,k}} G_{x_{i,j,k},m} + \sum_{\substack{i_1 \leq i \leq i_2 \\ j_1 \leq j \leq j_2}} E_{y_{i,j,k}} G_{y_{i,j,k},m} \tag{121}$$

with

$$
\begin{aligned}
G_{x_{i,j,k},m} &= \tfrac{x_{i,j,k}}{2\mu_0}\left(\tfrac{y_{i,j-1,k}}{\tilde{\mu}_{i,j-1,k}} + \tfrac{y_{i,j,k}}{\tilde{\mu}_{i,j,k}}\right)\tilde{B}_{y_{i,j,k},m} \quad , \\
G_{y_{i,j,k},m} &= -\tfrac{y_{i,j,k}}{2\mu_0}\left(\tfrac{x_{i-1,j,k}}{\tilde{\mu}_{i-1,j,k}} + \tfrac{x_{i,j,k}}{\tilde{\mu}_{i,j,k}}\right)\tilde{B}_{x_{i,j,k},m}
\end{aligned}
\tag{122}
$$

or

$$
\Im = \vec{E}_t \cdot \vec{G}_m \tag{123}
$$

with

$$
\vec{G}_m = (\vec{G}_{x,m}, \vec{G}_{y,m})^T \quad . \tag{124}
$$

If the boundary is considered to be electrically perfectly conducting, we have to take into account that $E_{x_{i,j,k}} = 0$ or $E_{y_{i,j,k}} = 0$ on the boundary.
If the boundary is considered to be magnetically perfectly conducting, we have to take into account that $x_{i,j,k} = 0$ or $x_{i-1,j,k} = 0$, and $y_{i,j,k} = 0$ or $y_{i,j-1,k} = 0$, respectively, that means, the areas which are located outside of the domain are not to be taken into consideration.
The orthogonality relation (88) is applied at two neighboring planes cutting the longitudinal transmission line axis at z_p and $z_p + \Delta z_p$. With (89) the transverse mode fields $\vec{E}_{t,m}^{(p)}$ are normalized, and we get from (121)

$$
\vec{E}_{t,m}^{(p)} \cdot \vec{G}_m^{(p)} = \eta_m \tag{125}
$$

with

$$
\vec{E}_{t,m}^{(p)} \cdot \vec{G}_m^{(p)} = \sum_{\substack{i_1 \le i \le i_2 \\ j_1 \le j \le j_2}} E_{x_{i,j,k},m}^{(p)} G_{x_{i,j,k},m}^{(p)} + \sum_{\substack{i_1 \le i \le i_2 \\ j_1 \le j \le j_2}} E_{y_{i,j,k,m}}^{(p)} G_{y_{i,j,k},m}^{(p)} \quad . \tag{126}
$$

We get from (97, 123, 125)

$$
w_m^{(p+\Delta p)} = \frac{1}{\eta_m} \int\limits_{\Omega} (\vec{E}_t^{(p+\Delta p)} \times \vec{H}_{t,m}^{(p)}) \cdot d\vec{\Omega} = \frac{1}{\eta_m} \vec{E}_t^{(p+\Delta p)} \cdot \vec{G}_m^{(p)} \quad . \tag{127}
$$

The components $G_{x_{i,j,k},m}$ and $G_{y_{i,j,k},m}$ contain the unknown quantities $\tilde{B}_{x_{i,j,k},m}$ and $\tilde{B}_{y_{i,j,k},m}$. We will see, they can be computed using the known

electric field.
Because of ([10], Eqn. (21)) we have for the mode m

$$\begin{aligned}
B_{z_{i,j,k},m} &= \frac{\jmath}{\omega}\Big(\frac{1}{y_{i,j,k}}E_{x_{i,j,k},m} - \frac{1}{y_{i,j,k}}E_{x_{i,j+1,k},m} + \\
&\quad \frac{1}{x_{i,j,k}}E_{y_{i+1,j,k},m} - \frac{1}{x_{i,j,k}}E_{y_{i,j,k},m}\Big) \ , \\
B_{z_{i,j-1,k},m} &= \frac{\jmath}{\omega}\Big(\frac{1}{y_{i,j-1,k}}E_{x_{i,j-1,k},m} + \frac{1}{x_{i,j-1,k}}E_{y_{i+1,j-1,k},m} - \\
&\quad \frac{1}{y_{i,j-1,k}}E_{x_{i,j,k},m} - \frac{1}{x_{i,j-1,k}}E_{y_{i,j-1,k},m}\Big) \ .
\end{aligned} \tag{128}$$

Because of the longitudinal homogeneity of the transmission lines we assume similar to ([10], section 8, Eqn. (87), see also Figure 4 in this paper)

$$\begin{aligned}
B_{y_{i,j,k},m} &= \tilde{B}_{y_{i,j,k},m} e^{-\jmath k_{z_m} h} \ , \\
B_{y_{i,j,k-1},m} &= \tilde{B}_{y_{i,j,k},m} e^{+\jmath k_{z_m} h}
\end{aligned} \tag{129}$$

with

$$h = \frac{z_{i,j,k}}{2} = \frac{z_{i,j,k-1}}{2} \ . \tag{130}$$

Using ([10], Eqns. (86) and (87)), and substituting $B_{z_{i,j,k},m}$, $B_{z_{i,j-1,k},m}$, $B_{y_{i,j,k},m}$ and $B_{y_{i,j,k-1},m}$ from (128, 129) into the first Maxwellian equation ([10], Eqn. (21)) gives

$$\begin{aligned}
&\frac{\jmath}{\omega}\frac{4h}{\tilde{\mu}_{i,j,k}}\Big(\frac{1}{y_{i,j,k}}E_{x_{i,j,k},m} - \frac{1}{y_{i,j,k}}E_{x_{i,j+1,k},m} + \frac{1}{x_{i,j,k}}E_{y_{i+1,j,k},m} - \frac{1}{x_{i,j,k}}E_{y_{i,j,k},m}\Big) - \\
&\Big(\frac{y_{i,j,k}}{\tilde{\mu}_{i,j,k}} + \frac{y_{i,j-1,k}}{\tilde{\mu}_{i,j-1,k}}\Big)\tilde{B}_{y_{i,j,k},m} e^{-\jmath k_{z_m} h} - \frac{\jmath}{\omega}\frac{4h}{\tilde{\mu}_{i,j-1,k}}\Big(\frac{1}{y_{i,j-1,k}}E_{x_{i,j-1,k},m} + \\
&\frac{1}{x_{i,j-1,k}}E_{y_{i+1,j-1,k},m} - \frac{1}{y_{i,j-1,k}}E_{x_{i,j,k},m} - \frac{1}{x_{i,j-1,k}}E_{y_{i,j-1,k},m}\Big) + \\
&\qquad \Big(\frac{y_{i,j-1,k-1}}{\tilde{\mu}_{i,j-1,k-1}} + \frac{y_{i,j,k-1}}{\tilde{\mu}_{i,j,k-1}}\Big)\tilde{B}_{y_{i,j,k},m} e^{+\jmath k_{z_m} h} = 2\jmath\omega\epsilon_0\mu_0 g^{y,z}_{i,j,k} E_{x_{i,j,k},m} \ .
\end{aligned} \tag{131}$$

Using ([10], Eqns. (35), (86), (88) and (100)) after transformation from (131) yields

$\frac{\jmath}{\omega}\left\{\frac{1}{\tilde{\mu}_{i,j,k}}\left[\frac{1}{x_{i,j,k}}(E_{y_{i+1,j,k},m}-E_{y_{i,j,k},m})+\frac{1}{y_{i,j,k}}(E_{x_{i,j,k},m}-E_{x_{i,j+1,k},m})\right]+\right.$

$\left.\frac{1}{\tilde{\mu}_{i,j-1,k}}\left[\frac{1}{x_{i,j-1,k}}(E_{y_{i,j-1,k},m}-E_{y_{i+1,j-1,k},m})+\frac{1}{y_{i,j-1,k}}(E_{x_{i,j,k},m}-E_{x_{i,j-1,k},m})\right]\right\}-$

$\frac{1}{2}\jmath\omega\epsilon_0\mu_0(y_{i,j,k}\tilde{\epsilon}_{i,j,k}+y_{i,j-1,k}\tilde{\epsilon}_{i,j-1,k})E_{x_{i,j,k},m}$

$$=-\frac{1}{2}\left(\frac{y_{i,j,k}}{\tilde{\mu}_{i,j,k}}+\frac{y_{i,j-1,k}}{\tilde{\mu}_{i,j-1,k}}\right)\left[\frac{1}{2h}(e^{+\jmath k_{z_m}h}-e^{-\jmath k_{z_m}h})\right]\tilde{B}_{y_{i,j,k},m}\ . \qquad (132)$$

Because of

$$\frac{1}{2h}(e^{+\jmath k_{z_m}h}-e^{-\jmath k_{z_m}h})=\jmath k_{z_m}\frac{\sin(k_{z_m}h)}{k_{z_m}h}, \qquad (133)$$

and

$$\lim_{h\to 0}\frac{\sin(k_{z_m}h)}{k_{z_m}h}=1, \qquad (134)$$

and using the definition of $\varkappa_0$ (see (6)), we have after multiplication of equation (132) with $(-\frac{x_{i,j,k}}{\jmath k_{z_m}\mu_0})$ (see (122) for the definition of $G_{x_{i,j,k},m}$):

$G_{x_{i,j,k},m}=\frac{x_{i,j,k}}{2\mu_0}\left(\frac{y_{i,j,k}}{\tilde{\mu}_{i,j,k}}+\frac{y_{i,j-1,k}}{\tilde{\mu}_{i,j-1,k}}\right)\tilde{B}_{y_{i,j,k},m}=$

$-\frac{x_{i,j,k}}{\omega k_{z_m}\mu_0}\left\{\frac{1}{\tilde{\mu}_{i,j,k}}\left[\frac{1}{x_{i,j,k}}(E_{y_{i+1,j,k},m}-E_{y_{i,j,k},m})+\frac{1}{y_{i,j,k}}(E_{x_{i,j,k},m}-E_{x_{i,j+1,k},m})\right]\right.$

$+\frac{1}{\tilde{\mu}_{i,j-1,k}}\left[\frac{1}{x_{i,j-1,k}}(E_{y_{i,j-1,k},m}-E_{y_{i+1,j-1,k},m})+\frac{1}{y_{i,j-1,k}}(E_{x_{i,j,k},m}-E_{x_{i,j-1,k},m})\right]$

$$\left.-\frac{1}{2}\varkappa_0^2(y_{i,j,k}\tilde{\epsilon}_{i,j,k}+y_{i,j-1,k}\tilde{\epsilon}_{i,j-1,k})E_{x_{i,j,k},m}\right\}\ . \qquad (135)$$

The corresponding formula for $G_{y_{i,j,k},m}$ can be derived in a similar way.

$G_{y_{i,j,k},m}=-\frac{y_{i,j,k}}{2\mu_0}\left(\frac{x_{i,j,k}}{\tilde{\mu}_{i,j,k}}+\frac{x_{i-1,j,k}}{\tilde{\mu}_{i-1,j,k}}\right)\tilde{B}_{x_{i,j,k},m}=$

$-\frac{y_{i,j,k}}{\omega k_{z_m}\mu_0}\left\{\frac{1}{\tilde{\mu}_{i,j,k}}\left[\frac{1}{y_{i,j,k}}(E_{x_{i,j+1,k},m}-E_{x_{i,j,k},m})+\frac{1}{x_{i,j,k}}(E_{y_{i,j,k},m}-E_{y_{i+1,j,k},m})\right]\right.$

$+\frac{1}{\tilde{\mu}_{i-1,j,k}}\left[\frac{1}{y_{i-1,j,k}}(E_{x_{i-1,j,k},m}-E_{x_{i-1,j+1,k},m})+\frac{1}{x_{i-1,j,k}}(E_{y_{i,j,k},m}-E_{y_{i-1,j,k},m})\right]$

$$\left.-\frac{1}{2}\varkappa_0^2(x_{i,j,k}\tilde{\epsilon}_{i,j,k}+x_{i-1,j,k}\tilde{\epsilon}_{i-1,j,k})E_{y_{i,j,k},m}\right\}\ . \qquad (136)$$

By using (135) and (136) we can compute the integral $\Im$ (see (121)).

6 Numerical Results

Comparisons between the original version of the program package F3D and the new version are presented.
The reduction of the computing times are demonstrated for the via hole represented in Figure 5.

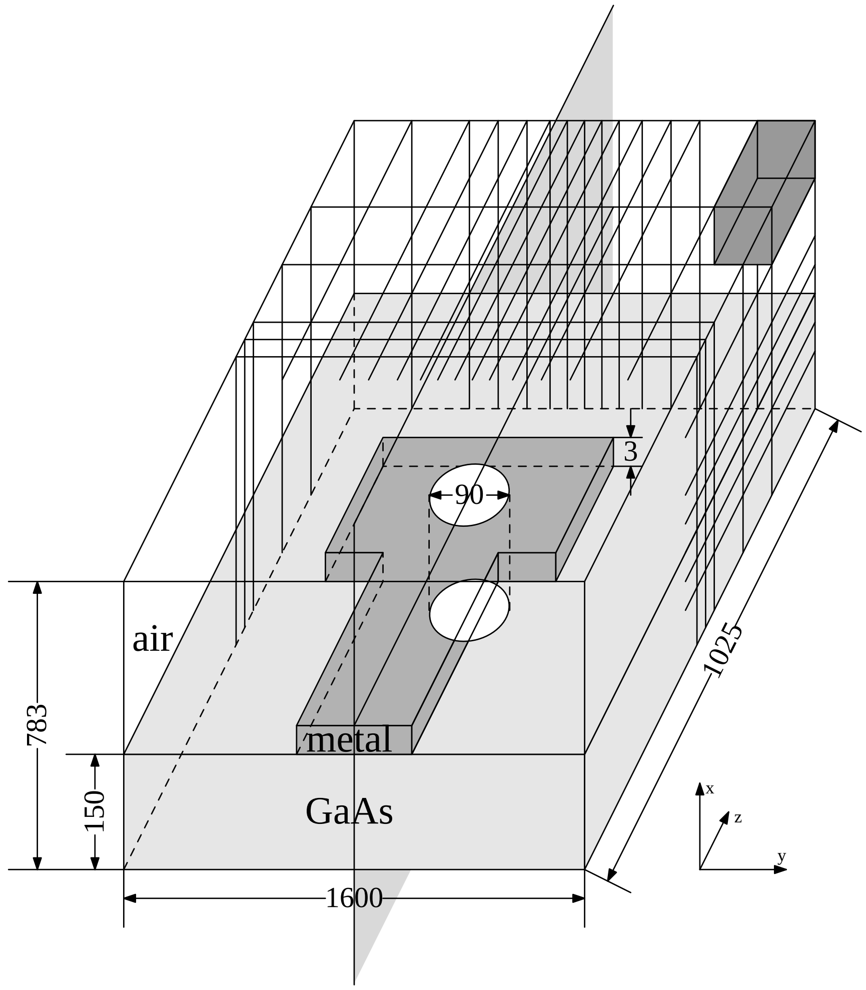

Figure 5: Via hole with a nonequidistant grid of rectangular parallelepipeds (dimensions in μm)

The structure represented in Figure 5 is symmetric along the z-direction. Using appropriate boundary conditions it will do to discretize the right-hand

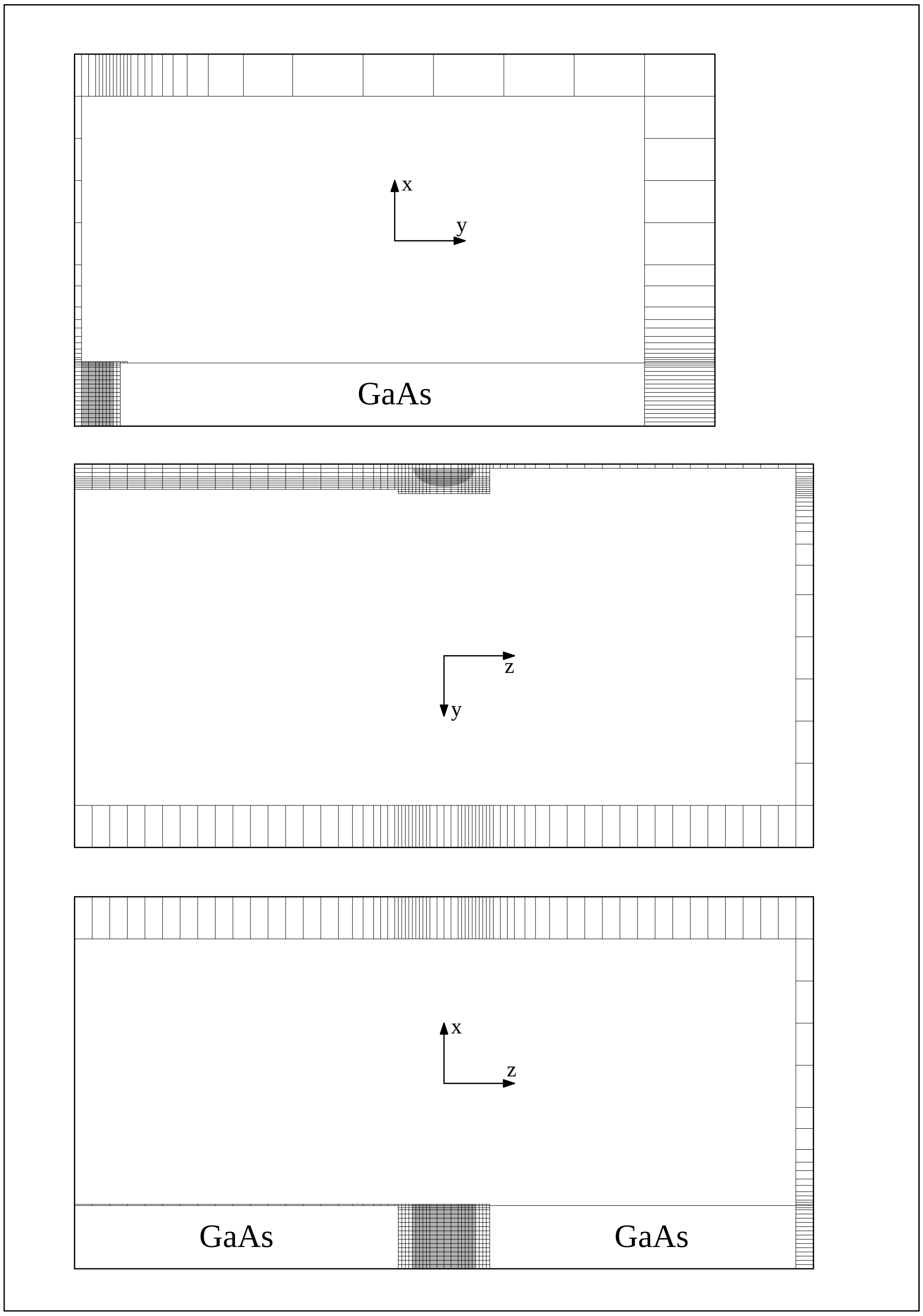

Figure 6: Discretization of the right-hand side of the structure via hole (dimensions in μm)

side of the structure only. This right-hand side of the structure is divided into $n_{xyz} = n_x n_y n_z$ ($n_x = 33, n_y = 28, n_z = 66$) elementary cells (see Figure 6).

The eigenvalue problem

The examples were computed on a SUN SPARC Server 630 with 128 MBytes memory.

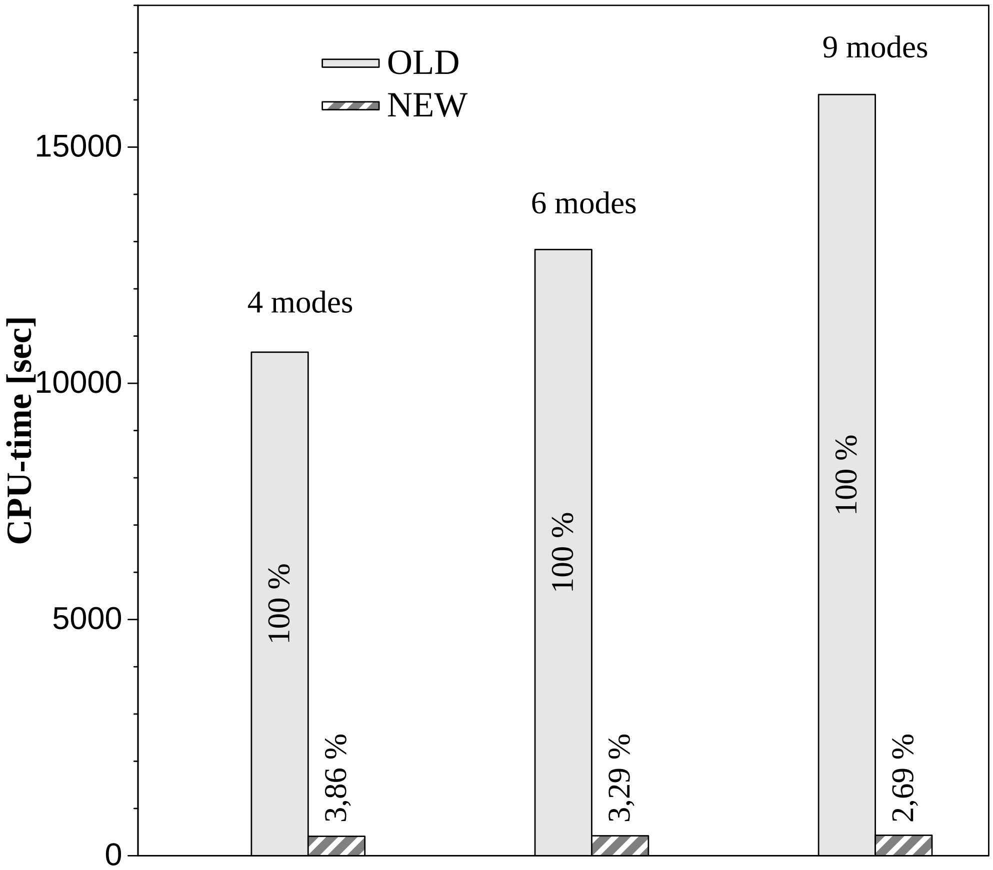

Figure 7: Comparisons of computing times for the eigenvalue problem between the original and the new version of F3D

The dimension of the eigenvalue problem amounts to $2n_x n_y - n_b = 1668$. n_b is caused by the boundary conditions (see section 2).

The algorithm of the original version of the program package F3D requires a storage of $2(2n_x n_y - n_b)^2$ elements for the eigenvalue matrix since the matrix

has to be stored twice.
The total storage requirements is reduced by a factor of 20 in the new version for this example, since the sparse storage technique is applied.
The maximum number of non-zeros in a row of the matrix is 9. The total number of non-zero elements of the matrix amounts to 11598 for the example via hole. The number of non-zeros of the matrix after the LU-decomposition (see section 2) is 89947 for this example.
The reduction of the computing time depends essentially on the number of required eigenvalues and eigenvectors. The reduction of the computing time is represented in Figure 7 for the calculation of 4, 6 and 9 eigenvectors. The time measurements involve matrix generation, solving the eigenvalue problem and computation of the mode fields.

The system of linear algebraic equations

The examples were computed on a SUN SPARCstation 2 with 128 MBytes memory.
The order of the system of linear algebraic equations is $3n_{xyz} = 182952$ (see section 3). The maximum number of non-zeros in a row of the matrix is 15. The total number of non-zero elements of the matrix amounts to 1328690 for the example via hole.
The reduction of the computing time is represented in Figure 8.

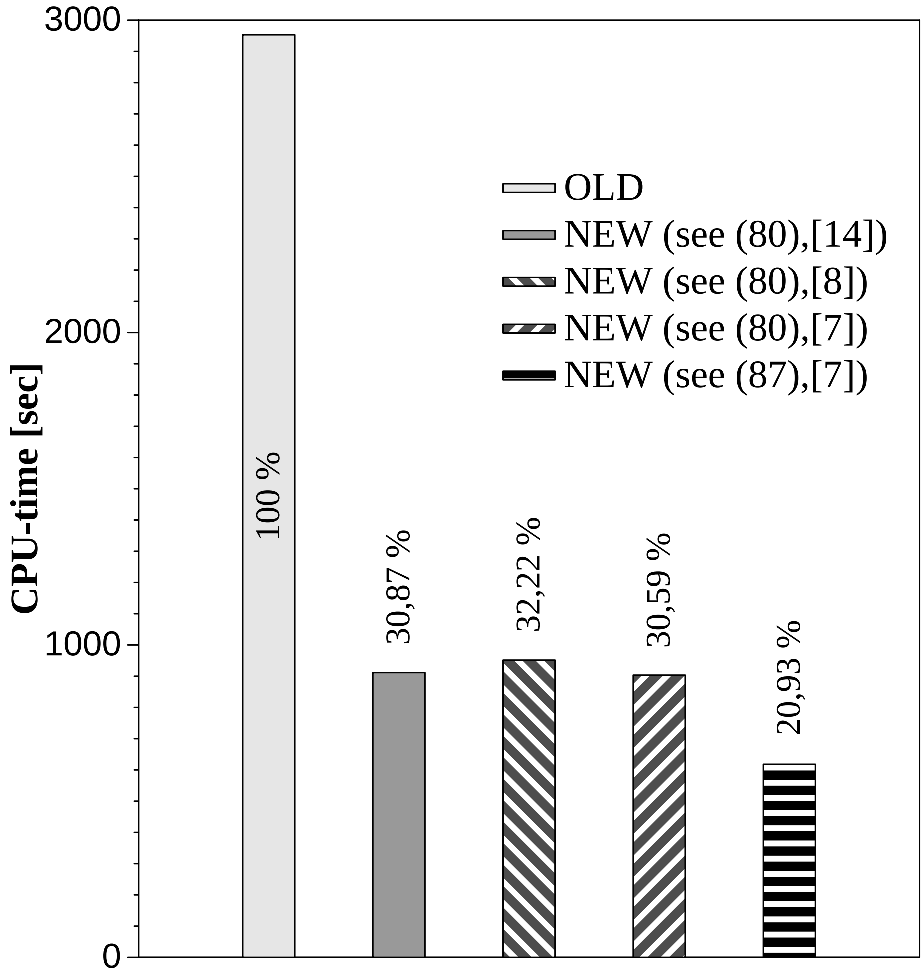

Figure 8: Comparisons of computing times for the system of linear algebraic equations

References

[1] Barret, R., Berry, M., Chan, T., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., van der Vorst, H., *Templates for the solution of linear systems: Building blocks for iterative methods*, SIAM, Philadelphia, Pennsylvania, 1993.

[2] Beilenhoff, K., Heinrich, W., Hartnagel, H. L., *Improved Finite-Difference Formulation in Frequency Domain for Three-Dimensional*

Scattering Problems, IEEE Transactions on Microwave Theory and Techniques, Vol. 40, No. 3, March 1992.

[3] Burmeister, W., *Programmpaket für lineare Gleichungssysteme mit schwach besetzten Matrizen, Moduln für lineare Gleichungssysteme*, in Numerische Verfahren in Aktion 1, Akademie-Verlag, Berlin, S. 125-154, 1980.

[4] Christ, A., Hartnagel, H. L., *Three-Dimensional Finite-Difference Method for the Analysis of Microwave-Device Embedding*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-35, No. 8, pp. 688-696, June 1987.

[5] Christ, A., *Streumatrixberechnung mit dreidimensionalen Finite-Differenzen für Mikrowellen-Chip-Verbindungen und deren CAD-Modelle*, Fortschrittberichte VDI, Reihe 21: Elektrotechnik, Nr. 31, S. 1-154, 1988.

[6] Collin, R. E., *Field Theory of Guided Waves*, McGraw-Hill Book Company New York, Toronto, London, 1960.

[7] Deuflhard, P., *A Study of Lanczos-Type Iterations for Symmetric Indefinite Linear Systems*, preprint SC 93-6, Konrad-Zuse-Zentrum für Informationstechnik Berlin, March 1993.

[8] Freund, R.W., Nachtigal, N.M., *A new Krylov-subspace method for symmetric indefinite linear systems*, AT&T Numerical Analysis Manuscript, Bell Labs, Murry Hill, NJ, 1994.

[9] Hebermehl, G., Binder, B., Brysch, R., Burmeister, W., Creutziger, J., Ehlert, J., Fiedler, O., Funke, R., Grohmann, U. R., Hübner, F. K., Jahnke, R., Kehl, U., Keusch, C., Kleemann, B., Knauf, L., Luber, H., Marx, M., Pohl, W., Reinhardt, G., Sandmann, H., Schalm, G., Schlundt, R., *NUMATH - Software for Numerical Mathematics*, Karl-Weierstraß-Institut für Mathematik, Report R-MATH-05/89, Berlin 1989.

[10] Hebermehl, G., Schlundt, R., Zscheile, H., Heinrich, W., *Simulation of Monolithic Microwave Integrated Circuits*, Preprint No. 235, Weierstraß-Institut für Angewandte Analysis und Stochastik im Forschungsverbund Berlin e.V., 1996.

[11] R.B. Lehoucq, R. B., *Analysis and Implementation of an Implicitly Restarted Arnoldi Iteration*, Rice University, Technical Report TR95-13, Department of Computational and Applied Mathematics, 1995.

[12] Markowitz, H. M., *An Algorithm for Least-squares Estimation on Nonlinear Parameters*, SIAM J. Appl. Math., 11, pp. 431-441, 1957.

[13] Nour-Omid, B., Parlett, B. N., Ericson, T., Jensen, P. S., *How to Implement the Spectral Transformation*, Math. Comp., 48, pp. 663-673, 1987.

[14] Paige, C.C., Saunders, M.A., *Solution of Sparse Indefinite Systems of Linear Equations*, SIAM J. Numer. Anal. 12, 4, pp. 617-629, September 1975.

[15] Parlett, B. N., Saad, Y., *Complex Shift and Invert Strategies for Real Matrices*, Linear Algebra and its Applications, vol 88/89, pp. 575-595, 1987.

[16] Rheinboldt, W., Mesztenyi, C. K., *Programs for the Solution of Large Sparse Matrix Problems Based on the Arcgraph Structure*, Univ. of Maryland, Computer Sc. Center, TR-262, 1973.

[17] Sorensen, D. C., *Implicit Application of Polynomial Filters in a k-Step Arnoldi Method*, SIAM J. Matr. Anal. Apps., 13, pp. 357-385, 1992.

[18] Sorensen, D. C., Lehoucq, R. B., Vu, Ph., *ARPACK*, Version-Number 2.2, 1995.

[19] Weiland, T., *Three Dimensional Resonator Mode Computation by Finite Difference Method*, IEEE Transactions on Magnetics, Vol. MAG-21, No. 6, pp. 2340-2343, November 1985.

[20] Weiland, T., *On the Unique Numerical Solution of Maxwell Eigenvalue Problems in Three Dimensions*, Particle Accelerators, Vol. 17, pp. 227-242, 1985.