Analytical Gradient Evaluation of Cost Functions in General Field Solvers: A Novel Approach for Optimization of Microwave Structures

P. Harscher, S. Amari* and R. Vahldieck and J. Bornemann*

Swiss Federal Institute of Technology, IFH
ETH Zentrum, Gloriastrasse 35, CH - 8092 Zürich
*Department of Electrical and Computer Engineering,
University of Victoria
BOX 3055, Victoria, B.C. V8W 3P6

Abstract

This paper introduces a method for the analytical calculation of gradients of a cost functions which is an attractive feature when optimizing microwave structures using field solvers. In contrast to utilizing finite differencing all gradients are computed from a single analysis of the structure regardless of its complexity. It is not even necessary to invert a large matrix; a linear system [A][x]=[b] is solved instead. No remeshing is required in the FEM and the gradient values are exact. The basic technique used in this new approach is applied to a moment method technique (CIET) and the finite element method (FEM). Both methods lend itself to the appropriate matrix equation.

Introduction

Evaluating the gradient of a cost function for the optimization of microwave circuits is usually based on the finite difference technique and can be a time consuming task. This is especially true when the circuit transfer function is calculated on the basis of a field-theory simulation tool since always two computations are necessary for one gradient. If, in addition, the number of independent variables is large, optimization can become an impossible task. In this contribution it will be shown that, under certain circumstances, the gradient of a cost function can be calculated analytically without using finite differences. The number of computations can be cut in half and well known disadvantages with finite differencing like inaccuracies at singularities in highly resonant structures are eliminated. The method has been applied successfully to the coupled integral equation technique (CIET) and the finite element method (FEM). Analytical calculation of the cost function is also possible with the adjoint network method (ANM) but requires a network representation of the structure to be optimized (and its adjoint). The mode matching technique (MMT) to calculate the fields is normally utilised to extract the network's representation in form of the admittance or scattering matrix. The ANM has been successfully applied to the optimization of filters and radiating structures.

Analytically calculating the gradient of a cost function directly in general numerical techniques without first deriving a network representation has not been published before. The possibility of doing so is of great interest as it offers a number of obvious advantages and also not so obvious ones, depending on the numerical method used. The new approach can not be applied to all numerical field computation methods since it requires a scattering problem representation of the whole micro-wave structure of the form which results directly from applying the FEM or a moment method (i.e. CIET), but is not necessarily limited to these methods. Here, [A] is a M x M matrix which depends on the independent variables and represents the structure to be optimized, [b] is the excitation and [x] is the response. For example, the vector [x] contains the expansion coefficients in the MoM or the nodal values in FEM.

It will be shown, that as long as the partial derivatives of the matrix [A] and the excitation [b] are known analytically, all sensitivities can be determined analytically. Up to now this approach has been successfully tested with the MoM in the optimization of waveguide filters and was subsequently applied to the FEM. It can be extended to other methods for which the scattering problem can be formulated as above.

The advantages of this approach are summarized as follows:

No network representation needed; only one cost function evaluation instead of two; higher accuracy compared to a finite difference scheme in particular in the vicinity of resonances; no remeshing of the structure required during gradient calculations; no matrix inversion necessary; reduced memory requirements; faster algorithms.

OUTLINE

- Introduction
- Review
- Analysis of Microwave Structures
- Optimization
- Analytic Gradient Evaluation
- Example CIET
- Example FEM
- Conclusion

INTRODUCTION

Design of Microwave Devices and Structures consists of:

- Modelling
- Analysis
- Optimization

Optimization can be done by:

- Stochastic Methods
- Deterministic Methods (i.e. Gradient Methods)

Goal: Minimize a Cost Function of the form:

$$F(a_i) = \sum_n K_n \left[\left| S_{11}(\omega_n, a_i) \right| - \left| S_{11}^{opt}(\omega_n, a_i) \right| \right]^2$$

K_n : constants

a_i : parameters

ω_n : frequencies

REVIEW

The problem is the calculation of the gradients $$\frac{\partial}{\partial a_i} F(a_i)$$

What has been done in the past?

- in general: calculation of the gradients with finite difference
- optimization of microwave filters and radiating structures by the Adjoint Network Method (ANM)
  - network representation of the microwave circuit (MMT)
  - analytical evaluation of network sensitivities (gradients)

NEW: Analytical calculation of gradients (due to geometric changes) of cost functions when using more general techniques like Finite Element Method (FEM) or Method of Moments (MoM).

ANALYSIS OF MICROWAVE STRUCTURES

numerical methods like:

- Finite Element Method (FEM)
- Coupled Integral Equation Technique (CIET)
- Spectral Domain Approach (SDA)
- Frequency Domain Transmission Line Matrix Method (FDTLM)
- Method of Moments (MoM)
- Method of Lines (MoL)

→ Matrix equation

$$[Z]\{k\} = \{S\}$$

OPTIMIZATION

We use: deterministic, gradient based method

minimize a cost function of the form:

$$F(a_i) = \sum_n K_n \left[\left| S_{11}(\omega_n, a_i) \right| - \left| S_{11}^{opt}(\omega_n, a_i) \right| \right]^2$$

with first derivatives:

$$\frac{\partial}{\partial a_i} F(a_i) = 2 \sum_n K_n \left(\left| S_{11}(\omega_n, a_i) \right| - \left| S_{11}^{opt}(\omega_n a_i) \right| \right) \frac{\partial \left| S_{11}(\omega_n, a_i) \right|}{\partial a_i}$$

a_i: variables, i.e. geometry parameters

Characterization of the problem:

- multidimensional
- nonlinear
- function of several variables
- constraint variables

Suitable methods:

- Gauss Newton
- Levenberg–Marquardt

→ use of first derivatives

(Numerical recipes, NAG library, Matlab optimization toolbox)

ANALYTIC GRADIENT EVALUATION

$$\frac{\partial}{\partial a_i} F(a_i) = 2\sum_n K_n \left(\left| S_{11}(\omega_n, a_i) \right| - \left| S_{11}^{opt}(\omega_n a_i) \right| \right) \frac{\partial \left| S_{11}(\omega_n, a_i) \right|}{\partial a_i}$$

Matrix equation: $[Z]\{k\} = \{S\}$

$$S_{11}(\omega_n, a_i) = const.* \, fct(a_i) * \{k\}_j$$

$$\frac{\partial |S_{11}|}{\partial a_i} = \mathrm{Re}\left\{ \frac{|S_{11}|}{S_{11}} \frac{\partial S_{11}}{\partial a_i} \right\}$$

$$\frac{\partial}{\partial a_i} S_{11}(\omega_n, a_i) = \ldots \frac{\partial}{\partial a_i}\{k\}_j$$

$$[Z]\{k\} = \{S\}$$

$$\frac{\partial}{\partial a_i}[Z]\{k\} + [Z]\frac{\partial}{\partial a_i}\{k\} = \frac{\partial}{\partial a_i}\{S\}$$

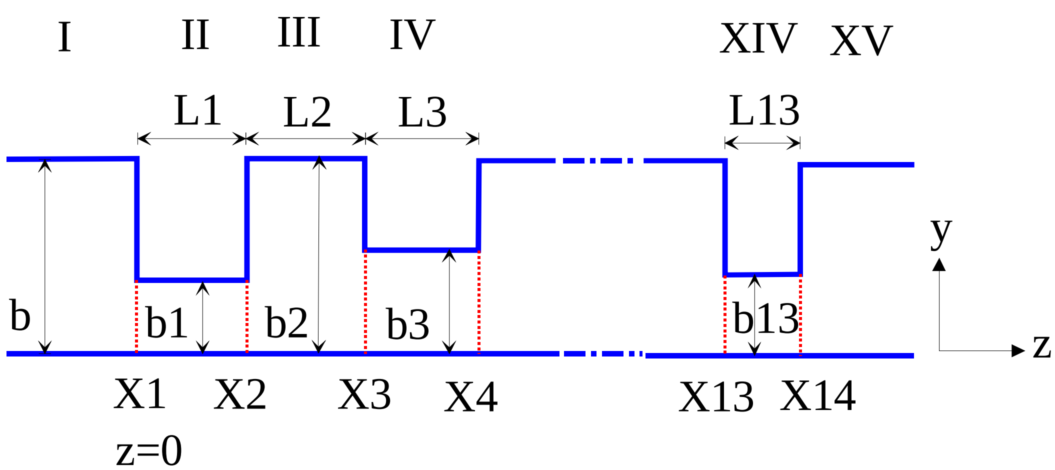
first possibility:

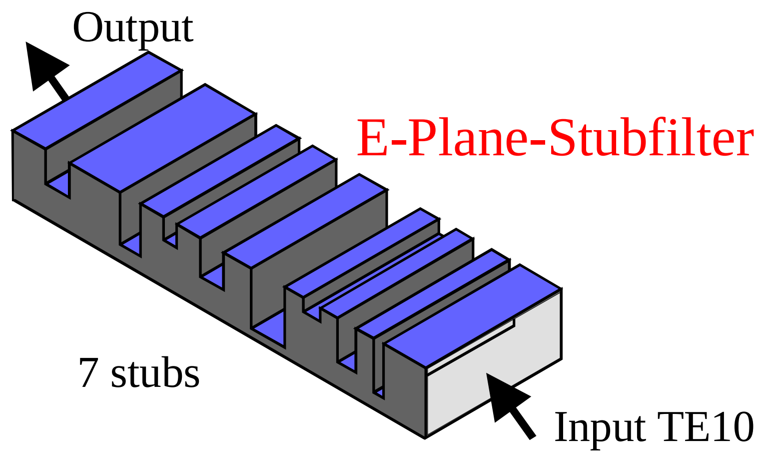
$$\frac{\partial}{\partial a_i}\{k\} = [Z]^{-1}\left(\frac{\partial}{\partial a_i}\{S\} - \frac{\partial}{\partial a_i}[Z]\{k\}\right)$$, with matrix inversion $[Z]^{-1}$

second possibility:

SOLVE linear system for $\frac{\partial}{\partial a_i}\{k\}$ without inverting matrix $[Z]$:

FIRST EXAMPLE: CIET





Equations for $\vec{E}$ and $\vec{H}$ in all regions

Longitudinal Section Electric (LSE)-Modes

example:

$$E_y^I = e^{-jk_1^I z} + \sum_{n=1}^{\infty} B_n^I \cos\left[(n-1)\frac{\pi}{b}y\right] e^{jk_n^I z}$$

$$H_x^I = -Y_1^I e^{-jk_1^I z} + \sum_{n=1}^{\infty} Y_n^I B_n^I \cos\left[(n-1)\frac{\pi}{b}y\right] e^{jk_n^I z}$$

$$k_m^i = \sqrt{k_0^2 - \left(\frac{\pi}{a}\right)^2 - \left(\frac{(m-1)\pi}{b_b}\right)^2}$$

$$Y_m^i = \frac{k_0^2 - \left(\frac{\pi}{a}\right)^2}{\omega\mu_0 k_m^i}$$

The result is a matrix equation of the form $[Z]\{k\}=\{S\}$

$$\begin{bmatrix} A & B & & & 0 & 0 \\ C & D & E & & 0 & 0 \\ & F & G & H & & \\ & & & \cdots & & \\ 0 & 0 & & AJ & AK & AL \\ 0 & 0 & & & AM & AN \end{bmatrix} \cdot \begin{Bmatrix} a \\ b \\ c \\ \cdots \\ m \\ n \end{Bmatrix} = \begin{Bmatrix} U \\ 0 \\ 0 \\ \\ \\ 0 \end{Bmatrix}$$

unknown coefficients:

$a_p, b_p \ldots n_p$, p=1 …M

M: number of basis functions

U_p : excitation

and the matrix elements $A_{ip}, B_{ip}, \ldots AN_{ip}$, example

$$A_{ip} = j\sum_{n=1}^{\infty} Y_n^I B_i^{b1b}(n) B_p^{b1b}(n) + \sum_{n=1}^{\infty} B_i^{b1b1}(n) B_p^{b1b1}(n) \cot\left(k_n^{II} L_1\right)\frac{b_1}{b}$$

$$U_p = jY_1^I B_i^{b1b1}(1)\frac{b_1}{b}$$

Basis functions:

$$B_i^{bjbl}(n) = \frac{b_j}{b_l(1+\partial_{n1})} \cdot \frac{\sqrt{\pi}}{2}\Gamma\left(\frac{2}{3}\right)\left[\frac{J_{1/6}\left(\pi\left|(i-1)-(n-1)\frac{b_j}{b_l}\right|\right)}{\left(\frac{\pi}{2}\left|(i-1)-(n-1)\frac{b_j}{b_l}\right|\right)^{1/6}} + \frac{J_{1/6}\left(\pi\left((i-1)+(n-1)\frac{b_j}{b_l}\right)\right)}{\left(\frac{\pi}{2}\left((i-1)+(n-1)\frac{b_j}{b_l}\right)\right)^{1/6}}\right]$$

Solve for the unknown coefficients $a_i, b_i, \ldots$

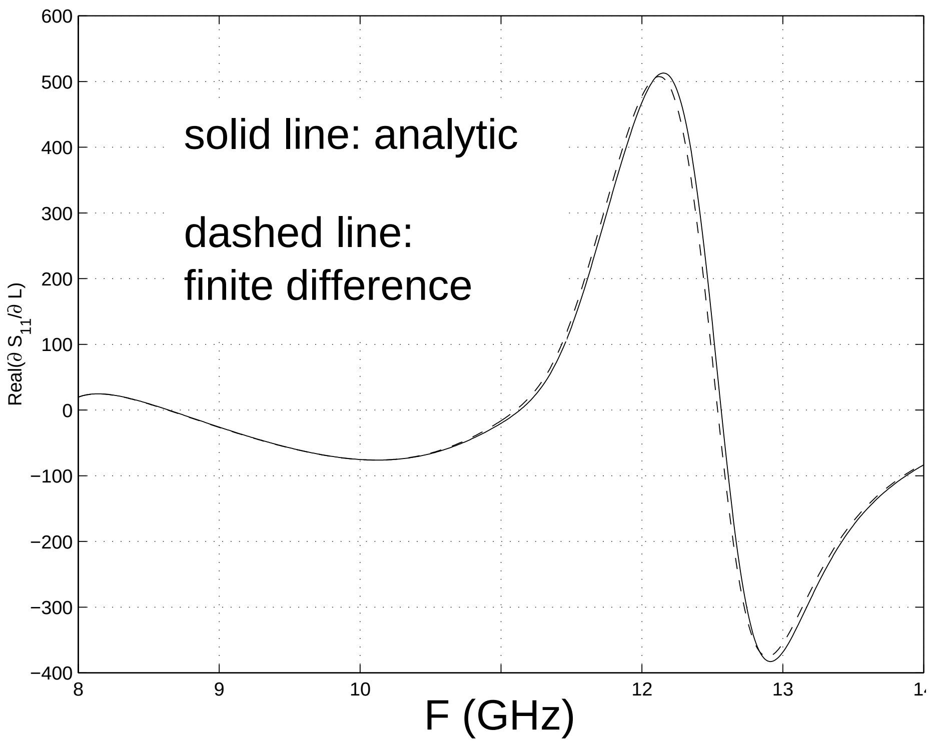
Within the CIET the reflection coefficient S_{11} is given in terms of the spectrum of the basis functions and the expansion coefficients a_i

$$S_{11} = -1 + \left[B^{bb}(1)\right]^T \{a\} = -1 + \sum_{i=1}^{M} a_i B_i^{bb}(1)$$

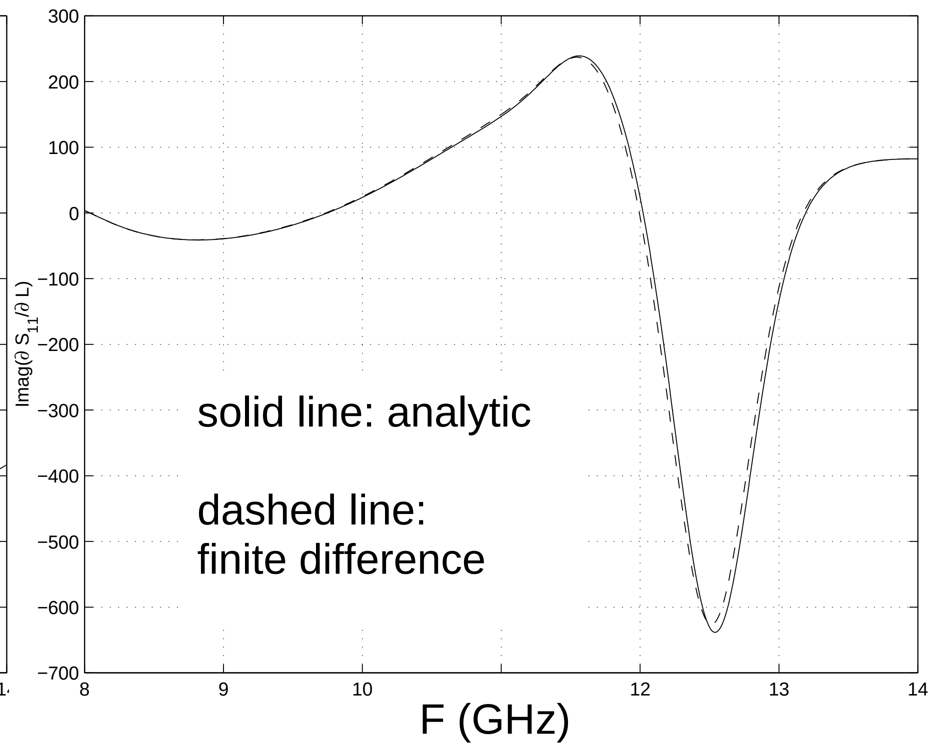
RESULTS

single E-plane stub in a rectangular waveguide

real part of $\partial S_{11}/\partial L$



imaginary part of $\partial S_{11}/\partial L$

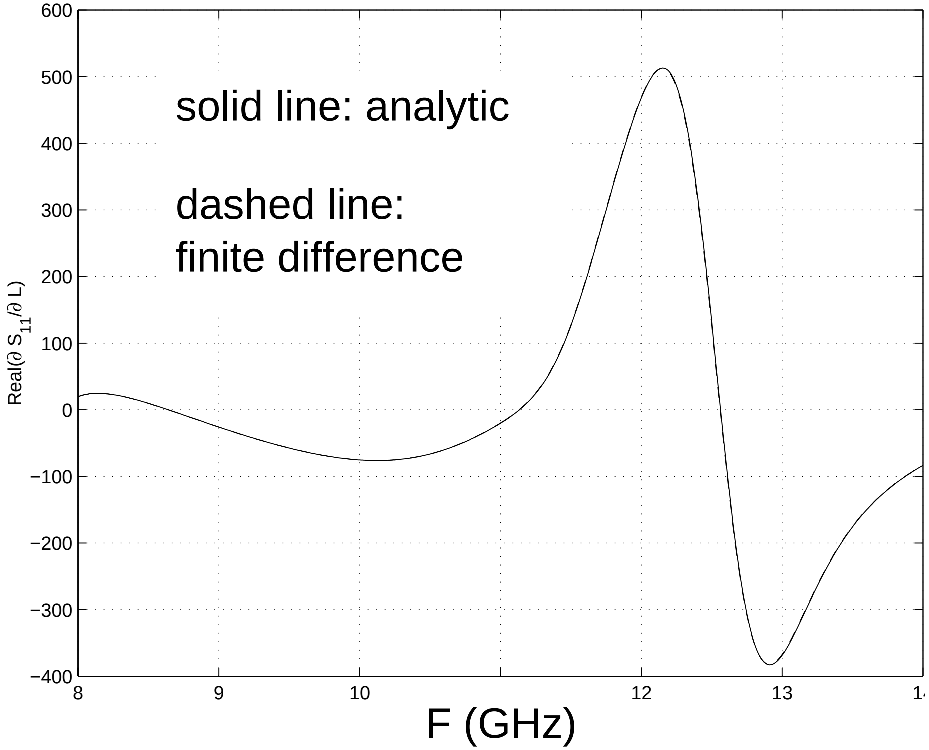


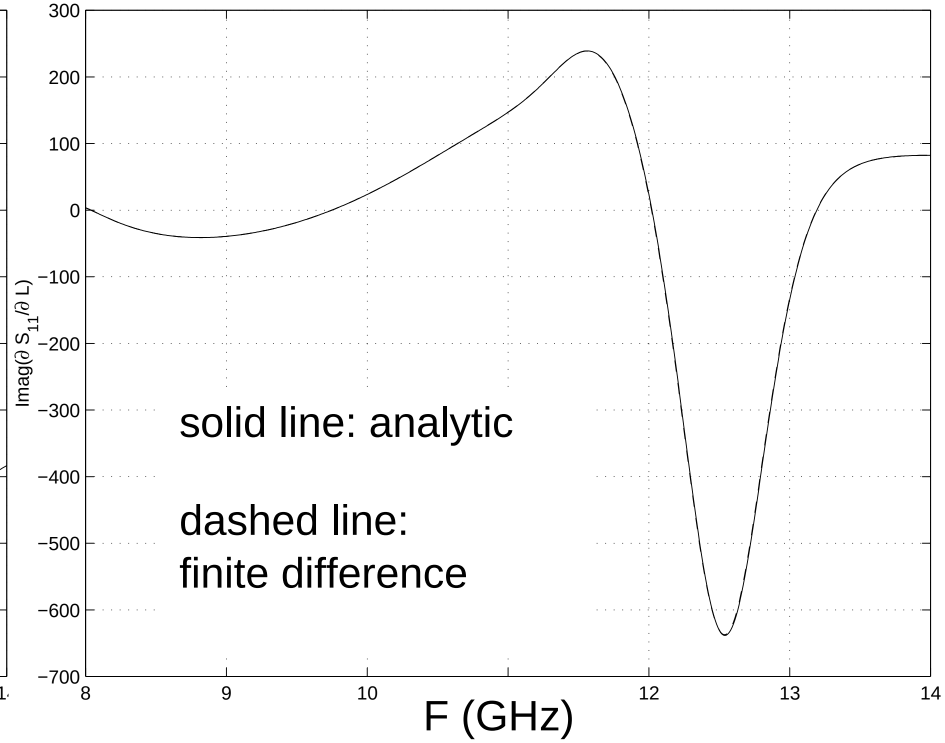
stub length L=a, finite difference increment $\partial L = a/100$

RESULTS

single E-plane stub in a rectangular waveguide

real part of $\partial S_{11}/\partial L$ | imaginary part of $\partial S_{11}/\partial L$

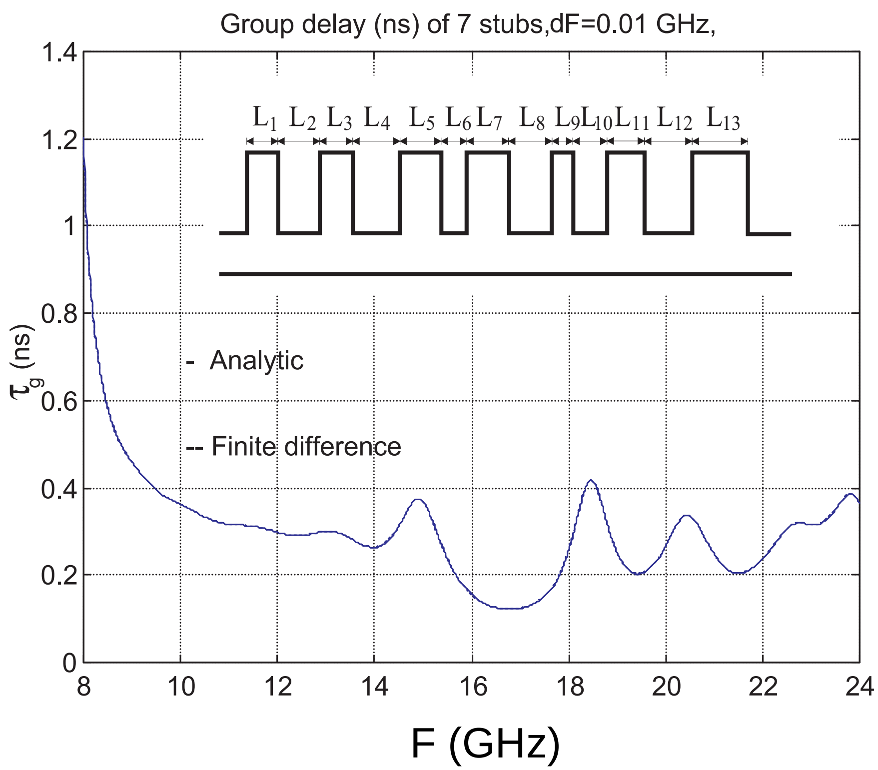




stub length L=a, finite difference increment $\partial L = a/1000$

RESULTS

7 E-plane stubs in a rectangular waveguide

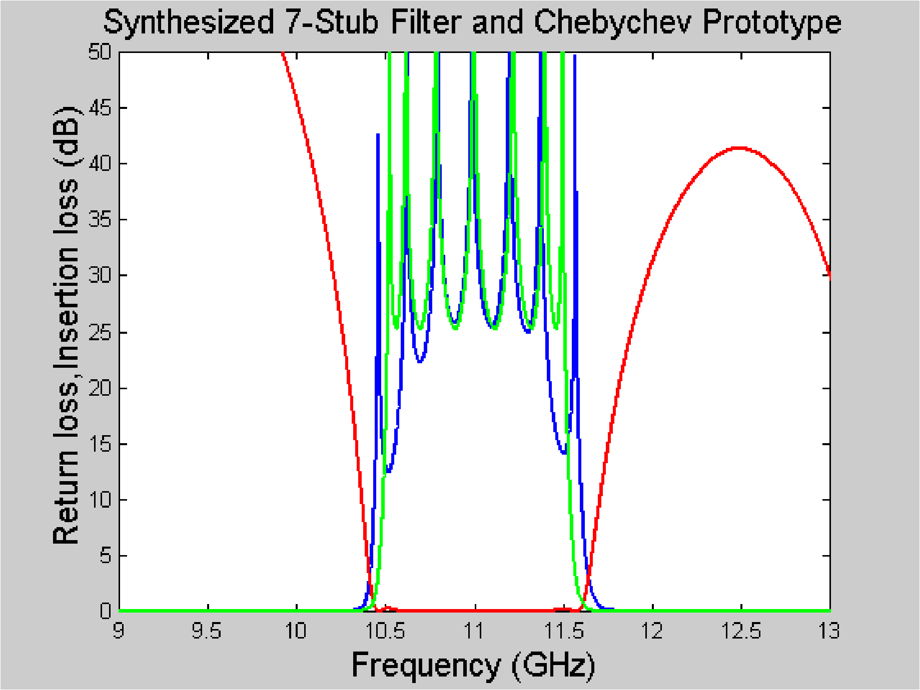


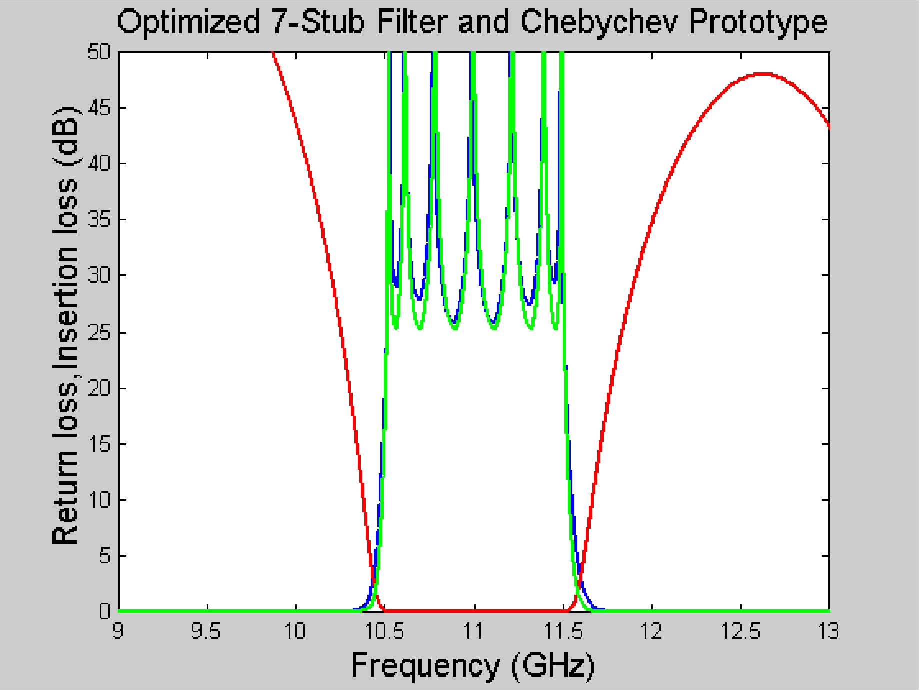
group delay

RESULTS

Bandpass Filter

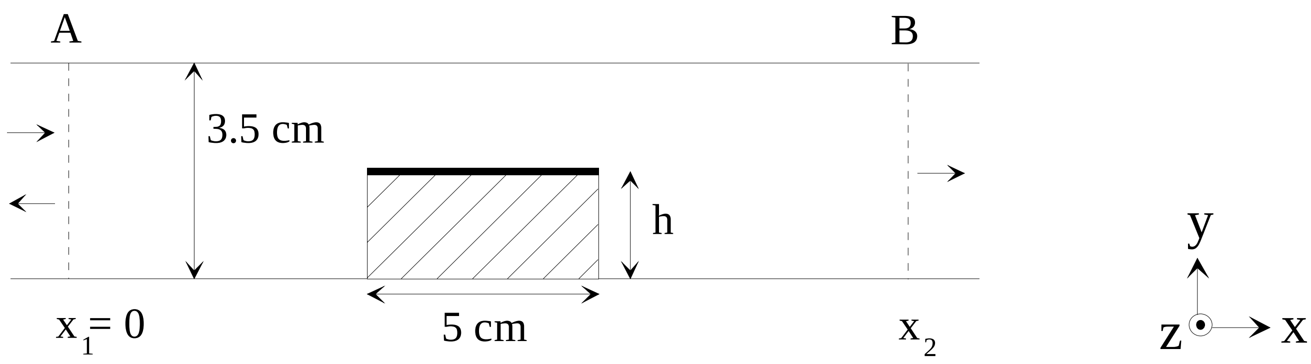
- Center frequency: 11 GHz
- Band width: 10%
- Minimum return loss in passband: 26 dB





SECOND EXAMPLE: FEM

Parallel plate waveguide with dielectric insert



Helmholtz equation, 2-D, cartesian coordinate system

$$\frac{\partial}{\partial x}\left(\frac{1}{\varepsilon_r}\frac{\partial H_z}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{1}{\varepsilon_r}\frac{\partial H_z}{\partial y}\right)+k_0^2\mu_r H_z=0$$ + boundary conditions

$S_{11} = fct(H_z)$ at plane A

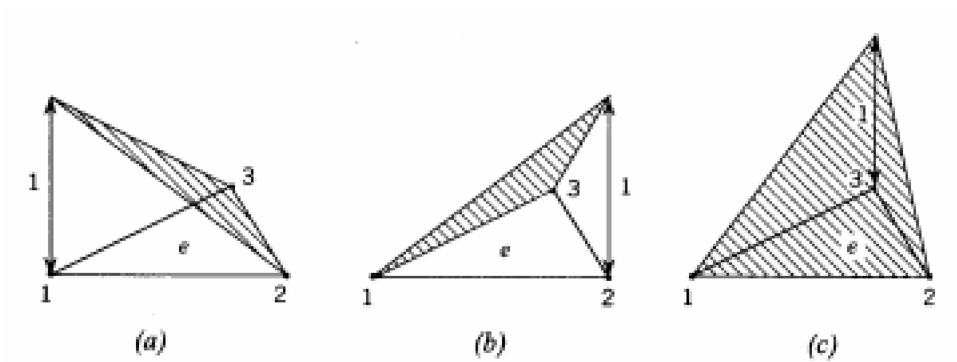
Use of a triangular mesh, linear interpolation functions

$$L_j^e(x,y)=\frac{1}{2\Delta^e}\left(a_j^e+b_j^e x+c_j^e y\right),\quad j=1,2,3$$

a_j^e, b_j^e, c_j^e : constant coefficients
= function of triangle coord.

example: $a_1^e = x_2^e y_3^e - y_2^e x_3^e$

Δ^e : area of a triangle



indices *i*: node number
e: element number

The H–Field within triangle e is:

$$H_z^e(x,y)=\sum_{j=1}^{3} L_j^e(x,y) H_{z,j}^e$$

Applying Galerkin Method the result is a matrix equation of the form:

$[K]\{H_z\} = \{b\}$ $[K]$: system matrix, with elements coming from elemental matrices $K_{ij}^e \Rightarrow K_{n(i,e),n(j,e)}$

$\{b\}$: excitation vector

elemental values:

$$K_{ij}^e = \frac{1}{4\Delta^e \varepsilon_r^e}\left(b_i^e b_j^e + c_i^e c_j^e\right) - k_0^2 \mu_r \frac{\Delta^e}{12}\left(1 + \delta_{ij}\right)$$

Solve matrix equation for unknown coefficients $\{H_z\}$

S_{11} can be formulated in terms of $\{H_z\}$ at cutting plane A:

$$S_{11} = \frac{H_z(x_1 = 0)}{H_0} - 1$$

ANALYTIC GRADIENT EVALUATION

$$S_{11} = \frac{H_z(x_1) - H_0 e^{-jk_0 x_1}}{H_0 e^{jk_0 x_1}} = \frac{H_z(x_1 = 0)}{H_0} - 1$$, x_1 is set to zero

$$\frac{\partial S_{11}}{\partial h} = \frac{1}{H_0} \frac{\partial H_z(x_1 = 0)}{\partial h}$$

→ calculate derivatives $\frac{\partial H_z(x_1 = 0)}{\partial h}$

linear set: $[K]\dfrac{\partial}{\partial h}\{H_z\} = \dfrac{\partial}{\partial h}\{b\} - \dfrac{\partial}{\partial h}[K]\{H_z\}$

$\dfrac{\partial H_z(x_1 = 0)}{\partial h}$ elements of $\dfrac{\partial}{\partial h}\{H_z\}$, nodes on cutting plane A.

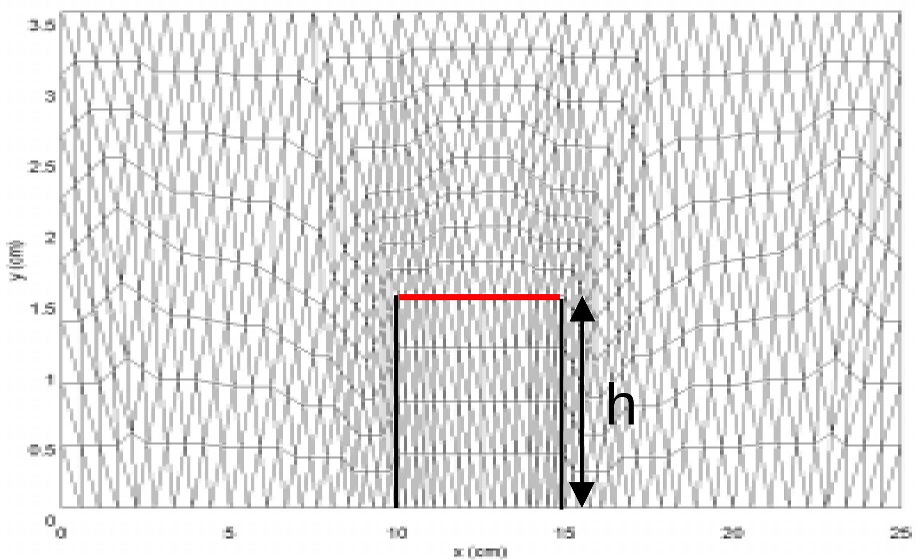
$\dfrac{\partial}{\partial h}\{b\} = 0$, because excitation is independent of h.

$\dfrac{\partial}{\partial h}[K]$, derivatives of the elements of the system matrix:

$$K_{ij}^e = \frac{1}{4\Delta^e \varepsilon_r^e}\left(b_i^e b_j^e + c_i^e c_j^e\right) - k_0^2 \mu_r \frac{\Delta^e}{12}\left(1 + \delta_{ij}\right)$$

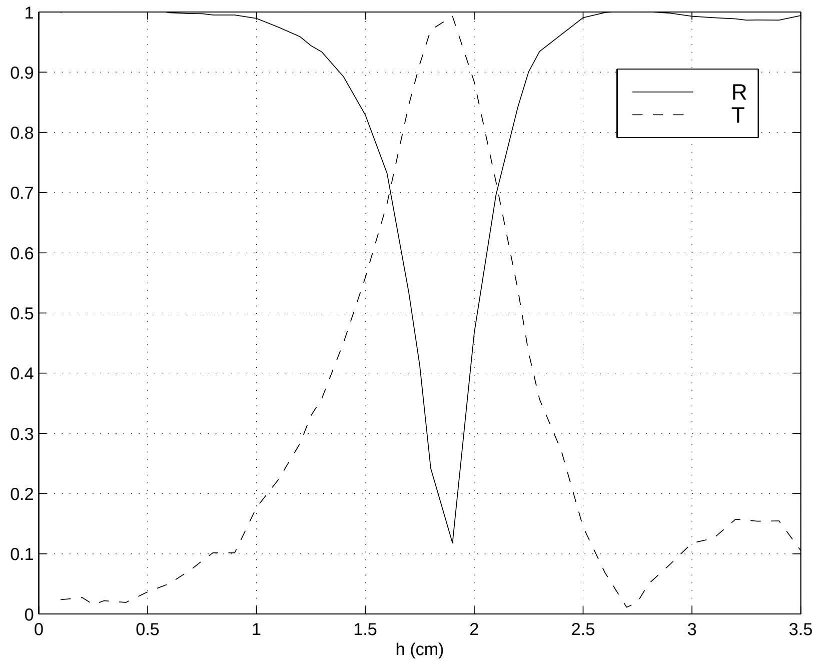
$$\frac{\partial}{\partial h}[K]_{ij}^e \;\Big|_{h=h_0} = \ldots.$$

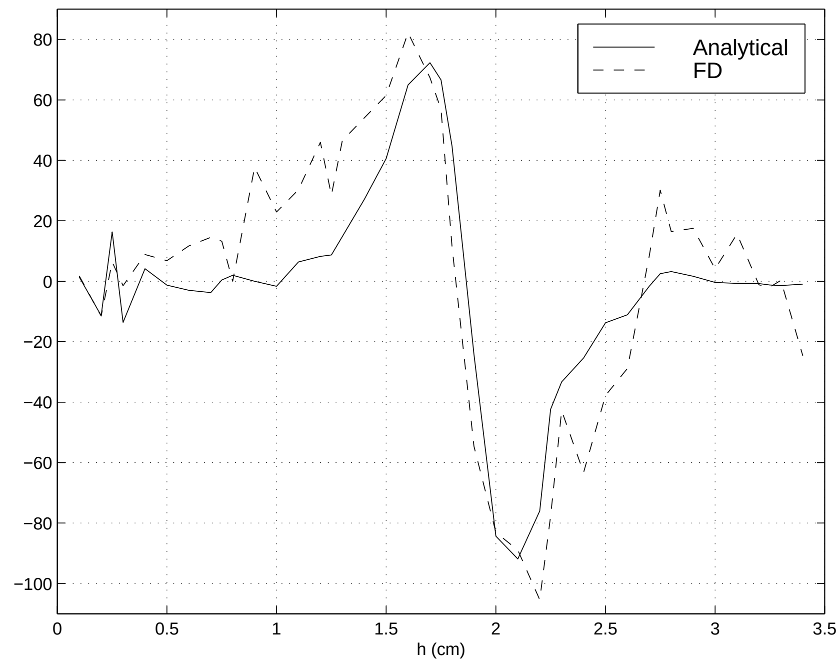
RESULTS



Mesh setting, triangular mesh with 1088 elements and 601 nodes

For calculation of $\frac{\partial}{\partial h}[K]_{ij}^{e}\ \Big|_{h=h_0}$ only nodes along the red edge needed.





$S_{11}(h), S_{21}(h)$

$\frac{\partial S_{11}}{\partial h}$ analytical

FD with $\Delta h = 0.1$ cm.

CONCLUSIONS

- novel technique to evaluate gradients for optimization
- general approach for problems that can be formulated in terms of general nonhomogeneous matrix equations (e.g. FEM, MoM)
- gradients are determined analytically (exact)
- no need for network representation of the problem
- no finite differencing
- only one function evaluation needed (only one mesh setting)
- no matrix inversion
- result: fast and accurate optimization of microwave structures
- excellent agreement between present approach and finite difference