

NOVEL ANALYTIC GRADIENT EVALUATION TECHNIQUES FOR OPTIMIZATION OF MICROWAVE STRUCTURES

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

Swiss Federal Institute of Technology, ETH Zentrum, Gloriastrasse 35, CH – 8092 Zürich,
<http://www.ifh.ee.ethz.ch>.

*Department of Electrical and Computer Engineering
University of Victoria, BOX 3055
Victoria, B.C. V8W 3P6
<http://www.ece.uvic.ca>

Introduction

ACCURATE design of modern and increasingly complex microwave devices and structures can often be achieved only through sophisticated optimization methods [1], [2]. It is generally accepted that efficient optimization methods are those exploiting not only the values of the cost function but its gradient and even its Hessian matrix as well. The gradient of the cost function is most of the time evaluated by repeated analyses at closely-spaced values of the independent variables; such a process can be prohibitively time consuming when the number of independent variables is large.

Recently, the Adjoint Network Method (ANM), which has been known for a long while in the area of circuit optimization, has been successfully used to analytically evaluate the gradient of cost functions for optimization of microwave filters and radiating structures [3], [4]. By its own nature, the ANM requires a network representation of the structure to be optimized (and its adjoint). The Mode-Matching Technique (MMT) was used to extract the network's matrix representation in form of, e.g., the admittance matrix [3] or the scattering matrix [4]. Starting from the scattering matrix of a network, including internal ports, a direct method of evaluating network sensitivities (gradient) was presented in [5].

Unfortunately more general numerical techniques such as the Method of Moments (MoM) [6] or the Finite Element Method (FEM) [7] do not directly yield a network representation in general.

Within these two methods, scattering problems are reduced to matrix equations of the form

$$[A][x] = [b]. \quad (1)$$

Here, $[A]$ is a $M \times M$ 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 MoM or the nodal values in FEM.

With these features in mind, the paper addresses the following questions:

1. *Is it possible to determine analytically the gradient of cost and response functions with respect to structural changes in the system without first deriving a network representation from $[x]$ and without finite differencing?*
2. *Is it possible to calculate group delays analytically from the matrix $[A]$ and excitation $[b]$?*

It is shown that, as long as the partial derivatives of the matrix $[A]$ and the excitation $[b]$ are known analytically, all relevant sensitivities can also be determined analytically.

Theory

We assume that the *analysis* of the structure to be optimized has been put in the form of equation (1). The optimization problem consists in determining the optimal values of the independent variables a_1, a_2, \dots, a_n such that the response function is equal to an optimal function x^{opt} . Note that the independent variables appear explicitly in both the matrix $[A]$ as well as the excitation $[b]$

in general. We also assume that the optimization problem consists in minimizing a cost function F which depends explicitly on the solution $[x]$. The partial derivative of the cost function with respect to a_i can be written in the form

$$\frac{\partial F}{\partial a_i} = \sum_{j=1}^M \frac{\partial F}{\partial x_j} \frac{\partial x_j}{\partial a_i} \quad (2)$$

Since the partial derivatives $\frac{\partial F}{\partial x_j}$ are known, or can be straightforwardly determined, the problem reduces to calculating the partial derivatives of the solution $[x]$ with respect to a_i . To evaluate these, we take the partial derivative of equation (1) to get

$$\frac{\partial[A]}{\partial a_i}[x] + [A] \frac{\partial[x]}{\partial a_i} = \frac{\partial[b]}{\partial a_i}. \quad (3)$$

Here, the derivative of a matrix is defined as a matrix whose elements are the derivatives of the original matrix. Rearranging this last equation and using equation (1), we get

$$\frac{\partial[x]}{\partial a_i} = [A]^{-1} \frac{\partial[b]}{\partial a_i} - [A]^{-1} \frac{\partial[A]}{\partial a_i} [A]^{-1} [b]. \quad (4)$$

This equation shows that the partial derivatives are all known once the inverse $[A]^{-1}$ is known along with the partial derivatives of the entries of the matrix $[A]$ and $[b]$. If the partial derivatives of $[A]$ and $[b]$ are known analytically, which is the case of a large number of structures, the gradient is then also known analytically from a *single* inversion of the matrix $[A]$. In the more general case, the partial derivatives of $[A]$ and $[b]$ can be computed from finite differences and still determine the gradient of the cost function from a *single* inversion of the matrix $[A]$, or equivalently from a *single* analysis of the structure.

Situations may arise where the matrix $[A]$ itself involves inversion of other matrices. In such a case, the following expression for the partial derivatives of the inverse of a matrix in terms of those of the matrix itself is to be used

$$\frac{\partial[B]^{-1}}{\partial a_i} = -[B]^{-1} \frac{\partial[B]}{\partial a_i} [B]^{-1} \quad (5)$$

We finally note that, if needed, the Hessian matrix can also be obtained from the present approach.

Results

The present approach is first applied to a simple structure to document the accuracy of the computed partial derivatives. We consider a single E-plane stub in a rectangular waveguide. The width of the stub is L . Assuming that the fundamental TE_{10} is incident on the stub, we propose to compute the partial derivative $\frac{\partial S_{11}}{\partial L}$ as a function of frequency using the present approach and conventional finite differencing. The formulation of this simple scattering problem within the Coupled-Integral-Equation Technique (CIET) is given in [8]. Edge-conditioned basis functions are used to accelerate convergence of the numerical solution. Figures 1 and 2 show the real and imaginary parts of this partial derivative as obtained from the present analytic approach (solid line) and finite difference with an increment $\delta L = a/100$. It is evident that both calculations agree well except for small deviations for frequencies between 11 and 13.8 GHz. However, when the increment is reduced to $\delta L = a/1000$, the two results agree and can not be distinguished as Figures 3 and 4 show.

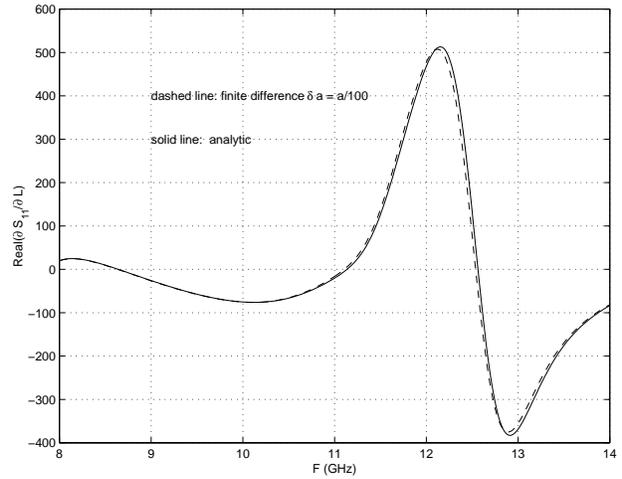


Fig. 1. Real part of $\frac{\partial S_{11}}{\partial L}$ as a function of frequency. Solid line: present method, dashed line: finite difference. $L = a = 2b = b_1 = 19.05mm$, $\delta L = a/100$.

The next example consists in a 7 stub E-plane structure. We first evaluate the partial derivatives of the scattering parameters of the TE_{10} mode with respect to the inner dimensions of the filters (13 in all). Table I shows these quantities as obtained

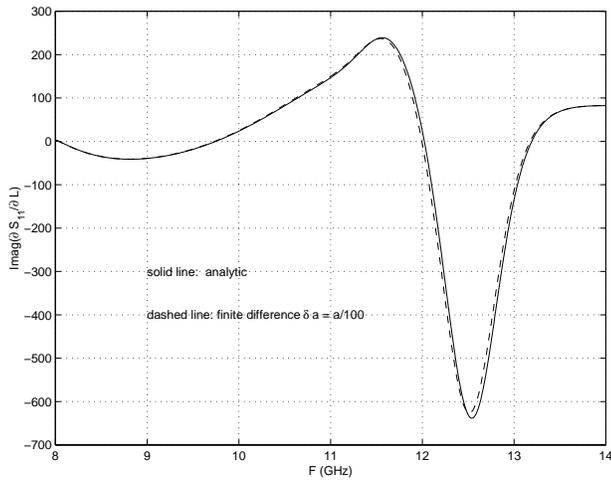


Fig. 2. Imaginary part of $\frac{\partial S_{11}}{\partial L}$ as a function of frequency. Solid line: present method, dashed line: finite difference. $L = a = 2b = b_1 = 19.05\text{mm}$, $\delta L = a/100$.

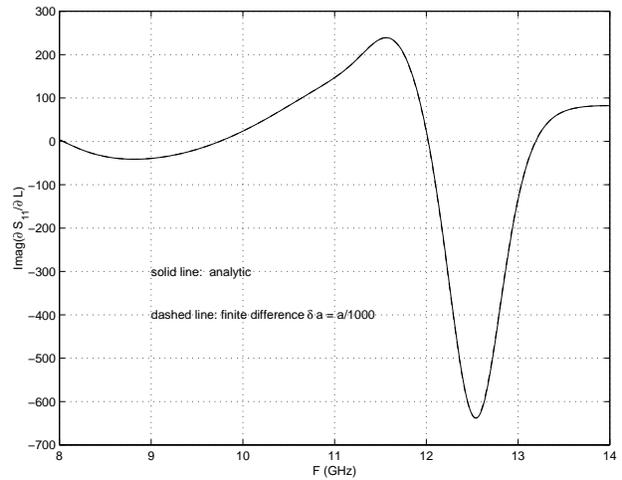


Fig. 4. Imaginary part of $\frac{\partial S_{11}}{\partial L}$ as a function of frequency. Solid line: present method, dashed line: finite difference. $L = a = 2b = b_1 = 19.05\text{mm}$, $\delta L = a/1000$.

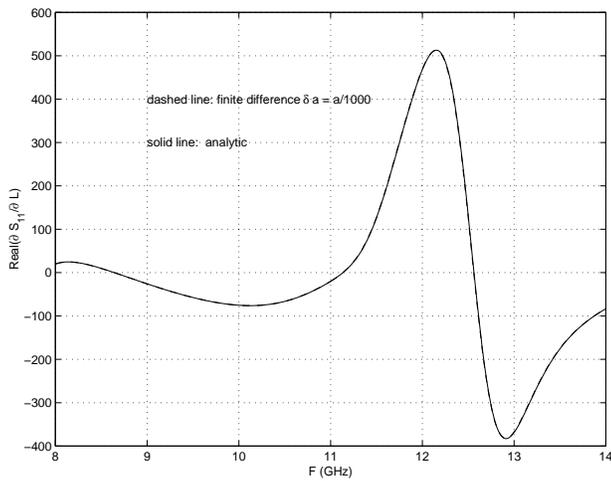


Fig. 3. Real part of $\frac{\partial S_{11}}{\partial L}$ as a function of frequency. Solid line: present method, dashed line: finite difference. $L = a = 2b = b_1 = 19.05\text{mm}$, $\delta L = a/1000$.

from the present approach as well as finite differencing for two values of the increment. It is again evident that the results from the conventional finite differencing converge to the results obtained from the proposed technique when small enough values of the increments in the independent variables are used. This constitutes an additional validation of this novel approach.

The present approach can also be applied to computing group delays. It is easily shown that the group delay is given by [1]

$$\tau_g = -\text{Im}\left[\frac{1}{S_{21}} \frac{\partial S_{21}}{\partial \omega}\right] \quad (6)$$

The group delay of of the 7-stub structure was

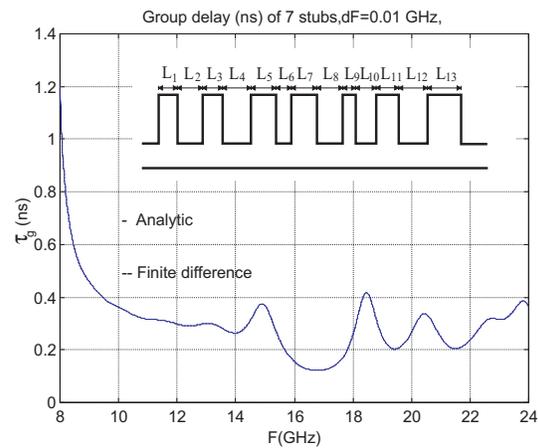


Fig. 5. Group delay of a 7-stub structure versus frequency. The analysis is performed using CIET with edge-conditioned basis functions [8].

computed using the present approach. Figure 5 shows that the results obtained from the present work (solid line) agree well with those from finite differencing the phase of S_{12} (dashed line). These results were obtained from 2 edge-conditioned basis functions at each discontinuity.

Table I
Partial derivatives $\frac{\partial |S_{11}|}{\partial L_i}$

i	This method	FD ¹	FD ²
1	-126.4792	-127.0087	-126.4798
2	0.7323	0.8009	0.7324
3	-126.9130	-127.4637	-126.9135
4	1.0517	1.1316	1.0518
5	-127.1495	-127.7106	-127.1500
6	1.2017	1.2859	1.2018
7	-127.2251	-127.7893	-127.2257
8	1.2017	1.2859	1.2018
9	-127.1495	-127.7106	-127.1500
10	1.0517	1.1316	1.0518
11	-126.9130	-127.4637	-126.9135
12	0.7323	0.8009	0.7324
13	-126.4792	-127.0087	-126.4798

¹ $\delta L_i = 0.1$ mm

² $\delta L_i = 0.1$ μ m

Conclusions

A novel technique to evaluate gradients for optimization of microwave structures and devices was presented. The approach only requires that the problem be formulated in terms of a general non-homogeneous matrix equation such as encountered in the Moment Method or the Finite Element Method. All partial derivatives are determined from a single analysis of the structure. Numerical results show excellent agreement between the present approach and the finite difference method when small enough increments are used.

References

1. K. C. Gupta, R. Garg and R. Chadha, "Computer Aided Design of Microwave Circuits", Artech House, 1981.
2. J. W. Bandler, Computer-Aided Circuit Optimization, Chapter 6 of "Modern Filter Theory and Design", G.C. Temes and S.K. Mittra, Ed., Wiley, New York, 1973.
3. F. Alessandri, M. Mongiardo and R. Sorrentino, "New efficient fullwave optimization of microwave circuits by the adjoint network method", IEEE Microwave Guided Wave Lett., vol. 3, pp. 414-416, Nov. 1993.
4. M. Mongiardo and T. Ravanelli, "Automated design of corrugated feeds by the adjoint network method", IEEE Trans. Microwave Theory Techn.,

vol. 45, pp. 787-792, May 1997.

5. J. W. Bandler and R. E. Seviara, "Direct method for evaluating scattering-matrix sensitivities", Electron. Lett., vol. 6, pp. 773-774, Nov. 1970.
6. R. F. Harrington, Field Computation by the Moment Method, Krieger, Malabar, FL., 1987.
7. J. B. Davies, The Finite Element Method, Chapter 2 of "Numerical Techniques for Microwave and Millimeterwave Passive Structures", T. Itoh Ed., Wiley, New York, 1989.
8. S. Amari, J. Bornemann and R. Vahldieck, "On the acceleration of the coupled-integral-equations technique and its application to multistub E-plane discontinuities", J. Electromag. Waves Appl., vol. 13, pp.537-552, 1999 (in press).