SENSITIVITY ANALYSIS WITH FINITE-ELEMENT METHOD FOR MICROWAVE DESIGN AND OPTIMIZATION

SENSITIVITY ANALYSIS WITH FINITE-ELEMENT METHOD FOR MICROWAVE DESIGN AND OPTIMIZATION

By

DONGYING LI, B. Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Applied Science

McMaster University

MASTER OF APPLIED SCIENCE (2006) McMASTER UNIVERSITY

(Electrical and Computer Engineering) Hamilton, Ontario

TITLE: **SENSITIVITY ANALYSIS WITH FINITE-ELEMENT METHOD FOR MICROWAVE DESIGN AND OPTIMIZATION**

AUTHOR: DONGYING LI

B. Sc. (Electrical Engineering), Shanghai Jiao Tong University

SUPERVISOR: Dr. Natalia K. Nikolova, Associate Professor

Department of Electrical and Computer Engineering

Dipl. Eng. (Technical University of Varna)

Ph. D. (University of Electro-Communication)

P. Eng. (Ontario)

CO-SUPERVISOR: Dr. James P. Reilly, Professor

Department of Electrical and Computer Engineering

B. A. Sc. (University of Waterloo)

M. Eng. (McMaster University)

Ph. D. (McMaster University)

P. Eng. (Ontario)

NUMBER OF PAGES: xiii, 84

ABSTRACT

The thesis proposes a novel method for the computation of the design sensitivity of microwave network parameters. The approach is based on the finite-element method. When combined with the iterative update method (the Broyden method) during the gradient-based optimization process, the approach requires practically no overhead for the computation of the response Jacobian, thus accelerating the optimization.

The efficiency and accuracy of the gradient-based optimization and the tolerance analysis greatly depend on the computation of the design sensitivity. However, common commercial full-wave electromagnetic solvers do not provide sensitivity information. With them, the design sensitivities are computed from the response themselves using finite-difference or higher-order approximations at the response level. Consequently, for each design parameter of interest, at least one additional full-wave analysis is performed.

The proposed self-adjoint sensitivity analysis (SASA) is so far the most efficient way to extract the sensitivity information for the network parameters with the finite-element method. As an improvement of the adjoint-variable method (AVM), it eliminates the additional system analyses. With one single

full-wave analysis, the sensitivities with respect to all design parameters are computed. This significantly improves the efficiency of the sensitivity computations.

When employed in gradient-based optimization, the computational overhead of the SASA can be further reduced. Instead of the finite-difference approximation, the system matrix derivatives are updated iteratively using the Broyden update. This reduces the computational overhead of the sensitivity analysis to practically zero. Further, several switching criteria between the Broyden update and the finite-difference approximation of the system matrix derivatives is proposed to guarantee the robust convergence of the optimization algorithm. This leads to our Broyden/finite-difference SASA (B/FD-SASA).

The efficiency in terms of CPU time as well as the accuracy of the SASA is verified by several numerical examples, where the reference results are provided through the traditional finite-difference approximations. Also, the efficiency of the B/FD-SASA is validated by a filter design example and a microwave imaging example, with implementations exploiting different gradient-based optimization algorithms.

ACKNOWLEDGEMENT

I wish to express my gratitude to Dr. N. K. Nikolova for her guidance and supervision during the two years of my master's study. Her selfless advices, valuable opinions as well as all her patience impact me greatly on my study and research in McMaster University, and what I learn from her will become a life-long wealth for my future research work.

I also want to thank Dr. M. H. Bakr for his advices and suggestions in the area of the optimization techniques. His insightful opinion is valuable and very important for the completion of this thesis.

Also I want to thank Jiang Zhu, Wenhuan Yu, Yan Li and Ying Li for spending two years closely as a research team in the Computational Electromagnetics Laboratory. Their co-operation and encouragement create a friendly and vigorous team-working environment from which I benefit a lot.

Finally I wish to thank my dearest mother, my family and all my friends for all the precious love and support they give to me.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AVM	Adjoint Variable Method
B-SASA	Broyden Self-adjoint Sensitivity Analysis
B/FD-SASA	Broyden/Finite-Difference Self-Adjoint Sensitivity Analysis
BFGS	Broyden-Fletcher-Goldfarb-Shannon
EM	Electromagnetics
FD	Feasible Adjoint Sensitivity Technique
FD-SASA	Finite Difference
FDTD	Finite-Difference Time Domain
FEM	Finite Element Method
MoM	Method of Moment
SQP	Sequential Quadratic Programming
TLM	Transmission-Line Method
TR	Trust Regions

CHAPTER 1

INTRODUCTION

Design sensitivity analysis is aimed at the evaluation of the derivatives of the system response with respect to the design parameters. In microwave structure design, these are typically shape and material parameters. The overall design process, including design optimization, yield and tolerance analysis, as well as statistical analysis, can greatly benefit from the knowledge of the sensitivity information [1]. Sensitivity analysis technique is crucial, especially in numerical microwave problems which are electrically large and where analytical sensitivity solutions are impossible.

The adjoint-variable method (AVM) technique is known as the most efficient method for the response sensitivity computation of complex linear and nonlinear problems [1], [2]. The AVM has a history of applications in the area of control theory [1], as well as in finite-element analysis in structural [1], [3] and electrical engineering [4], [5], [6], [7], [8], [9]. The application of the AVM in the microwave area emerges at the early 1970's with the computation of the network sensitivities based on voltage/current state variables and responses [5],

[6], [7] and *S*-parameters [8], [10]. Yet, the computation is based on Tellegen's theory of circuit concepts, not field solutions.

Recently, an adjoint-variable method has been proposed for the sensitivity computation in numerical electromagnetic (EM) problems both in the time domain (the transmission line method, TLM [11], and the finite-difference in time-domain method, FDTD [12]) and in the frequency domain (the frequency domain TLM [13], [14], the method of moments, MoM, and the finite-element method, FEM [15]). For the frequency-domain application, this method uses the finite-difference approximation of the system matrix derivatives, which eliminates the need for analytical pre-processing [15]. The computational load is significantly reduced without sacrificing the accuracy.

However, all traditional AVM techniques still require one more full-wave analysis (the adjoint system analysis) in addition to the original system analysis. Moreover, the excitation in the adjoint system analysis depends on the response and its relation to the solution of the original problem. This may cause potential difficulties in the formulation of this additional analysis problem, and its implementation in the framework of a commercial high-frequency CAD package.

In this thesis, a general approach, named self-adjoint sensitivity analysis (SASA), is proposed for the sensitivity analysis of microwave network parameters, i.e., *S*-parameters. With this approach, we eliminate the additional (adjoint) system analysis and further improve the efficiency of the sensitivity

computation [16], [17], [18]. This possibility has been first discussed in Akel *et al.* [19], in the case of the FEM formulation based on the tetrahedral edge elements. With a commercial high-frequency FEM solver, this standalone algorithm runs independently of the underlying analysis algorithm. The only information it needs is the field solution at specific grid points. Thus, it is easily integrated into any kind of design automation process with a commercial solver.

The most promising application of the SASA is in the area of gradient-based optimization [15], [20]. The optimization process can be significantly accelerated with the Jacobian provided by the SASA, since the computation of the Jacobian and/or Hessian is the bottleneck of the optimization efficiency with full wave EM solvers.

In this thesis, the advantage of the SASA is validated by its application in gradient based optimization. Notably, the efficiency of the SASA based optimization can be further improved by applying the Broyden update [21] at the level of the system matrix derivatives. As the overhead of the Broyden update is negligible compared to the finite-difference approximation of the system matrix, the computational load for the Jacobian calculation is practically zero.

The accumulation of inaccuracies during the iterative Broyden update of the system matrix derivatives may lead to wrong solutions, or even to divergence of the optimization process. To guarantee robustness of the convergence, the computation of the system matrix derivatives should switch between the fast

iterative method and the robust finite-difference approximation. Certain switching criteria are suggested in this thesis [22]. The resulting algorithm is named the Broyden/finite-difference SASA (B/FD-SASA). The B/FD-SASA is validated by the optimization of a microwave filter and by the solution of an inverse imaging problem using various gradient based optimization algorithms.

Chapter 2 introduces the basics of the AVM methodology. We review the basic sensitivity expression of the AVM and discuss its implementation in the sensitivity analysis of complex linear systems. The computational overhead of the traditional AVM is analyzed and compared with the overhead of the finite-difference approximation.

Chapter 3 introduces the SASA with the finite-element method and explores its application in *S*-parameter sensitivity computation for microwave structures. The general SASA expression is derived from the 3-D FEM formulation. The algorithm is validated by a rectangular waveguide bend example and a dielectric coupling filter design example. Error estimation and computational overhead evaluation are conducted. Comparisons with the traditional finite-difference approximation are provided.

Chapter 4 integrates the SASA method with the Broyden update in the framework of gradient-based optimization. The B/FD-SASA method is implemented through switching criteria which ensure the reliability of the sensitivity results. The B/FD-SASA method is validated through two

optimization examples. The convergence of the optimization using this method is compared with that using FD-SASA as well as the finite differences at the response level.

An overall conclusion is made in Chapter 5 and suggestions for future development are given.

The contributions of this research can be summarized as follows:

(1) Development of an efficient sensitivity computation algorithm with the finite-element method, i.e., SASA [16], [18].

(2) Implementation of the SASA algorithm with the commercial finite-element solver FEMLAB® [16], [18].

(3) Validation of the efficiency of SASA through numerical design examples [16].

(4) Development of a Broyden-update-based sensitivity computation algorithm for gradient based optimization [22].

(5) Validation of the efficiency of the sensitivity computation algorithm with different optimization algorithms [22].

REFERENCES

[1] D. G. Cacuci, *Sensitivity & Uncertainty Analysis, Volume 1: Theory*. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[2] A. D. Belegundu and T.R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice Hall, 1999.

[3] E. J. Haug, K.K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando: Academic Press Inc., 1986.

[4] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, pp.318-323, Aug. 1969.

[5] J. W. Bandler and R. E. Seviora, "Current trends in network optimization," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp.1159-1170, Dec. 1970.

[6] V. A. Monaco and P. Tiberio, "Computer-aided analysis of microwave circuits," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp.249-263, Mar. 1974.

[7] J. W. Bandler, "Computer-aided circuit optimization," in *Modern Filter Theory and Design*, G. C. Temes and S. K. Mitra, Eds. New York: Wiley, 1973, ch. 6.

[8] K. C. Gupta, R. Garg, and R. Chadha, *Computer-Aided Design of Microwave Circuits*. Norwood, MA: Artech House, 1981.

[9] J. Vlach and K. Singhal, *Computer Methods for Circuit Analysis and Design*. New York: Van Nostrand, 1983.

[10] J. W. Bandler and R. E. Seviora, "Wave sensitivities of networks," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 138-147, Feb. 1972.

[11] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for time-domain TLM with wide-band Johns matrix boundaries," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 678-685, Feb. 2004.

[12] N. K. Nikolova, H. W. Tam, and M. H. Bakr, "Sensitivity analysis with the FDTD method on structured grids," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 1207-1216, Apr. 2004.

[13] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for frequency domain TLM problems with conducting boundaries," *IEEE Microwave and Wireless Component Letters*, vol. 13, pp. 408-410, Sept. 2003.

[14] S. M. Ali, N. K. Nikolova, and M. H. Bakr, "Central adjoint variable method for sensitivity analysis with structured grid electromagnetic solvers," *IEEE Trans. Magnetics*, vol. 40, pp. 1969-1971, Jul. 2004.

[15] N. K. Georgieva, S. Glavic, M. H. Bakr and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751-2758, Dec. 2002.

[16] N. K. Nikolova, J. Zhu, D. Li, M. Bakr, and J. Bandler, "Sensitivities of network parameters with electromagnetic frequency domain simulator," *IEEE Trans. Microwave Theory Tech.* vol. 54, pp. 670-681, Feb, 2006.

[17] N. K. Nikolova, J. Zhu, D. Li, et al., "Extracting the derivatives of network parameters from frequency domain electromagnetic solutions," *General Assembly of Intl. Union of Radio Science CDROM*, Oct. 2005.

[18] D. Li and N. K. Nikolova, "*S*-parameter sensitivity analysis of waveguide structures with FEMLAB," *COMSOL Multiphysics Conf. 2005 Proceedings*, Cambridge, MA, Oct. 2005, pp. 267-271.

[19] H. Akel and J. P. Webb, "Design sensitivities for scattering-matrix calculation with tetrahedral edge elements," *IEEE Trans. Magnetics*, vol. 36, pp. 1043-1046, Jul. 2000.

[20] N. K. Nikolova, *et al.*, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antenna Propagat.*, vol. 52, pp. 2147-2157, Aug. 2004.

[21] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, pp. 577-593, 1965.

[22] D. Li, N. K. Nikolova, and M. H. Bakr, "Optimization using Broyden-update self-adjoint sensitivities," *IEEE Antenna Propagat. Symposium 2006*, accepted.

CHAPTER 2

METHODOLOGY OF THE ADJOINT VARIABLE METHOD

2.1 INTRODUCTION

Commercial high-frequency electromagnetic (EM) solvers usually do not compute sensitivity information, i.e., the Jacobian of the objective function, with respect to the design parameters. For design purposes, when sensitivity information is required, a finite-difference approximation at the response level is usually performed as a simple although inefficient way to obtain the response derivatives. The finite-difference approximation is highly inefficient in numerical computations, since it requires at least N+1 additional full system analyses for a problem with N designable parameters [1]. With higher-order approximations, the number of analyses increases. The feasibility of this approach becomes questionable when the design-variable space is large.

The adjoint-variable method (AVM) is proved to be the most efficient method for sensitivity analysis [1], [2], [3], as it requires only one additional

system analysis to compute all sensitivities. The additional analysis is known as adjoint system analysis, with the adjoint system matrix being the transpose of the system matrix of the original problem [1], [3], [4]. Thus, the AVM improves the efficiency of the sensitivity computation by a factor of N in comparison with the forward or backward finite-difference approximations. The performance of the AVM has been validated in control theory [3], structural engineering [4], and in circuit and computational EM applications in electrical engineering [3], [5].

In this chapter, we give a brief introduction into the methodology of the AVM, especially its applications with frequency-domain numerical EM solvers. Most of the discussion in this chapter and in the rest of the thesis focuses on applications with the finite-element method (FEM).

In Section 2.2, we present the concept of the frequency-domain AVM for linear systems. Then, we give a general discussion on the sensitivity analysis of complex linear systems in Section 2.3. Section 2.4 discusses the difficulties in the implementation of the AVM with commercial solvers. The efficiency and the required computational resources are discussed in Section 2.5 along with a comparison with the finite-difference approximation at the response level. The advantages and the drawbacks of the AVM are also discussed.

2.2 FREQUENCY-DOMAIN ADJOINT-VARIABLE METHOD

After proper discretization, a time-harmonic linear EM problem (by linear, we refer to the fact that the problem has linear material properties) can be written as a linear system of equations [6]:

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}. \tag{2.1}$$

Here, $\boldsymbol{A}$ is the M by M system matrix, $\boldsymbol{x}$ is the 1 by M state variable vector, in FEM, known as the field solution vector, and $\boldsymbol{b}$ is the 1 by M excitation vector, which can be derived from the electromagnetic sources and the inhomogeneous boundary conditions according to the FEM formulation. The system matrix is a function of the vector of design (shape or material) parameters $\boldsymbol{p}$, i.e., $\boldsymbol{A}(\boldsymbol{p})$. Thus, the field solution vector $\boldsymbol{x}$ is an implicit function of $\boldsymbol{p}$.

For sensitivity analysis purposes, we need to determine the gradient of a user defined response function $f(\boldsymbol{p},\bar{\boldsymbol{x}}(\boldsymbol{p}))$ with respect to $\boldsymbol{p}$ at the field solution $\bar{\boldsymbol{x}}$ of (2.1):

$$\nabla_{\boldsymbol{p}} f(\boldsymbol{p},\bar{\boldsymbol{x}}(\boldsymbol{p})) \text{ subject to } \boldsymbol{A}\bar{\boldsymbol{x}}=\boldsymbol{b}. \tag{2.2}$$

Here, the gradient of the response function $f(\boldsymbol{p},\bar{\boldsymbol{x}}(\boldsymbol{p}))$ is defined as a row vector [4], [6]

$$\nabla_{\boldsymbol{p}} f=\left[\frac{\partial f}{\partial p_1}\ \frac{\partial f}{\partial p_2}\ \cdots\ \frac{\partial f}{\partial p_N}\right]. \tag{2.3}$$

Note that the response function $f(\boldsymbol{p},\bar{\boldsymbol{x}}(\boldsymbol{p}))$ is formulated so that it may have an explicit dependence on the design variables in addition to its implicit dependence on $\boldsymbol{p}$ through the field solution vector $\boldsymbol{x}$. In some situations, both dependences exist.

We first constrain our problem as a real-number problem, i.e., both the system matrix $\boldsymbol{A}(\boldsymbol{p})$ and the response function $f(\boldsymbol{p},\bar{\boldsymbol{x}}(\boldsymbol{p}))$ are real. The sensitivity analysis with complex numbers is discussed in the next section.

According to [7], an AVM sensitivity expression can be formulated as:

$$\nabla_p f=\nabla_p^e f+\hat{\boldsymbol{x}}^T\left[\nabla_p \boldsymbol{b}-\nabla_p(\boldsymbol{A}\bar{\boldsymbol{x}})\right]. \tag{2.4}$$

Here, we divide the gradient of the response function into two parts:

$$\nabla_p f=\nabla_p^e f+\nabla_p f\cdot\nabla_p \boldsymbol{x}. \tag{2.5}$$

$\nabla_p^e f$ stands for the explicit dependence of the response function f on the design variables $\boldsymbol{p}$, and $\nabla_x f\cdot\nabla_p \boldsymbol{x}$ reflects the implicit dependence on $\boldsymbol{p}$ through the field solution $\boldsymbol{x}$. The vector $\hat{\boldsymbol{x}}$ is the adjoint solution, which is the solution of the adjoint system of equations [4], [8]:

$$\boldsymbol{A}^T\hat{\boldsymbol{x}}=\left[\nabla_x f\right]^T. \tag{2.6}$$

The adjoint system excitation is

$$\hat{\boldsymbol{b}}=\left[\nabla_x f\right]^T. \tag{2.7}$$

Here, we need to compute the original system solution $\bar{x}$, the adjoint system solution $\hat{x}$, and the derivative of the system matrix with respect to each design variable $\partial A/\partial p_i$, i=1,...,N. Thus, with only two full-wave simulations, namely, the original system simulation and the adjoint system simulation, we can compute the sensitivities.

It is important to notice that we need to compute the system matrix derivative with respect to each design parameter $\partial A/\partial p_i$, i=1,...,N. In some rare cases, the matrix derivatives may be analytically available [9], [10]. Then, the sensitivities are exact. According to [9], the time needed for the analytical computation of one system matrix derivative is comparable with one system matrix fill. Thus, the sensitivity computation of a problem with N design parameters leads to an overhead of N matrix fills in addition to the original and adjoint system analyses.

For most of the full-wave EM analysis methods, the system matrix derivatives are either not analytically available or too complicated to be analytically derived for general design software. In these cases, the system matrix derivatives are computed by the finite-difference approximation [11]:

$$\frac{\partial \boldsymbol{A}(\boldsymbol{p})}{\partial p_i} = \frac{\boldsymbol{A}(\boldsymbol{p}+\Delta p_i \cdot \boldsymbol{e}_i) - \boldsymbol{A}(\boldsymbol{p})}{\Delta p_i}. \tag{2.8}$$

Here, $\boldsymbol{e}_i$ is the unit vector whose ith element equals 1 and all others equal 0:

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} i\text{th element} \tag{2.9}$$

and Δp_i is the finite-difference perturbation of the ith design variable. This approximation also requires N additional matrix fills, similarly to the exact method. Our studies have shown that the accuracy for the sensitivity computation using this system-matrix-level finite-difference approximation is satisfactory, with a relative error well below 1%, compared with the exact sensitivity computation [1].

2.3 SENSITIVITIES OF COMPLEX LINEAR SYSTEMS

The derivations in the above section apply to real-number problems only. However, in electromagnetic frequency-domain sensitivity analysis, the system equations are complex, and, often, the responses are complex, too. It can be shown that the sensitivity formula in the complex case can be derived from the real-number sensitivity formula.

A complex linear system of equations in the form of (2.1) can be reformulated in a real-valued form [11]:

$$\begin{bmatrix} \Re\boldsymbol{A} & -\Im\boldsymbol{A} \\ \Im\boldsymbol{A} & \Re\boldsymbol{A} \end{bmatrix} \begin{bmatrix} \Re\boldsymbol{x} \\ \Im\boldsymbol{x} \end{bmatrix} = \begin{bmatrix} \Re\boldsymbol{b} \\ \Im\boldsymbol{b} \end{bmatrix}. \tag{2.10}$$

Here, $\Re$ and $\Im$ stand for the real and the imaginary parts of a matrix, respectively. We can re-write (2.10) as

$$\boldsymbol{A}_r \boldsymbol{x}_r = \boldsymbol{b}_r \tag{2.11}$$

where

$$\boldsymbol{A}_r = \begin{bmatrix} \Re\boldsymbol{A} & -\Im\boldsymbol{A} \\ \Im\boldsymbol{A} & \Re\boldsymbol{A} \end{bmatrix}, \quad \boldsymbol{x}_r = \begin{bmatrix} \Re\boldsymbol{x} \\ \Im\boldsymbol{x} \end{bmatrix}, \quad \boldsymbol{b}_r = \begin{bmatrix} \Re\boldsymbol{b} \\ \Im\boldsymbol{b} \end{bmatrix}. \tag{2.12}$$

The size of the real-valued system of equations is twice the size of the original one.

With this real-valued system of equations, the AVM sensitivity expression for a real-valued response f becomes:

$$\nabla_p f = \nabla_p^e f + \hat{\boldsymbol{x}}_r^T \left[\nabla_p \boldsymbol{b}_r - \nabla_p (\boldsymbol{A}_r \bar{\boldsymbol{x}}_r) \right]. \tag{2.13}$$

Here, $\hat{\boldsymbol{x}}_r$ is the solution of the corresponding adjoint system of equations

$$\boldsymbol{A}_r^T \hat{\boldsymbol{x}}_r = \hat{\boldsymbol{b}}_r = \left[\nabla_{x_r} f \right]^T. \tag{2.14}$$

The real-valued adjoint system excitation $\hat{\boldsymbol{b}}_r$ is

$$\left[\nabla_{x_r} f \right]^T = \left[\nabla_{\Re x} f \;\; \nabla_{\Im x} f \right]^T. \tag{2.15}$$

The adjoint system of equations can be written in a complex form as

$$\boldsymbol{A}^H \hat{\boldsymbol{x}} = \hat{\boldsymbol{b}} = \left[\nabla_x f \right]^H \tag{2.16}$$

where

$$\nabla_x f = \nabla_{\Re x} f + j \nabla_{\Im x} f. \tag{2.17}$$

Note that here we use the Hermitian transpose of $\boldsymbol{A}$, which is the complex conjugate transpose. For complex matrices, the Hermitian transpose has analogous properties as those of the direct transpose for real-number matrices.

Thus, with equation (2.16), we can write the AVM sensitivity expression in its complex form [6]:

$$\nabla_p f = \nabla_p^e f + \Re\left\{\hat{\boldsymbol{x}}^H\left[\nabla_p \boldsymbol{b} - \nabla_p(\boldsymbol{A}\bar{\boldsymbol{x}})\right]\right\}. \tag{2.18}$$

Usually, the response function is also complex, i.e.

$$f = f_R + jf_I. \tag{2.19}$$

Here, f_R and f_I represent the real and imaginary parts of f. In most of the cases, the response function is analytic, i.e., it satisfies the Cauchy-Riemann equations [12]:

$$\begin{aligned} \nabla_{\Re x} f_R &= \nabla_{\Im x} f_I = \Re \nabla_x f \\ -\nabla_{\Im x} f_R &= \nabla_{\Re x} f_I = \Im \nabla_x f. \end{aligned} \tag{2.20}$$

In these cases, the AVM sensitivity formula becomes [6], [7]:

$$\nabla_p f = \nabla_p^e f + \hat{\boldsymbol{x}}^H\left[\nabla_p \boldsymbol{b} - \nabla_p(\boldsymbol{A}\bar{\boldsymbol{x}})\right]. \tag{2.21}$$

Still, as in the real-number cases, we only require one additional adjoint system analysis to compute the full sensitivity information.

In some rare cases, the response functions are complex but not analytic [6]. In this situation [11], we must perform two separate AVM sensitivity analyses for the real and the imaginary parts of the response function:

$$\nabla_p f_R = \nabla_p^e f_R + \Re\left\{\hat{\boldsymbol{x}}_R{}^H\left[\nabla_p \boldsymbol{b} - \nabla_p(\boldsymbol{A}\bar{\boldsymbol{x}})\right]\right\} \tag{2.22}$$

$$\nabla_p f_I = \nabla_p^e f_I + \Re\left\{\hat{\boldsymbol{x}}_I{}^H\left[\nabla_p \boldsymbol{b} - \nabla_p(\boldsymbol{A}\bar{\boldsymbol{x}})\right]\right\} \tag{2.23}$$

where $\hat{\boldsymbol{x}}_R$ is the solution of the real-part adjoint problem,

$$A^H \hat{x}_R = \nabla_{\Re x} f \tag{2.24}$$

and $\hat{x}_I$ is the solution of the imaginary-part adjoint problem,

$$A^H \hat{x}_I = \nabla_{\Im x} f . \tag{2.25}$$

2.4 DIFFICULTIES IN THE AVM IMPLEMENTATION

As stated above, commercial EM solvers cannot compute sensitivities. The major difficulty that prevents the integration of the AVM in commercial full-wave EM solvers is the unavailability of the system matrix derivatives with respect to the design parameters $\partial A/\partial p_i$, $i=1,...,N$. The computation of the system matrix derivatives involves complicated manipulation of the mesh structure even when using a finite-difference approximation.

Therefore, we aim at performing the AVM sensitivity analysis outside the EM simulator. The simulator provides the system matrix and the field solutions.

2.5 COMPUTER RESOURSES AND THE ADJOINT-VARIABLE METHOD

The sensitivity analysis using response level finite-difference approximations is equivalent to N additional full-wave analyses. Each full-wave analysis consists of two stages of calculations. The first stage involves the matrix fill, and the second one involves solving the system of equations (2.1). The latter often involves of the LU decomposition of the system matrix and the forward and

backward substitutions. Each of these stages is repeated N times if forward (or backward) finite-difference approximations are used at the response level for the purpose of sensitivity analysis.

For sensitivity analysis using the AVM, only one additional full-wave analysis is performed. For this additional adjoint system analysis, a matrix fill is not required, as the system matrix of the adjoint problem is the transpose of the original one. The computational time required to transpose a matrix is negligible. Thus, the overhead related to this additional analysis is only the time and memory required to solve (2.16). This overhead may actually be minimal if the original problem (2.1) has already been solved by LU decomposition and the L and U factors are re-used in the solution of (2.16) [7].

The remaining overhead of the AVM is to compute the sensitivity with respect to each design parameters using (2.21). This also involves two parts: 1) the computation of the system matrix derivatives using (2.8), which is equivalent to a matrix fill for each perturbed system of equations, i.e. $A(\boldsymbol{p}+\Delta p_i \cdot \boldsymbol{e}_i)$, and 2) the calculation of the AVM sensitivity expression (2.21).

We compare the overhead of the sensitivity analysis using the traditional finite-difference (FD) approximation at the response level, and that of the AVM, for a problem with N design parameters in Table 2.1.

TABLE 2.1
COMPUTATIONAL RESOURCE COMPARISON BETWEEN THE FD AND THE AVM

Method	Matrix fills	System solutions	Sensitivity formula computation
FD	N	N	0
AVM	N	1	N

For problems with FEM formulation, the system matrix is often sparse due to the nature of the numerical method. Thus, the time required by a matrix fill is usually far less than the time required by a system solution, especially for an electrically large problem. Also, the time required for the calculation of the sensitivity formula (2.21) is negligible compared with the other two computations. Thus, the AVM is significantly more efficient than the traditional FD method.

REFERENCES

[1] N. K. Nikolova, *et al.*, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antenna Propagat.*, vol. 52, pp. 2147-2157, Aug. 2004.

[2] D. G. Cacuci, *Sensitivity & Uncertainty Analysis, Volume 1: Theory*. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[3] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice Hall, 1999.

[4] E. J. Haug, K. K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando: Academic Press Inc., 1986.

[5] J. W. Bandler, "Computer-aided circuit optimization," in *Modern Filter Theory and Design*, G. C. Temes and S. K. Mitra, Eds. New York: Wiley, 1973, ch. 6.

[6] S. M. Ali, N. K. Nikolova, and M. H. Bakr, "Recent advances in sensitivity analysis with frequency-domain full-wave EM solvers," *Applied Computational Electromagnetics Society Journal*, vol. 19, pp. 147-154, Nov. 2004.

[7] N. K. Nikolova, J. W. Bandler, and M. H. Bakr, "Adjoint techniques for sensitivity analysis in high-frequency structure CAD," *IEEE Trans. Microwave Theory Tech.*, vol. 52, No. 1, pp. 403-419, Jan. 2004.

[8] M. D. Greenberg, *Advanced Engineering Mathematics*. Upper Saddle River, NJ: Prentice Hall, 1998, pp. 605.

[9] J. Ureel and D. De Zutter, "Shape sensitivities of capacitances of planar conducting surfaces using the method of moments," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 198-207, Feb. 1996.

[10] J. Ureel and D. De Zutter, "A new method for obtaining the shape sensitivities of planar microstrip structures by a full-wave analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 249-260, Feb. 1996.

[11] N. K. Georgieva, S. Glavic, M. H. Bakr, and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751-2758, Dec. 2002.

[12] E. B. Saff and A. D. Snider, *Fundamentals of Complex Analysis*. Engle wood Cliffs, NJ: Prentice Hall, Inc., 1976.

CHAPTER 3

SELF-ADJOINT SENSITIVITY ANALYSIS FOR THE FINITE-ELEMENT METHOD

3.1 INTRODUCTION

The importance of the design sensitivity analysis of distributed systems stems from the need to improve their performance or to know their uncertainties. The design sensitivity comprises the response derivatives with respect to shape or material parameters. Manufacturing and yield tolerances, design of experiments and models, design optimization, etc., are aspects of the overall design, which can greatly benefit from the availability of the response sensitivity.

The adjoint-variable method (AVM) is known to be the most efficient approach to design sensitivity analysis for problems of high complexity where the number of state variables is much greater than the number of the required response derivatives [1]–[3]. General adjoint-based methodologies have been available for some time in control theory [1], and techniques complementary to

the finite-element method (FEM) have been developed in structural [2][3] and electrical [4]-[9] engineering. Yet, feasible implementations remain a challenge. The reason lies mainly in the complexity of these techniques.

A simpler and more versatile approach has been adopted [10]-[12] for analyses in the frequency domain, which is described in Chapter 2. The effort to formulate analytically the system matrix derivative — an essential component of the sensitivity formula — was abandoned as impractical for a general purpose sensitivity solver. Instead, approximations of the system-matrix derivatives are employed using finite differences [10]. Neither the accuracy nor the computational speed are sacrificed.

All of the above approaches require the analysis of an adjoint problem whose excitation is response dependent. Not only does this mean one additional full-wave simulation but it also requires modification of the electromagnetic (EM) analysis engine due to the specifics of the adjoint-problem excitation. Notably, Akel *et al.* [6] have pointed out that in the case of the FEM with tetrahedral edge elements, the sensitivity of the *S*-matrix can be derived without an adjoint simulation.

In this chapter, we formulate a general self-adjoint approach to the sensitivity analysis of network parameters. It requires neither an adjoint problem nor analytical system matrix derivatives. We focus on the linear problem in the frequency domain, which is at the core of a number of commercial high-frequency

simulators. Thus, for the first time, we suggest practical and fast sensitivity solutions realized entirely outside the framework of the EM solver. These standalone algorithms can be incorporated in an automated design to perform optimization, modeling, or tolerance analysis of high-frequency structures with any commercial solver, which exports the system matrix and the solution vector.

In Section 3.2, we briefly introduce the formulation of the finite-element method. Then in Section 3.3, we introduce the definition of a self-adjoint problem and show the self-adjoint formulas for network-parameter sensitivity calculations, particularly for the *S*-parameter sensitivities, based on the finite-element method. Section 3.4 outlines the features of the commercial FEM solvers, which enable independent network-parameter sensitivity analysis, and gives a general procedure for the implementation of the SASA with commercial software. Numerical validation and comparisons are presented in Section 3.5. Section 3.6 discusses the relative error and the the computational overhead associated with the sensitivity analysis.

3.2 FINITE-ELEMENT METHOD FOR EM PROBLEMS

The finite-element method (FEM) is a numerical technique for obtaining approximate solutions to boundary-value problems of mathematical physics. The FEM was first introduced in the 1950's, mainly in the structural design area. Through more than 50 years of development, FEM has been widely applied in all kinds of areas such as mechanical engineering, structural engineering, thermal

dynamic engineering, as well as electrical engineering. The FEM has been recognized as a general method widely applicable to engineering and physics problems.

The FEM was first used to solve electromagnetic problems in 1965, by O. C. Zienkiewicz and Y. K. Cheung [13], by discretizing 2-D problems and using Poisson's equation. Since then, FEM has been developed as a major numerical method in computational electromagnetics and has been applied to a wide variety of problems for different frequency bands.

The FEM implementation mainly consists of four stages [14]: 1) domain descretization, 2) selection of interpolation functions, 3) system equation formulation, and 4) system equation solution.

In the first stage, the whole computational domain, denoted as Ω, is subdivided into a number of small domains, or the "elements", as Ω^e, e=1,2,...,K, where K is the number of the elements. A typical element can be a line segment in one-dimensional domains, a triangle or a rectangle in two-dimensinal domains, and a tetrahedron, a triangular prism or a rectangular brick in three-dimensional domains, as Figure 3.1 shows [14].

In scalar finite element method, the problem is formulated in terms of the unknown function ϕ computed at the nodes of each element, while in vector finite element method it is computed at the edge of each element. In electromagnetics, ϕ is the field solution. The complete description of an element node in the FEM

includes the coordinates of the node, its local position in the element and its global position in the entire system (referred to as "local number" and "global number").

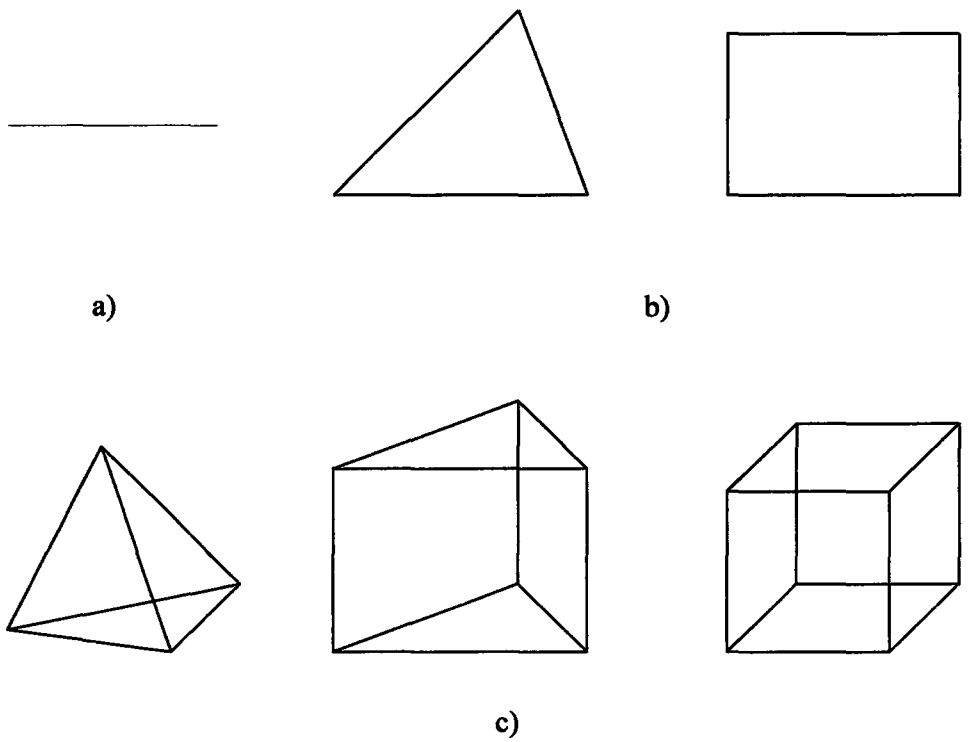

Figure 3.1 Basic elements in the FEM mesh: a) one-dimensional, b) two-dimensional, c) three-dimensional.

The second stage is to select an interpolation function, which provides an approximation of the solution at coordinates other than the element nodes, or element edges in vector finite element method. The interpolation function can be a polynomial, which is linear, quadratic, or higher-order. Thus, the solution anywhere in an element can be expressed as:

$$\tilde{\phi}^e = \sum_{j=1}^{n} N_j^e \phi_j^e . \tag{3.1}$$

Here n is the number of the nodes in one element, ϕ_j^e is the value of ϕ at the jth node, and N_j^e is the interporlation coefficient for the jth node. The latter equals 1 at the jth node and 0 at all other nodes of the element.

After proper discretization and interpolation, we can formulate the system of equations using the Ritz method or the Galerkin method, both of which solve the boundary value problem

$$\mathcal{L}\phi = f . \tag{3.2}$$

Here, $\mathcal{L}$ is the integro-differential operator, and f is the function determined by the excitation or the boundary conditions.

The Ritz method or the Galerkin method cast (3.2) in the general matrix form:

$$\boldsymbol{K}\boldsymbol{\varphi} = \boldsymbol{b} . \tag{3.3}$$

Here, $\boldsymbol{K}$ is an M by M matrix where M is the number of the total nodes, $\boldsymbol{\varphi}$ is an M by 1 vector of the solution values at the nodes, $\boldsymbol{b}$ is an M by 1 vector computed from the desired excitation f and the boundary conditions.

For consistency of notations, in this thesis, the solution vector $\boldsymbol{\varphi}$ is expressed as the field solution vector $\boldsymbol{x}$, and $\boldsymbol{K}$ is expressed as the system matrix $\boldsymbol{A}$. Then, equation (3.3) has exactly the same as equation (2.1).

3.3 SELF-ADJOINT SENSITIVITIES FOR *S*-PARAMETERS IN THE FINITE-ELEMENT METHOD

The AVM introduced in Chapter 2 needs an additional adjoint system analysis for the sensitivity computation. It is well known that this adjoint system analysis is hard to set up and carry out in commercial EM software. This limits the practical applications of the AVM.

In this section, we give a detailed derivation of the self-adjoint sensitivity analysis (SASA) for the scattering matrix, i.e., the *S*-parameters. We show that SASA needs only the original system analysis to compute the sensitivity information [15]. We also comment on how to use SASA to compute other network parameter sensitivities.

Recall the AVM sensitivity formulas discussed in (2.4). For a network parameter sensitivity, the gradients $\nabla_p \boldsymbol{b}$ and $\nabla_p^e F$ in (2.4) vanish, and the formula becomes:

$$\nabla_p F = -\hat{\boldsymbol{x}}^T \cdot \nabla_p \left(\boldsymbol{A}\bar{\boldsymbol{x}} \right). \tag{3.4}$$

Again, in (3.4), $\bar{\boldsymbol{x}}$ is fixed, and only $\boldsymbol{A}$ is differentiated.

To obtain the full scattering matrix of a K-port structure for a particular mode υ, K solutions of the system of equations (2.1) are carried out with one of the ports being excited while the rest of the ports are matched. We assume that the jth port is excited and define the S_{kj} parameters as

$$S_{kj}^{(\upsilon)} = \frac{\iint\limits_{k\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{(\upsilon)}) \cdot (\mathbf{a}_n \times \mathbf{e}_k^{(\upsilon)}) ds_k}{\iint\limits_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{(\upsilon)inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j^{(\upsilon)}) ds_j} - \delta_{kj}, \delta_{kj} = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases}. \tag{3.5}$$

Here, $\mathbf{E}_j^{(\upsilon)inc}$ is the incident field of the υ mode at the jth port, $\mathbf{E}_j^{(\upsilon)}$ is the resulting **E**-field solution, $\mathbf{a}_n$ is the unit normal to the respective port surface, and $\mathbf{e}_\xi^{(\upsilon)}$, $\xi = j,k$, are the normalized real vectors representing the υ-mode **E**-field distribution across the respective ports. The modal vectors $\mathbf{e}_\xi^{(\upsilon)}$ form an orthonormal basis:

$$\iint\limits_{\xi\text{-port}} \left(\mathbf{e}_\xi^{(\upsilon)} \cdot \mathbf{e}_\xi^{(\upsilon')} \right) ds_\xi = \delta_{\upsilon\upsilon'} \tag{3.6}$$

where $\delta_{\upsilon\upsilon'} = 1$ if the modes υ and υ' are the same, and $\delta_{\upsilon\upsilon'} = 0$ otherwise. They are obtained either analytically or numerically [14], [16]. The analytical expressions for the modes $\mathbf{e}^{(\upsilon)}$ of a rectangular waveguide can be found in [14].

We note here that for (3.5), if the *S*-parameters are computed at planes different from their respective ports, de-embedding is applied. It is in the form of an additional exponential factor; e.g., for the reflection coefficient, it is $e^{2\gamma L_d}$ where γ is the propagation constant and L_d is the distance between the port and the plane of de-embedding. This factor is parameter-independent and does not change the derivations which follow. It is omitted for the sake of simpler notations.

Our formulation (3.5) uses the approach in [14] where the output power wave in the kth port is obtained by projecting the transverse components of the transmitted/reflected **E**-field onto the transverse components of the port modal vector $\mathbf{e}_k^{(\upsilon)}$. In the denominator, the input power wave in the jth port is obtained in the same manner. For a single-mode analysis, the typical incident field setup is $\mathbf{E}_j^{(\upsilon)inc} = E_{0j}\mathbf{e}_j^{(\upsilon)}$ where E_{0j} is a user defined magnitude. Usually, $E_{0j} = 1$. We note that an alternative formulation, see, e.g., [16], uses the orthonormal set (3.6) as well as its dual (**H**-field) vector set. Both S-parameter definitions lead to the same final sensitivity result. We choose to work with (3.5).

Since we consider the S-parameter sensitivities of a single mode, for simpler notations, the superscript (υ) is omitted but implied in all formulas hereafter. Thus, with the jth port being excited, the respective right-hand side of (2.1) is denoted by $\boldsymbol{b}_j$, and its respective solution vector is $\boldsymbol{x}_j$. It represents the field solution $\mathbf{E}_j$. K such solutions $\bar{\boldsymbol{x}}_j$, $j = 1,\ldots,K$, are available from the S-parameter analysis of the structure.

In the FEM, within each surface element s at a port, the field $\mathbf{E}^s$ is approximated via the **E**-field components x_i^s, $i = 1,\ldots,n_s$, tangential to the n_s edges of the element [14]:

$$\mathbf{a}_n^s \times \mathbf{E}^s = \sum_{i=1}^{n^s} \mathbf{B}_i^s x_i^s = (\boldsymbol{x}^s)^T \cdot \{\mathbf{B}^s\} = \{\mathbf{B}^s\}^T \cdot \boldsymbol{x}^s . \tag{3.7}$$

Here, $\mathbf{a}_n^s$ is the unit normal to the surface element, $\boldsymbol{x}^s = [\, x_1^s \cdots x_{n_s}^s \,]^T$, and $\mathbf{B}_i^s$ are the vector basis functions of the element. The column $\{\mathbf{B}^s\}$ has the $\mathbf{B}_i^s$ vectors as its elements, $\{\mathbf{B}^s\} = [\mathbf{B}_1^s \cdots \mathbf{B}_{n_s}^s]^T$. Note that the vector of edge field components $\boldsymbol{x}^s$ is a subset of the solution $\boldsymbol{x}$ of (2.1).

If S_{kj} is the response whose sensitivities we need, i.e., $F = S_{kj}$, we must consider the solution of the adjoint problem:

$$A^T \hat{\boldsymbol{x}} = [\nabla_x f]^T \tag{3.8}$$

where the respective adjoint excitation becomes $\hat{\boldsymbol{b}}_{kj} = [\nabla_x S_{kj}]^T$. Instead of dealing with the global adjoint excitation vector $[\nabla_x S_{kj}]^T$ we can consider its elemental subset $[\nabla_{x^s} S_{kj}]^T$.

From (3.5), we see that S_{kj} is a linear and, therefore, analytic function of the field solution $\mathbf{E}_j$, and, therefore, of $\boldsymbol{x}_j$, as implied by the linear relation (3.7). Then the analysis in (3.8) and (3.4) applies. We first write (3.5) in terms of the field of the surface elements of the kth port:

$$S_{kj} = \frac{\sum_{sk} \iint_{sk\text{-element}} (\mathbf{a}_n^s \times \mathbf{E}_j^s) \cdot (\mathbf{a}_n^s \times \mathbf{e}_k) ds_{sk}}{\iint_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j) ds_j} - \delta_{kj} \tag{3.9}$$

where δ_{kj} has been already defined in (3.5). Making use of (3.7),

$$S_{kj} = \frac{\sum_{sk} \iint_{sk\text{-element}} (\{\mathbf{B}^{sk}\}^T \cdot \mathbf{x}^{sk}) \cdot (\mathbf{a}_n^{sk} \times \mathbf{e}_k) ds_{sk}}{\iint_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j) ds_j} - \delta_{kj}. \tag{3.10}$$

From (3.10), we find the derivatives of S_{kj} with respect to the edge field components of the s_k element at the kth port:

$$\frac{\partial S_{kj}}{\partial x_i^{sk}} = \frac{\iint_{sk\text{-element}} \mathbf{B}_i^{sk} \cdot (\mathbf{a}_n^{sk} \times \mathbf{e}_k) ds_{sk}}{\iint_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j) ds_j}, i = 1, \ldots, n_{s_k} \tag{3.11}$$

where n_{s_k} is the number of edges of the s_k element. In gradient notation, (3.11) becomes

$$[\nabla_{x^{sk}} S_{kj}]^T = \frac{\iint_{sk\text{-element}} \{\mathbf{B}^{sk}\} \cdot (\mathbf{a}_n^{sk} \times \mathbf{e}_k) ds_{sk}}{\iint_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j) ds_j}. \tag{3.12}$$

After the assembly of the FEM equations, each of the elements of $[\nabla_{x^{sk}} S_{kj}]^T$ becomes an element of the global adjoint excitation vector $\hat{\boldsymbol{b}}_{kj} = [\nabla_x S_{kj}]^T$.

We now compare the elements of the adjoint excitation (3.12) with the elements of the excitation for the s_k element of the kth port in the original FEM problem [14]:

$$\boldsymbol{b}^{sk} = \iint_{s_k\text{-element}} \{\mathbf{B}^{sk}\} \cdot (\mathbf{U}_k^{inc} \times \mathbf{a}_n^{sk}) \, ds_{s_k} \tag{3.13}$$

where

$$\mathbf{U}_k^{inc} = -2\gamma_k E_{0k} \mathbf{e}_k \tag{3.14}$$

for a single-mode incident field. Here, $\mathbf{e}_k$ is the normalized modal vector, E_{0k} is a user defined magnitude (usually set as 1), and γ_k is the modal propagation constant of the port. The comparison reveals a simple linear relationship between the original and adjoint excitation vectors, $\boldsymbol{b}_k$ and $\hat{\boldsymbol{b}}_{kj}$:

$$\hat{\boldsymbol{b}}_{kj} = \frac{1}{2\gamma_k E_0 \iint\limits_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j) ds_j} \cdot \boldsymbol{b}_k . \tag{3.15}$$

Both $\boldsymbol{b}_k$ and $\hat{\boldsymbol{b}}_{kj}$ are obtained from their respective elemental excitations, $\boldsymbol{b}^{sk}$ and $[\nabla_{x^{sk}} S_{kj}]^T$, through identical system-assembly procedures.

Next, we turn to the adjoint solutions $\hat{\boldsymbol{x}}_{kj}$ resulting from $\hat{\boldsymbol{b}}_{kj}$ ($k, j = 1, \ldots, K$). We note that the FEM system matrix $\boldsymbol{A}$ is symmetric (see, e.g., [14]),

$$\boldsymbol{A} = \boldsymbol{A}^T . \tag{3.16}$$

From (3.15) and (3.16), we conclude that all adjoint solutions $\hat{\boldsymbol{x}}_{kj}$ needed for the S-parameter sensitivities can be calculated from the K original solution vectors $\boldsymbol{x}_k$, $k = 1, \ldots, K$, by a simple multiplication with a known complex constant:

$$\hat{\boldsymbol{x}}_{kj} = \kappa_{kj} \cdot \boldsymbol{x}_k , \kappa_{kj} = \frac{1}{2\gamma_k E_0 \iint\limits_{j\text{-port}} (\mathbf{a}_n \times \mathbf{E}_j^{inc}) \cdot (\mathbf{a}_n \times \mathbf{e}_j) ds_j} . \tag{3.17}$$

They are then substituted in (3.4) where F can be any of the elements of the S-matrix.

The self-adjoint nature of the solution derived above shows that the information necessary to compute the *S*-parameter sensitivities is already contained in the full-wave solution provided by the FEM simulator. The sensitivity analysis is thus reduced to a relatively simple and entirely independent post-process, which does not require additional full-wave solutions.

As a conclusion, we state the sensitivity formula for the self-adjoint *S*-parameter problem:

$$\nabla_p S_{kj} = -\kappa_{kj}(\bar{x}_k)^T \cdot \nabla_p \left(A\bar{x}_j\right), j,k = 1,\ldots,K. \tag{3.18}$$

Here, κ_{kj} is a constant, which depends on the powers incident upon the *j*th and *k*th ports.

We also notice that the *S*-parameters relate to all other types of network parameters through known analytical formulas. Thus, the *S*-parameter sensitivities can be converted to any other type of network-parameter sensitivities using chain differentiation.

3.4 GENERAL PROCEDURES AND SOFTWARE REQUIREMENTS FOR SASA

Assume that the basic steps in the EM structure analysis have already been carried out. These include: 1) a geometrical model of the structure has been built through the graphic user interface of the simulator, 2) a mesh has been generated, 3) the system matrix A has been assembled, 4) the system equations have been

solved for all K port excitations, and the original solution vectors $\bar{x}_k$, $k=1,\ldots,K$, of the nominal structure have been found with sufficient accuracy. The self-adjoint sensitivity analysis is then carried out with the following steps:

(1) *Parameterization*: Identify design parameters p_i, $i=1,\ldots,N$.

(2) *Generation of Matrix Derivatives*: For each p_i, perturb the structure slightly (with about 1 % of the nominal p_i value) while keeping the other parameters at their nominal values. Re-generate the system matrix $A_i = A(\boldsymbol{p}+\Delta p_i \cdot \boldsymbol{u}_i)$, where $\boldsymbol{u}_i$ is a $N\times 1$ vector whose elements are all zero except the ith one, $u_i = 1$. Compute the N derivatives of the system matrix via finite differences:

$$\frac{\partial A}{\partial p_i} \approx \frac{\Delta A}{\Delta p_i} = \frac{A_i - A}{\Delta p_i}, i=1,\ldots,N. \tag{3.19}$$

Note that (3.19) is applicable only if A and A_i are of the same size, i.e., the two respective meshes contain the same number of nodes and elements. Moreover, the numbering of these nodes and elements must correspond to the same locations (within the prescribed perturbation) in the original and perturbed structures.

(3) *Sensitivity Computations*: Use (3.18) with the proper constant κ.

The above steps show that the EM simulator must have certain features, which enable the self-adjoint sensitivity analysis. First, it must be able to export the system matrix so that the user can compute the system matrix derivatives with (3.19). Second, it must allow the control over the mesh generation so that (3.19)

is physically meaningful. Third, it must export the field/current solution vector $\bar{x}$ so that we can compute the sensitivities with (3.18). The second and third features are available with practically all commercial EM simulators. The first feature deserves more attention. The system matrix is typically very large. Fortunately, in the FEM it is usually sparse and can be compressed and further stored in the computer RAM or in a disk file without excessive time delay. Only a few of the commercial simulators give access to the generated system matrices. This is the reason why our numerical experiments are carried out with FEMLAB® [17]. FEMLAB® supports all required features. In fact, we can access its system and solution matrices directly without the need to write to the disk[1].

3.5 EXAMPLES

We compute the network-parameter sensitivities with our self-adjoint formula and compare the results with those obtained by a forward finite-difference approximation applied directly at the level of the response. This second approach requires a full-wave simulation for each designable parameter. In all plots, our results are marked with SASA (for self-adjoint sensitivity analysis), while the results obtained through direct finite differencing are marked with FD.

[1] This process is done through the FEMLAB® Matlab interface and all the sensitivity computation is completed in Matlab.

3.5.1 Rectangular Waveguide Bend

We first validate our theory using a 2-D example of a rectangular waveguide bend with a 45° miter. Its H-plane view is shown in Figure 3.2. The cross-section of the rectangular waveguide is $a \times b = 3.48 \times 1.6$ cm. The structure is analyzed in its dominant TE_{10} mode. The only design parameter in the waveguide-bend example is the miter length d, i.e. $\boldsymbol{p} = [d]$. We compute the S-parameter derivatives with respect to d in a frequency band from 5.16 GHz to 7.74 GHz (15 frequency points with uniform distribution). The range of parameter values is $1 \le d \le 3.2$ cm with a step of 1 mm. Figure 3.3 shows the derivatives of the real and imaginary parts of S_{11} with respect to d at 5.16 GHz. Figure 3.4 shows the derivatives of S_{21} at the same frequency. Figure 3.5 and Figure 3.6 show the same values of S_{11} and S_{21} at 7.74 GHz. The perturbation of d used in the computation of the system matrix derivative is 1 %. The same perturbation is used in the FD computations as well. The agreement between the two sets of data plotted in Figure 3.3 to Figure 3.6 is excellent. This is true for the whole frequency band of interest.

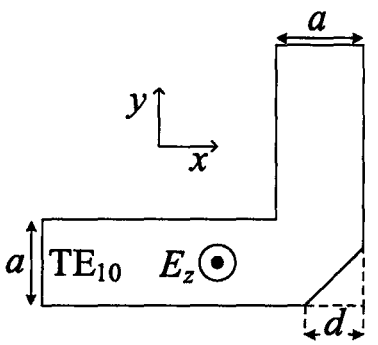

Figure 3.2 Top view of the H-plane waveguide bend structure used to validate the sensitivity analysis.

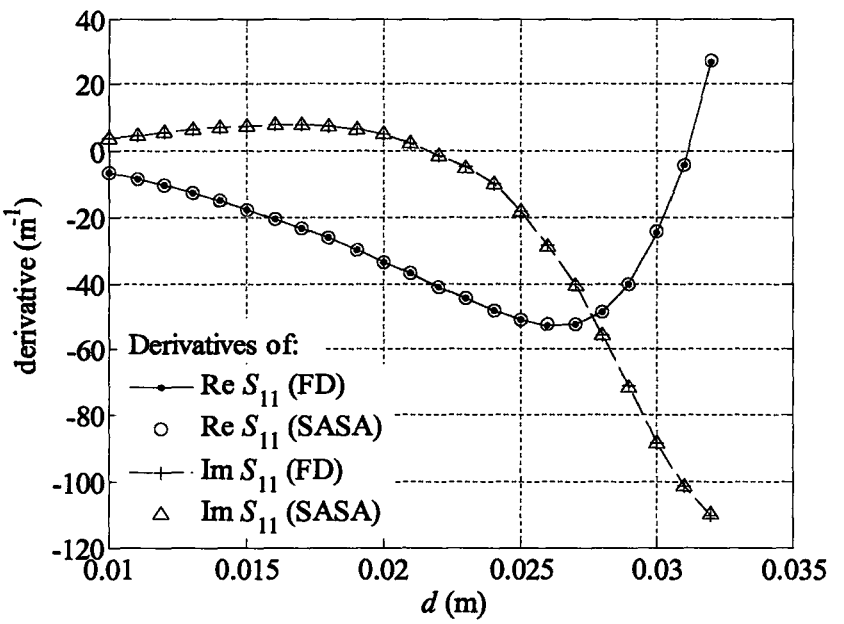

Figure 3.3 Derivatives of $\mathrm{Re}(S_{11})$ and $\mathrm{Im}(S_{11})$ with respect to d in the waveguide-bend example at $f = 5.16$ GHz.

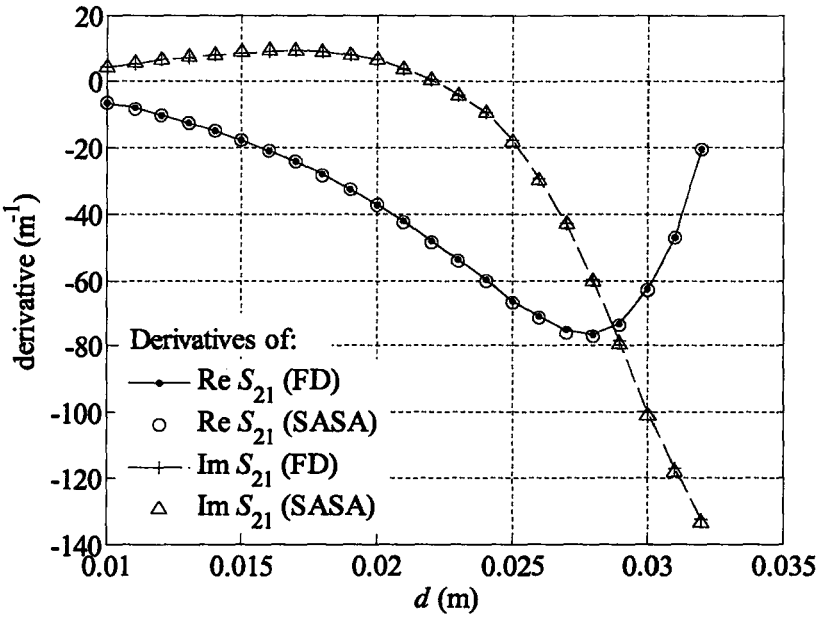

Figure 3.4 Derivatives of $\mathrm{Re}(S_{21})$ and $\mathrm{Im}(S_{21})$ with respect to d in the waveguide-bend example at $f = 5.16$ GHz.

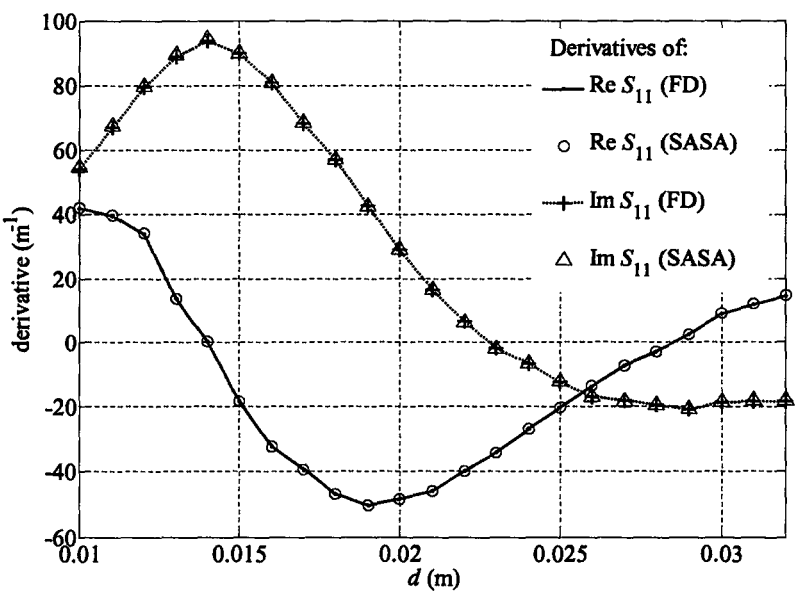

Figure 3.5 Derivatives of $\mathrm{Re}(S_{11})$ and $\mathrm{Im}(S_{11})$ with respect to d in the waveguide-bend example at $f = 7.74$ GHz.

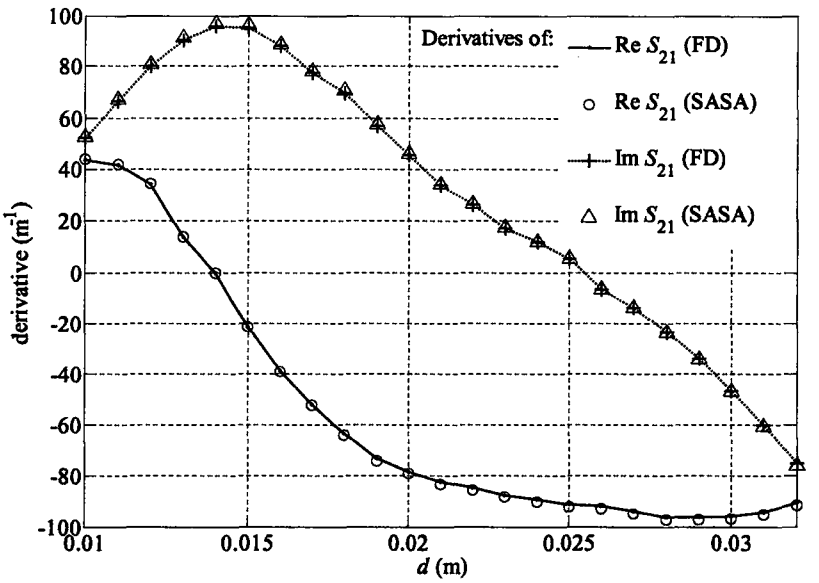

Figure 3.6 Derivatives of $\mathrm{Re}(S_{21})$ and $\mathrm{Im}(S_{21})$ with respect to d in the waveguide-bend example at $f = 7.74$ GHz.

3.5.2 Dielectric Coupling Filter

Figure 3.7 shows the top view of the dielectric filter. It is also analyzed in its dominant TE_{10} mode. The dielectric-resonator filter [18] is built from three identical rectangular ceramic posts. The material of the posts has a complex dielectric permittivity $\tilde{\varepsilon} = \varepsilon_0 \times 38.5(1 - j2\times10^{-4})$, where ε_0 is the permittivity of vacuum. The design parameters are: the width of the posts normalized to the waveguide width r/a, the distance between the posts normalized with respect to the guided wavelength s/λ_g, and the length of the posts normalized with respect to the waveguide width t/a. Figure 3.8 shows the$|S_{11}|$ and $|S_{21}|$ of a typical design with respect to the frequency. Figure 3.9 to Figure 3.11 show the derivatives of

$|S_{11}|$ and $|S_{21}|$ with respect to r/a, s/λ_g and t/a, respectively. The results are for a frequency $f = 6.88$ GHz. Again, very good agreement is observed between the self-adjoint derivatives and the respective finite-difference estimates.

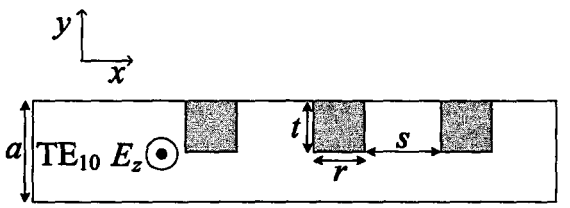

Figure 3.7 Top view of the dielectric coupling filter structure used to validate the sensitivity analysis.

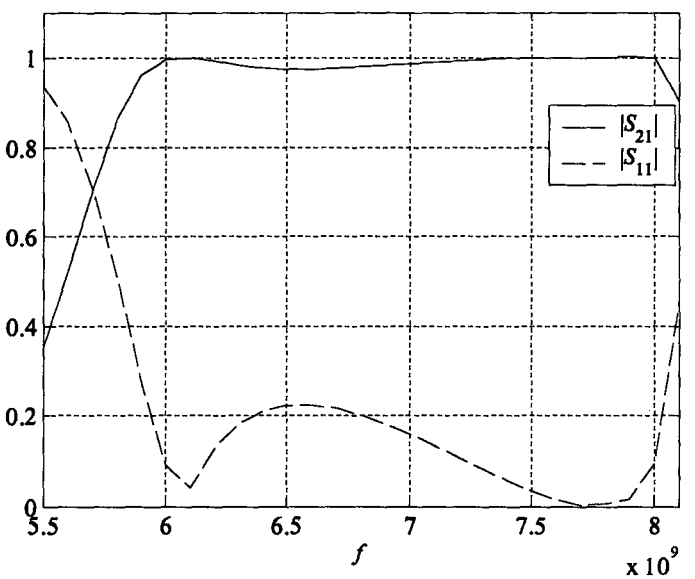

Figure 3.8 $|S_{11}|$ and $|S_{21}|$ with respect to the frequency with the design parameters of $r/a = 0.06$, $t/a = 0.2$, $s/\lambda_g = 0.32$.

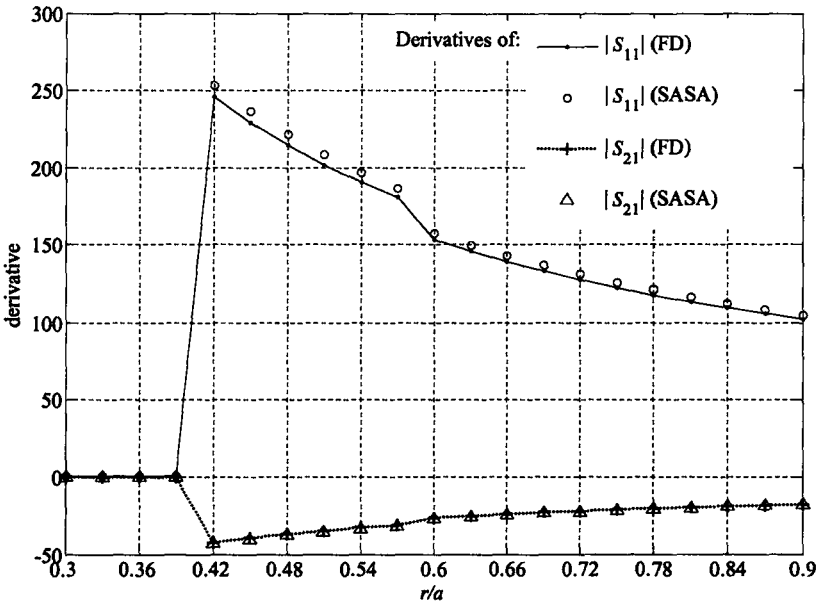

Figure 3.9 Derivatives of $|S_{11}|$ and $|S_{21}|$ at 6.88 GHz for the dielectric-resonator filter with respect to r/a. The other design parameters are fixed at $t/a = 0.2$, $s/\lambda_g = 0.32$.

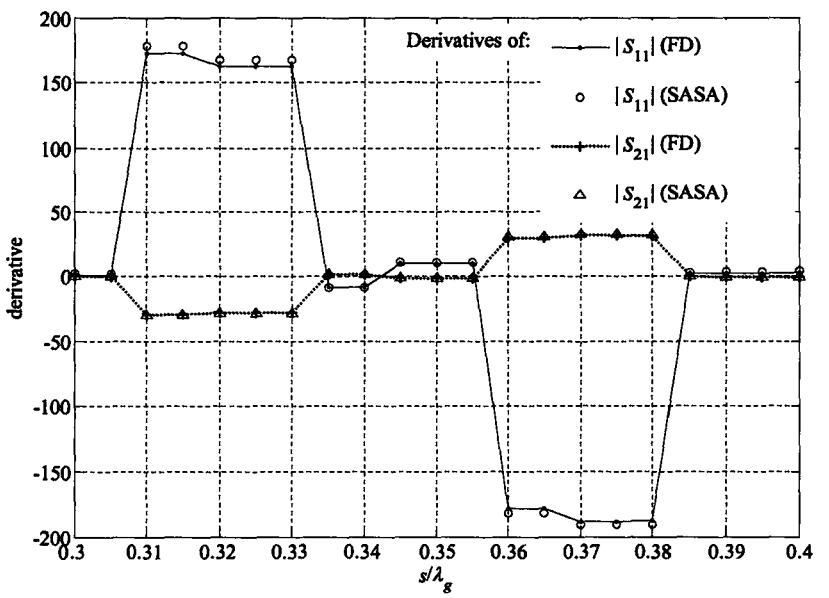

Figure 3.10 Derivatives of $|S_{11}|$ and $|S_{21}|$ at 6.88 GHz for the dielectric-resonator filter with respect to s/λ_g. The other design parameters are fixed at $r/a = 0.06$, $t/a = 0.2$.

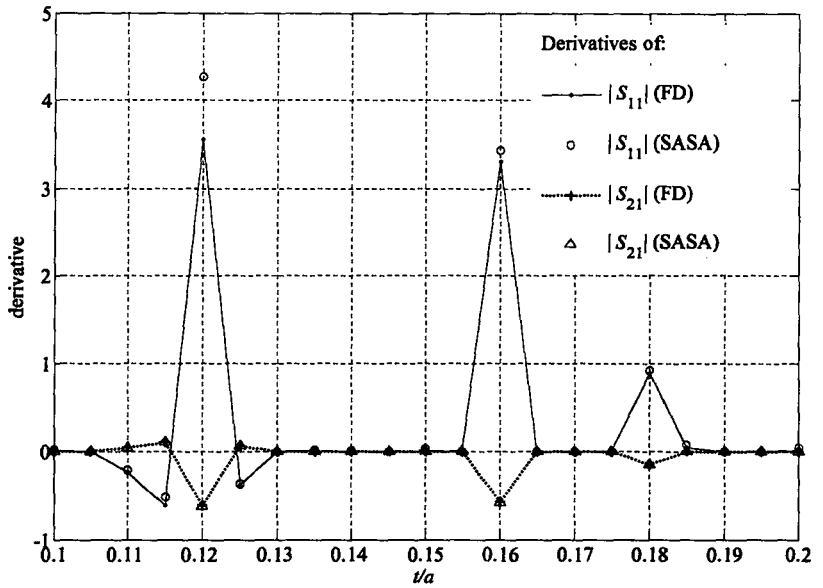

Figure 3.11 Derivatives of $|S_{11}|$ and $|S_{21}|$ at 6.88 GHz for the dielectric-resonator filter with respect to t/a. The other design parameters are fixed at $r/a = 0.06$, $s/\lambda_g = 0.32$.

3.6 ERROR ESTIMATION AND EFFICIENCY COMPARISON

3.6.1 Error Estimation

To estimate the relative error of the SASA and compare it with the error of the finite-difference approximation, first we need to establish the reference. As it is nearly impossible to compute the exact sensitivities in FEM, we choose to use the finite difference approximation with an extra fine mesh as the reference. Theoretically, the solutions of the problem converge as the mesh density approaches infinity.

The reference we use should fulfill the following requirements: 1) the convergence error of the response with respect to the mesh must be small enough; and 2) the difference between the derivatives computed by the forward finite difference, the backward finite difference, and the central finite difference must also be small. With these criteria fulfilled, we can guarantee that the finite-difference approximation, which is used as a reference, is accurate enough. Here, we choose both of the errors to be under 1 %.

To achieve this, we perform mesh refinement from a relatively coarse mesh. We refine the mesh during each iteration with a decrease of the length of the maximum mesh element edge by 50 %. The convergence error between the kth and (k+1)st iteration is computed using the following expression:

$$E^{(k)} = \left| \frac{F^{(k+1)} - F^{(k)}}{F^{(k+1)}} \right| \cdot 100\%. \tag{3.20}$$

Here F is the objective function for our sensitivity analysis. In our examples, F is S_{kj}.

We perform the error estimation with the example in Section 3.5.1. We keep refining the mesh until the convergence criterion is fulfilled with a mesh of about 36,000 elements, and using 1 % perturbation for the finite-difference approximation. The convergence error is 0.58 %, and the least square error between the derivatives using forward finite difference and backward finite difference is 0.76 %. We use this result as a reference when we compute the

relative error of the finite-difference approximation and the self-adjoint method, both of which use a coarser mesh than the mesh in the reference simulation.

Figure 3.12 to Figure 3.15 show the comparison between the reference derivatives and the derivatives computed by forward finite difference as well as by the SASA at 5.16 GHz. TABLE 3.1 shows the comparison of the average estimated errors of the FD and the SASA.

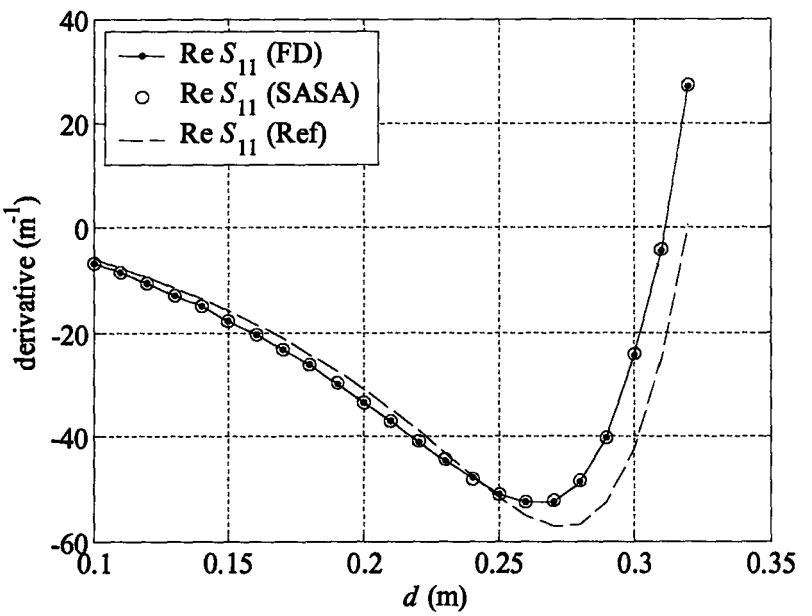

Figure 3.12 Derivatives of $\mathrm{Re}(S_{11})$ with respect to d in the waveguide-bend example at $f = 5.16$ GHz compared with the reference.

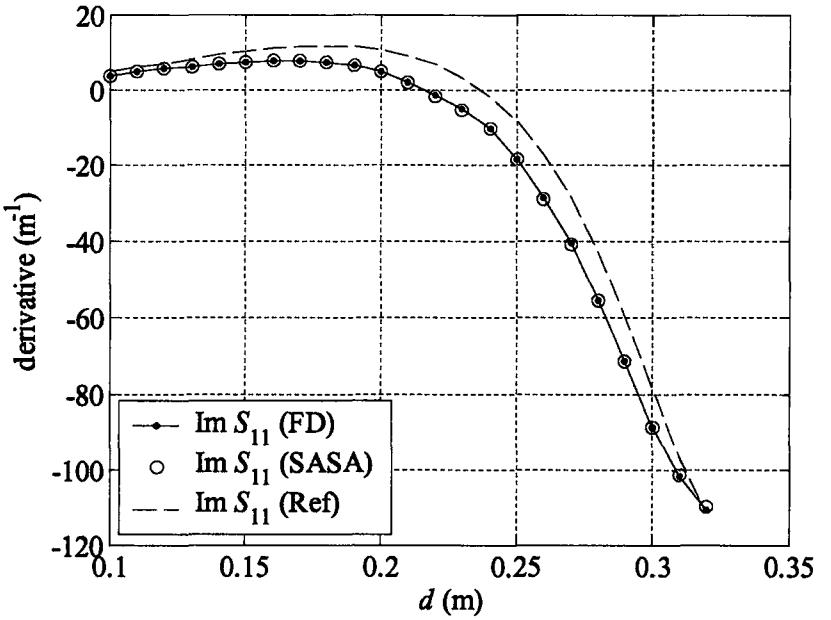

Figure 3.13 Derivatives of $\mathrm{Im}(S_{11})$ with respect to d in the waveguide-bend example at $f = 5.16$ GHz compared with the reference.

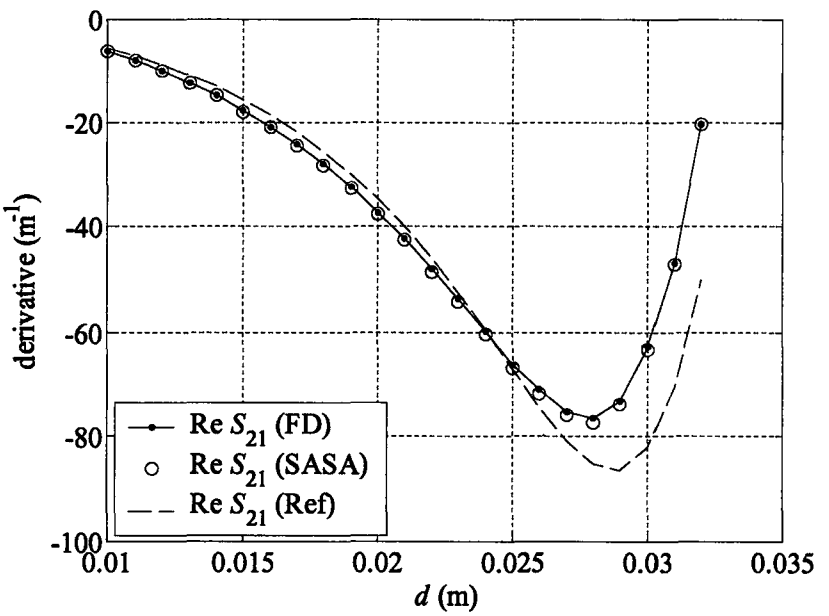

Figure 3.14 Derivatives of $\mathrm{Re}(S_{21})$ with respect to d in the waveguide-bend example at $f = 5.16$ GHz compared with the reference.

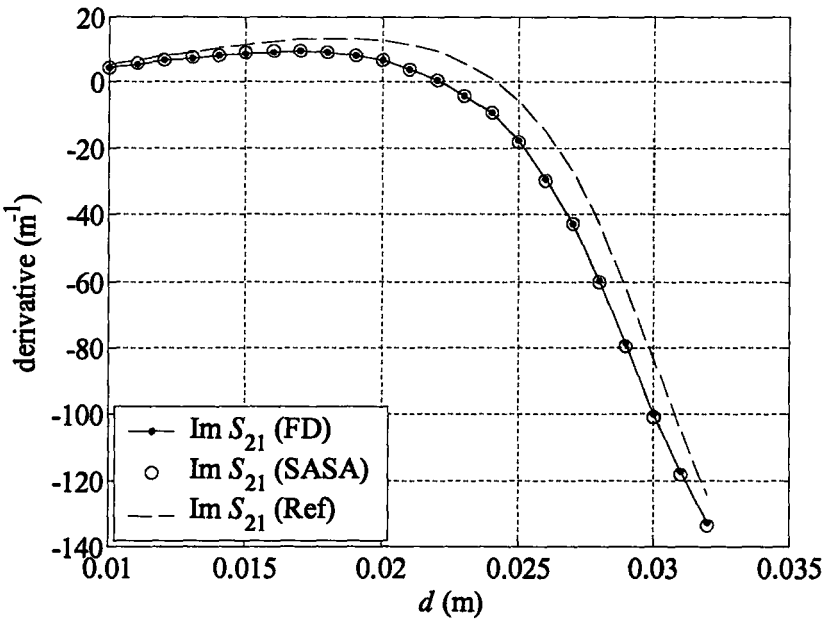

Figure 3.15 Derivatives of $\mathrm{Im}(S_{21})$ with respect to d in the waveguide-bend example at $f = 5.16$ GHz compared with the reference.

TABLE 3.1

ERROR ESTIMATION OF THE FD AND THE SASA COMPARED WITH THE REFERENCE

F	$\mathrm{Re}S_{11}$	$\mathrm{Im}S_{11}$	$\mathrm{Re}S_{21}$	$\mathrm{Im}S_{21}$
$e_{r(FD)}$	3.20%	2.80%	2.38%	1.99%
$E_{r(SASA)}$	3.31%	2.82%	2.39%	1.93%

3.6.2 Efficiency Comparison

The computational overhead associated with the self-adjoint sensitivity analysis is due to two types of calculations: 1) the system matrix derivatives, $\partial A / \partial p_i$, $i = 1,\ldots,N$, and 2) the row-matrix-column multiplications involved in the sensitivity formula (3.18). Compared to the full-wave analysis, the sensitivity

formula (3.18) requires insignificant CPU time, which is often neglected. We denote the time required to compute one derivative with the sensitivity formula as T_{SF}. In comparison, the calculation of the N system matrix derivatives is much more time consuming. Whether it employs finite differences or analytical expressions, it is roughly equivalent to N matrix fills. We denote the time for one matrix fill as T_{MF}. Thus, the overhead time required by the self-adjoint sensitivity analysis is

$$T_{SASA} = N \cdot T_{MF} + N \cdot T_{SF} . \tag{3.21}$$

On the other hand, if we employ forward finite differences directly at the level of the response in order to compute the N derivatives of the network parameters, we need N additional full analyses, each involving a matrix fill and a linear system solution. Thus, the overhead of the finite-difference sensitivity analysis is

$$T_{FD} = N \cdot T_{MF} + N \cdot T_{LS} \tag{3.22}$$

where T_{LS} is the time required to solve (2.1).

We can define a time-saving factor as the ratio $S_T = T_{FD} / T_{SASA}$, which is a measure of the CPU savings offered by our sensitivity analysis approach:

$$S_T = \frac{T_{MF} + T_{LS}}{T_{MF} + T_{SF}} . \tag{3.23}$$

Since T_{SF} is negligible in comparison with T_{MF},

$$S_T \approx 1 + \frac{T_{LS}}{T_{MF}} . \tag{3.24}$$

Evidently, the larger the ratio $R = T_{LS} / T_{MF}$, the larger our time savings. Notice that $S_T \geq 1$, i.e., our approach would never perform worse than the finite-difference approach. R depends on the size of the problem — it grows as the number of unknowns M increases.

Figure 3.16 shows the ratio $R = T_{LS} / T_{MF}$ of the FEMLAB® solver. The data are generated with a dielectric-slab waveguide filter [19] where we increase the number of unknowns from 254 to 16495 through mesh refinement. The plotted ratios are only representative since they depend on the type of the mesh and on the type of the linear-system solver (direct or iterative). The trend of the ratio increasing with the size of the problem is general. We also emphasize that we record the CPU time only. With large matrices, a computer may run out of memory (RAM), in which case, part of the data is swapped to the disk. This causes a significant increase of T_{LS} and R. This usually does not happen in the FEM solver since the system matrix of FEM is sparse.

In

TABLE 3.2, we show the actual CPU time spent for response sensitivity calculations with our self-adjoint approach and the finite-difference approximation using the FEMLAB® solver. We consider the case of one design parameter ($N = 1$), i.e., a single derivative is computed. The size of the system M varies. The increase of the time-saving factor S_T as the number of unknowns increases corresponds closely to the ratio curve $R = T_{LS} / T_{MF}$ plotted in Figure

3.16 in accordance with (3.24). The analyzed structures are the same as those used to investigate the T_{LS}/T_{MF} ratios.

We also carry out a time comparison between our approach and the finite-difference approach when the size of the system M is fixed but the number of design parameters N varies. The results are summarized in TABLE 3.3, respectively. As predicted by (3.24), the time savings are practically independent of the number of design parameters.

We re-iterate that in optimization, the Broyden update is a far more efficient alternative to the computation of the system matrix derivatives [20][21]. With it, S_T becomes roughly proportional to $(T_{MF}+T_{LS})/T_{SF}$, which is normally a very large ratio. The application of this algorithm in optimization is to be discussed in Chapter 4.

It is important to understand that in our implementation we do not have access to the meshing and matrix-assembly modules of the EM simulators. As a result, full re-meshing and a matrix fill are required to obtain the system matrix derivative for each of the N design parameters. If the self-adjoint algorithm is to be implemented within an EM solver, which already has parameterization available, the time required for a perturbed-geometry matrix fill T_{MF} may be drastically reduced [10]. This is because the parameterization necessarily flags all mesh elements affected by a parameter perturbation. It is then a relatively simple task to link the affected mesh elements to the affected matrix elements, and re-

compute only the affected elements instead of re-computing the whole system matrix. The number of affected matrix elements is very small and the system matrix derivative is very sparse when a shape parameter is perturbed. When global material parameters are perturbed, many or all of the matrix elements change and the system matrix derivative [12] is dense.

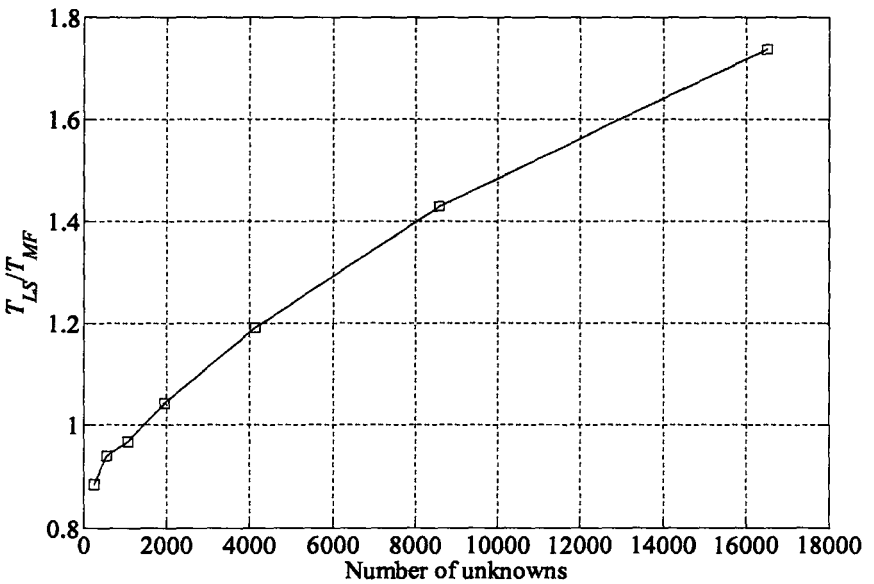

Figure 3.16 The ratio between the time required to solve the linear system and the time required to assemble the system matrix in FEMLAB®.

TABLE 3.2

FEMLAB® COMPUTATIONAL OVERHEAD OF SENSITIVITY ANALYSIS WITH THE SELF-ADJOINT METHOD AND WITH THE FINITE DIFFERENCES ($N = 1$)

M	254	555	1060	1951	4129	8557	16495
T_{SASA} (s)	17.4	17.9	18.2	19.4	21.1	25.8	32.4
T_{FD} (s)	18.2	19.3	20.9	23.7	31.1	47.3	81.8
S_T	1.04	1.08	1.15	1.21	1.47	1.88	2.52

TABLE 3.3

FEMLAB® COMPUTATIONAL OVERHEAD OF SENSITIVITY ANALYSIS WITH THE SELF-ADJOINT METHOD AND WITH THE FINITE DIFFERENCES ($M = 50000$)

N	1	2	3	4	5	6	7
T_{SASA} (s)	32.6	66.0	99.0	135.0	167.0	200.2	236.2
T_{FD} (s)	129.1	261.0	394.6	521.6	646.1	790.0	919.5
S_T	3.96	3.95	3.99	3.86	3.87	3.95	3.89

REFERENCES

[1] D. G. Cacuci, *Sensitivity & Uncertainty Analysis, Volume 1: Theory*. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[2] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice Hall, 1999.

[3] E. J. Haug, K. K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando: Academic Press Inc., 1986.

[4] P. Neittaanmäki, M. Rudnicki, and A. Savini, *Inverse Problems and Optimal Design in Electricity and Magnetism*. New York: Oxford University Press, 1996, Chapter 4.

[5] Hong-bae Lee and T. Itoh, "A systematic optimum design of waveguide-to-microstrip transition," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 803-809, May 1997.

[6] H. Akel and J. P. Webb, "Design sensitivities for scattering-matrix calculation with tetrahedral edge elements," *IEEE Trans. Magnetics*, vol. 36, pp. 1043-1046, July 2000.

[7] J. P. Webb, "Design sensitivity of frequency response in 3-D finite-element analysis of microwave devices," *IEEE Trans. Magnetics*, vol. 38, pp. 1109-1112, Mar. 2002.

[8] Y. S. Chung, C. Cheon, I. H. Park and S. Y. Hahn, "Optimal design method for microwave device using time domain method and design sensitivity analysis—part I: FETD case," *IEEE Trans. Magnetics*, vol. 37, pp. 3289-3293, Sep. 2001.

[9] N. K. Nikolova, J. W. Bandler, and M. H. Bakr, "Adjoint techniques for sensitivity analysis in high-frequency structure CAD," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 403-419, Jan. 2004.

[10] N. K. Georgieva, S. Glavic, M. H. Bakr and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751-2758, Dec. 2002.

[11] S. M. Ali, N. K. Nikolova, and M. H. Bakr, "Recent advances in sensitivity analysis with frequency-domain full-wave EM solvers," *Applied Computational Electromagnetics Society Journal*, vol. 19, pp. 147-154, Nov. 2004.

[12] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for frequency domain TLM problems with conducting boundaries," *IEEE Microwave and Wireless Components Letters*, vol. 13, pp. 408-410, Sept. 2003.

[13] O. C. Zienkiewicz and Y. K. Cheung, "Finite elements in the solution of field problems," *The Engineer*, vol. 200, pp. 507-510, 1965.

[14] J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. New York: John Wiley & Sons, 2002.

[15] N. K. Nikolova, J. Zhu, D. Li, M. H. Bakr, and J. W. Bandler, "Sensitivity analysis of network parameter with electromagnetic frequency-domain simulators," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 670-681, Feb. 2006.

[16] M. Salazar-Palma, T. K. Sarkar, L.-E. García-Castillo, T. Roy, A. Djordjević, *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modeling*. Norwood, MA: Artech, 1998, pp. 465-466.

[17] FEMLAB®3.1 User's Guide 2004, COMSOL, Inc., 8 New England Executive Park, Suite 310, Burlington, MA 01803, USA, http://www.comsol.com.

[18] L. Minakova and L. Rud, "Spectral approach to the synthesis of the waveguide bandstop filters based on dielectric rectangular posts," *Int. Conference on Mathematical Methods in Electromagnetic Theory MMET*, Sept. 2000, vol. 2, pp. 479-481.

[19] A. Abdelmonem, J. -F. Liang, H. -W. Yao, and K. A. Zaki, "Spurious free D.L. TE mode band pass filter," *IEEE MTT-S Int. Microwave Symposium Digest*, vol. 2, May 1994, pp. 735–738.

[20] N. K. Nikolova, R. Safian, E. A. Soliman, M.H. Bakr, and J.W. Bandler, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antennas Propagat.*, vol. 52, pp. 2147 – 2157, Aug. 2004.

[21] E. A. Soliman, M. H. Bakr, and N. K. Nikolova, "Accelerated gradient-based optimization of planar circuits," *IEEE Trans. Antennas Propagat.*, vol. 53, pp. 880-883, Feb. 2005.

CHAPTER 4

GRADIENT BASED OPTIMIZATION USING SENSITIVITY ANALYSIS

4.1 INTRODUCTION

The optimization algorithms used in computer-aided design can be divided into two categories: those which require only objective function values during the optimization process and those which require the objective function and its derivatives with respect to the optimizable parameters. Examples of the former are the traditional pattern search [1], genetic and particle swarm algorithms [1], [2], as well as some neural-network based algorithms [3]. Some of these, such as the genetic and particle swarm algorithms, are preferable when there is little or no information about an initial design. Their drawback is that they require a large number of system analyses. With 3-D full wave electromagnetic (EM) solvers, which require extensive simulation time, such approaches are often impractical. The second category includes gradient-based algorithms based on quasi-Newton,

sequential quadratic programming (SQP) and trust-region methods. They need the objective function Jacobian and/or Hessian in addition to the objective function itself. These optimization methods search for a local optimal point. A gradient-based algorithm is expected to converge much faster, i.e., with fewer system analyses, than an algorithm in the first category. Its drawback is that a global minimum is not guaranteed, and a failure to converge is a possibility. Naturally, the solution provided by a gradient-based local optimization algorithm depends on the quality of the initial design. For a realistic 3-D EM-based design problem with an acceptable starting point, gradient-based optimization is to be preferred.

The efficiency of a successful gradient-based optimization process depends mainly on two factors: (i) the number of iterations required to achieve convergence, and (ii) the number of simulation calls per iteration. The first factor depends largely on the nature of the algorithm, on the proper formulation of the objective or cost function, and on the accuracy of the response Jacobians and/or Hessian. The second factor depends mostly on the method used to compute the Jacobians and/or Hessian, which are necessary to determine the search direction and the step in the design parameter space. The sensitivity analysis, which provides Jacobians, is very time consuming when finite differences or higher-order approximations are used at the response level. At least N+1 full wave simulations are needed to obtain a Jacobian for N design parameters. This is unacceptable when N is large.

In Chapter 3, we proposed a self-adjoint sensitivity analysis (SASA) method for the efficient computation of network parameter sensitivities in the frequency domain [4], [5], [6]. We refer to this method as finite-difference SASA (FD-SASA) since it uses finite-difference approximation to compute the system matrix derivatives. The SASA produces the response and its Jacobian with a single full-wave analysis when the objective function depends on the network parameters, e.g., the *S*-parameters. The major overhead of the sensitivity computation with the FD-SASA is due to the computation of the system matrix derivatives via finite differences. This overhead is equivalent to N matrix fills.

In this chapter, we investigate the feasibility of the Broyden update in the computation of the system matrix derivatives [7], [8] for use with the self-adjoint formula during optimization. We refer to this approach as Broyden-update self-adjoint sensitivity analysis (B-SASA). It is applicable to sensitivity analysis for optimization purposes due to the iterative nature of Broyden's formula [9]. With it, the sensitivity analysis has practically no computational overhead as the N additional matrix fills are unnecessary. This improvement is significant compared with the FD-SASA. In comparison with the response-level finite-difference approximations of the Jacobian, the B-SASA overhead is negligible.

The B-SASA method may produce inaccurate gradient information under certain conditions. This drawback has been noticed in [7]. Here, we develop a set of criteria for switching back and forth throughout the optimization process between the robust but more time-demanding FD-SASA and the B-SASA. This

hybrid approach (B/FD-SASA) guarantees good accuracy of the gradient information with minimal computational time.

We validate and compare our method using two kinds of optimization algorithms: a minimax algorithm, which is suitable for filter and impedance-transformer design, and a least-squares algorithm, which is suitable for inverse problems. Different gradient-based search algorithms are tested such as the trust-region and the SQP. All these methods require the Jacobian, which is computed using the FD-SASA, B-SASA or B/FD-SASA methods. In the trust-region and the SQP algorithms, the Hessian is also needed but it is estimated using the classical Broyden-Fletcher-Goldfarb-Shannon (BFGS) update.

In Section 4.2, we discuss the B-SASA and the B/FD-SASA implementation in gradient-based optimization. Section 4.3 gives two numerical examples. We compare the performance of the optimization processes in Section 4.4, using the sensitivity-analysis approaches discussed above.

4.2 OPTIMIZATION WITH SELF-ADJOINT SENSITIVITIES AND BROYDEN UPDATE

We recall that in Chapter 3, the system matrix derivatives are computed by the finite difference approximation, i.e.,

$$\frac{\partial A}{\partial p_i} \approx \frac{\Delta A}{\Delta p_i} = \frac{A_i - A}{\Delta p_i}, i = 1, \ldots, N. \tag{4.1}$$

Such computation requires at least N matrix fills. To eliminate this overhead, here, we compute the system matrix derivative applying Broyden's formula. The original Broyden's formula has the form [9]:

$$\boldsymbol{G}_{k+1}=\boldsymbol{G}_k+\frac{\boldsymbol{F}(\boldsymbol{p}_k+\boldsymbol{h}_k)-\boldsymbol{F}(\boldsymbol{p}_k)-\boldsymbol{G}_k\boldsymbol{h}_k}{\boldsymbol{h}_k^T\boldsymbol{h}_k}\boldsymbol{h}_k^T. \tag{4.2}$$

Here, $\boldsymbol{G}_k$ is the approximated Jacobian $\nabla_x \boldsymbol{F}$ at $\boldsymbol{p}_k$. $\boldsymbol{G}_{k+1}$ is the approximated Jacobian at $\boldsymbol{p}_k+\boldsymbol{h}_k$. $\boldsymbol{F}$ is the vector of response functions, and $\boldsymbol{h}_k$ is the increment of the design parameter vector between two iterations, i.e., $\boldsymbol{h}_k=\boldsymbol{p}_{k+1}-\boldsymbol{p}_k$.

We re-write the formula in its matrix form, with respect to the elements of $\boldsymbol{A}$:

$$\left(\frac{\partial \boldsymbol{A}}{\partial p_n}\right)^{(k+1)}=\left(\frac{\partial \boldsymbol{A}}{\partial p_n}\right)^{(k)}+\frac{\boldsymbol{A}(\boldsymbol{p}^{(k)}+\boldsymbol{h}^{(k)})-\boldsymbol{A}(\boldsymbol{p}^{(k)})-\sum_j\left(\frac{\partial \boldsymbol{A}}{\partial p_j}\right)^{(k)}h_j^{(k)}}{\boldsymbol{h}^{(k)T}\boldsymbol{h}^{(k)}}h_n^{(k)},\ n=1,\ldots,N. \tag{4.3}$$

$\boldsymbol{A}(\boldsymbol{p}^{(k)})$ is the system matrix at the kth iteration, when the design parameter space is $\boldsymbol{p}^{(k)}$, and $\boldsymbol{h}^{(k)}$ is the increment vector in the design parameter space between the kth and (k+1)st iteration. The resulting sensitivity-analysis algorithm is referred to as B-SASA. With it, the derivatives of the system matrix in the first optimization iteration are obtained using a forward finite-difference approximation. They are updated iteratively thereafter. The iterative update requires negligible computational resources compared to a matrix fill.

The derivatives of the Broyden update are less accurate than those in the FD-SASA [7]. The inaccuracy tends to be significant when the increment of the design parameters is very small, e.g., near a local minimum, as catastrophic cancellation occurs. We propose two criteria to switch from B-SASA to FD-SASA: 1) $G(\boldsymbol{p}^{(k)})>G(\boldsymbol{p}^{(k-2)})$, or 2) $\|\boldsymbol{h}_k\|\leq d$. Here G is the objective function and d is the minimum edge length of the mesh elements. The algorithm checks the criteria during each iteration. After a switch from B-SASA to FD-SASA occurs, only one optimization iteration is performed with the FD-SASA, after which the algorithm returns to B-SASA. This B/FD-SASA method is simple and guarantees acceptable accuracy of the system matrix derivatives even for small increments in the design-parameter space.

4.3 EXAMPLES

We validate our algorithm with a finite-element method (FEM) solver FEMLAB® [10] by two numerical examples: an H-plane waveguide filter and an inverse imaging problem. In both examples, we perform gradient-based optimization using the response Jacobian provided by: (1) the proposed hybrid B/FD SASA approach, (2) the FD-SASA, and (3) the forward finite-difference approximation at the response level denoted by FD.

4.3.1 Six-Section H-Plane Waveguide Filter

The six-section H-plane filter is shown in Figure 4.1 [11]. The rectangular waveguide is of width 3.485 cm and height 1.58 cm. The cutoff frequency of the TE_{10} mode is 4.3 GHz. The 6 resonators are separated by 7 septa of finite thickness $\delta=0.625$ mm. The design parameters are the resonator lengths L_1, L_2 and L_3, and the septa widths W_1, W_2, W_3 and W_4. A minimax objective function is used with the design specifications

$$\begin{aligned} &|S_{21}|\le 0.52 \quad f\le 5.0 \text{ GHz}, \\ &|S_{21}|\ge 0.98 \quad 5.0\le f\le 9.0 \text{ GHz}, \\ &|S_{21}|\le 0.70 \quad f\ge 0.0 \text{ GHz}. \end{aligned} \tag{4.4}$$

We choose 12 uniformly distributed points in the frequency range from 4.5 GHz to 10.0 GHz. The initial design is given by $\boldsymbol{p}^{(0)T}=[L_1\ L_2\ L_3\ W_1\ W_2\ W_3\ W_4]$ = [12 14 18 14 11 11 11] (all in mm).

We use Madsen's minimax optimization algorithm [12], which employs a trust region, and the Matlab minimax algorithm, which employs SQP. We refer to these algorithms as TR-minimax and SQP-minimax, respectively. For TR-minimax, the initial trust-region radius is set to $r_0=0.03\cdot\|\boldsymbol{p}^{(0)}\|$, where $\boldsymbol{p}^{(0)}$ is the initial design parameter vector. We use the FD-SASA, the B/FD-SASA, and the response-level finite differences (FD) to supply the Jacobian in the three separate optimization processes.

Figure 4.2 shows the $|S_{21}|$ of the initial design and the optimal design using FD and B/FD-SASA method. Figure 4.3 and Figure 4.4 show the parameter step

size and the objective function versus the iterations when using TR-minimax with all three sensitivity-analysis techniques. Figure 4.5 and Figure 4.6 illustrate the same three cases with SQP-minimax. In TABLE 4.1, the optimal designs achieved by the three approaches in the case of TR-minimax are listed.

We also note that in the TR-minimax optimization process of the H-plane filter, the B/FD-SASA method switches from B-SASA to FD-SASA once at the 5th iteration due to criterion 1, while with SQP-minimax the B/FD-SASA does not switch. The optimal points of the different methods are practically the same.

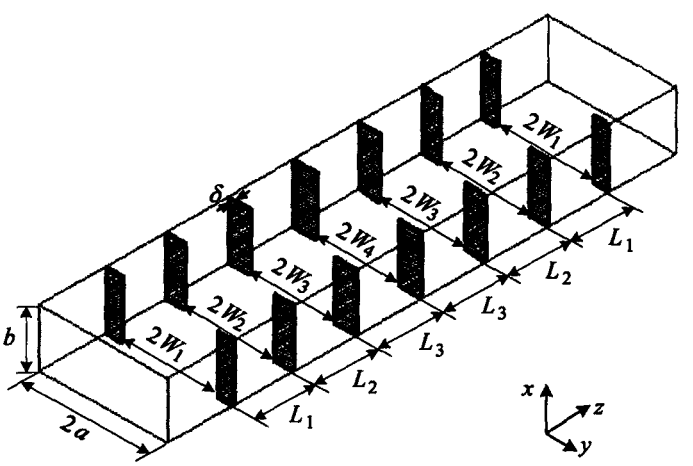

Figure 4.1 Six-section H-plane filter.

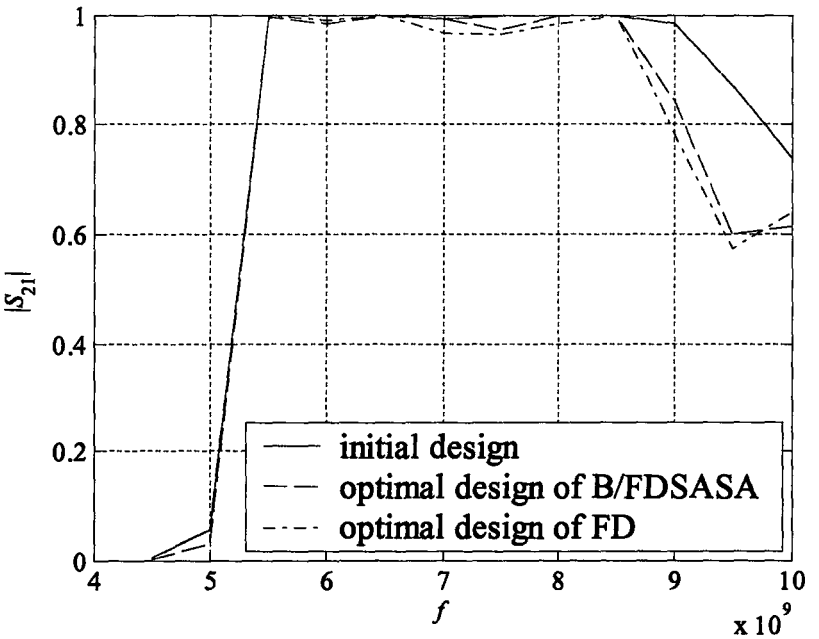

Figure 4.2 The $|S_{21}|$ with respect to frequency at initial design and optimal design.

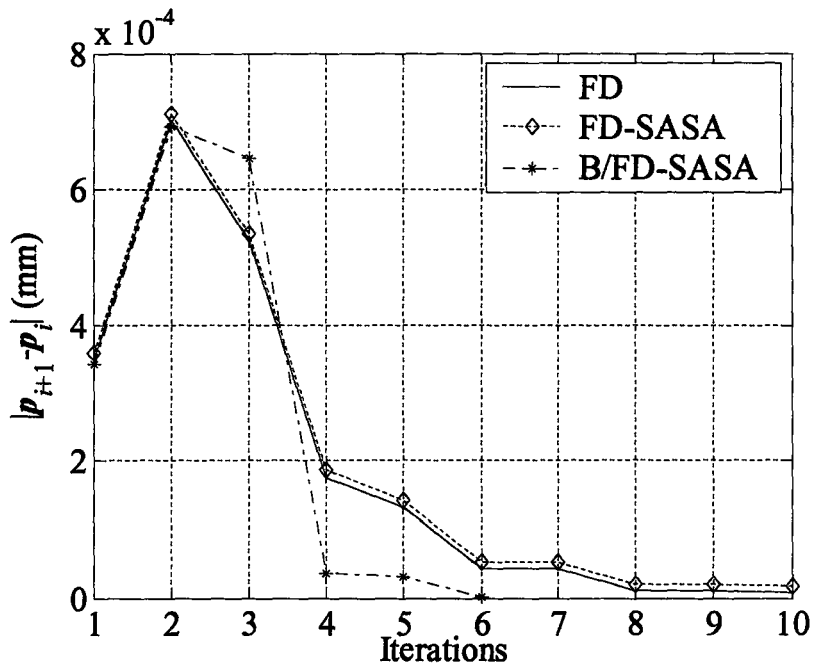

Figure 4.3 Parameter step size vs. optimization iterations in the TR-minimax optimization of the H-plane filter.

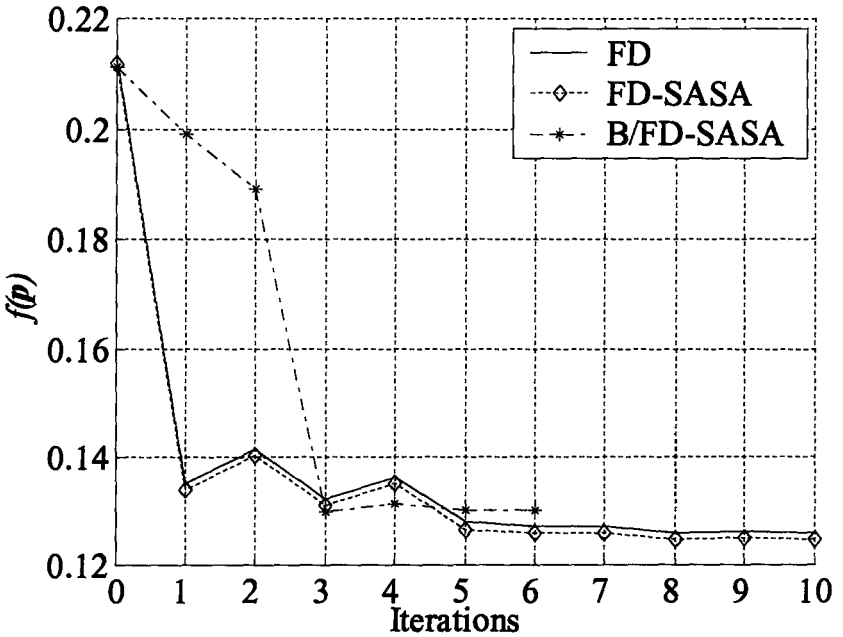

Figure 4.4 Objective function vs. optimization iterations in the TR-minimax optimization of the H-plane filter.

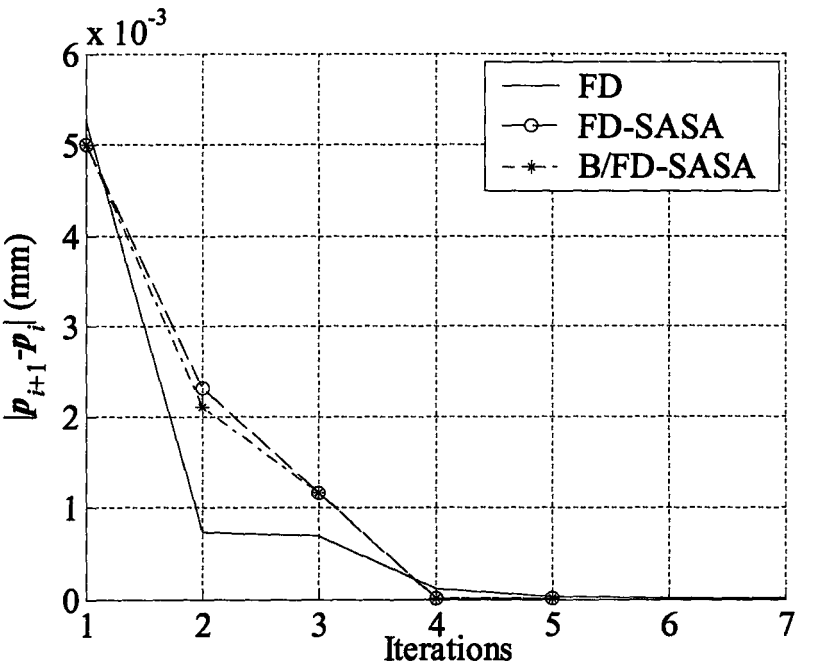

Figure 4.5 Parameter step size vs. optimization iterations in the SQP-minimax optimization of the H-plane filter.

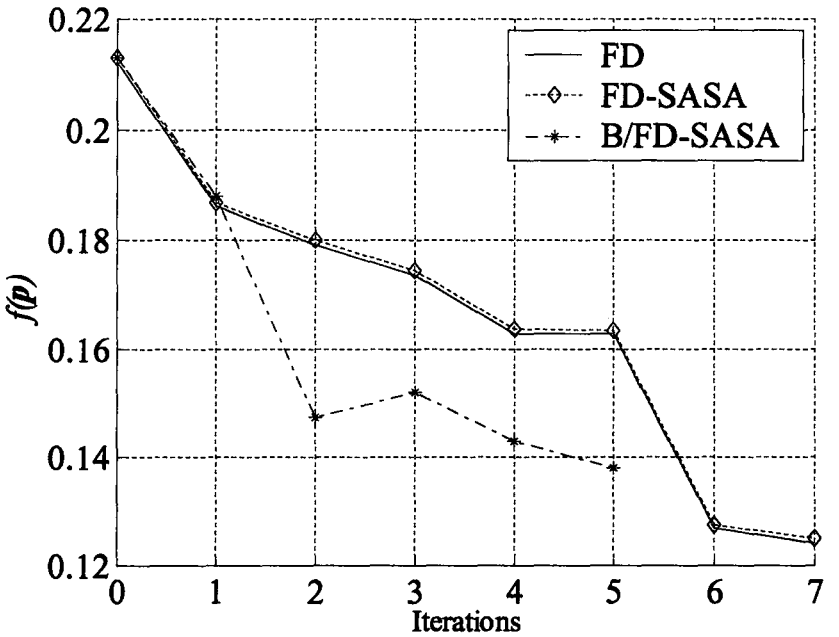

Figure 4.6 Objective function vs. optimization iterations in the SQP-minimax optimization of the H-plane filter.

TABLE 4.1

OPTIMAL DESIGN USING DIFFERENT SENSITIVITY ANALYSIS METHODS WITH TR-MINIMAX

(all in mm)	L_1	L_2	L_3	W_1	W_2	W_3	W_4
FD	12.226	14.042	17.483	14	11	10.922	11.341
FD-SASA	12.233	14.088	17.485	14	11	10.987	11.378
Mixed B/FD-SASA	12.131	13.855	17.809	14.01	11.1	11.098	11.191

4.3.2 Inverse Image problem

The 2-D inverse imaging problem is shown in Figure 4.7. The lossy inhomogeneous structure is illuminated by a TEM wave for frequencies from 5 GHz to 9 GHz, with an interval of 0.5 GHz. The objective is to determine the position and size of an object immersed in the host medium. The object is

modeled as a rectangular area with width w and length l. The distance to the interface is d. The object has a relative permittivity $\varepsilon_{r2}=45$ and specific conductivity $\sigma_2=4.5$. The host medium is characterized by $\varepsilon_{r1}=12$ and $\sigma_1=0.5$. The computational domain is surrounded by absorbing boundaries.

To obtain a target response, we perform a simulation with the target shape parameters $\bar{\boldsymbol{p}}^T=[w\ l\ d]=[44\ \ 55\ \ 5]$ (in mm). The target response is the magnitude of the reflection coefficient. It is denoted as $|\bar{S}_{11}|$.

We optimize the three shape parameters so that the simulated response matches the target response. The objective function for the optimization is the least-square error function

$$G(\boldsymbol{p})=\sqrt{\sum_{i=1}^{9}\left[|S_{11}(\boldsymbol{p},f_i)|-|\bar{S}_{11}(f_i)|\right]^2}\ . \tag{4.5}$$

Here, $|S_{11}(\boldsymbol{p},f_i)|$ is the response from the FEM forward solution at the frequency f_i, and $|\bar{S}_{11}(f_i)|$ is the target response at f_i.

We use Matlab's *lsqnonlin* optimization algorithm, which employs a trust region, with the initial trust-region radius set as $r_0=0.03\cdot\left\|\boldsymbol{p}^{(0)}\right\|$. The objective function and the parameter step size versus iterations are shown in Figure 4.8 and Figure 4.9.

We use the FD-SASA and B/FD-SASA methods to supply the Jacobian to the optimization algorithm. The initial guess of the shape parameter is $\boldsymbol{p}^{(0)T}=[w\ l\ d]=[40\ \ 40\ \ 20]$ (in mm) with both methods. Both FD-SASA and

B/FD SASA result in an optimization which takes 8 iterations to converge to an optimal point. The optimal point with the FD-SASA is $\boldsymbol{p}^{*(8)}_{\text{FD-SASA}} = [42.8\ 57.7\ 5.51]$ (in mm), while that of the B/FD-SASA is $\boldsymbol{p}^{*(8)}_{\text{B/FD-SASA}} = [44.2\ 54.1\ 5.97]$ (in mm). The B/FD-SASA switches to the FD-SASA method once at the 4th iteration due to criterion 1.

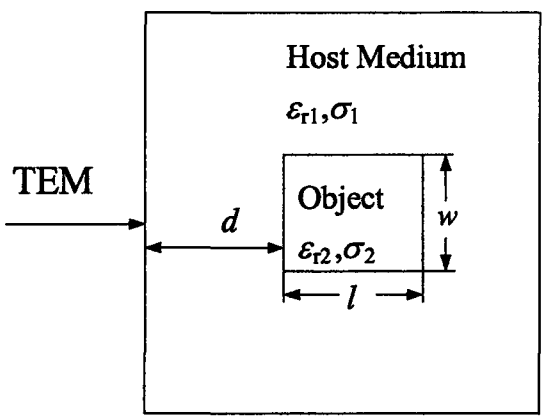

Figure 4.7 2-D inverse imaging problem.

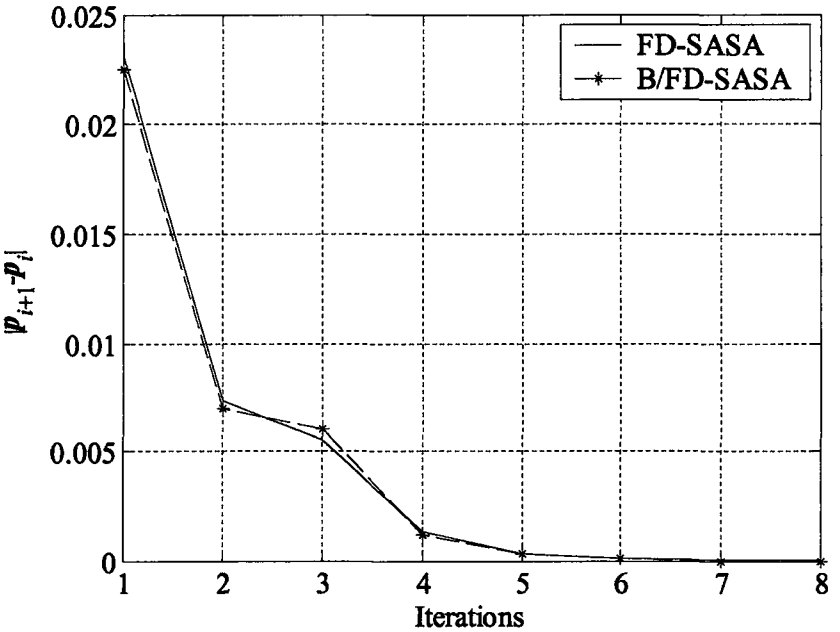

Figure 4.8 Parameter step size vs. optimization iterations using the least-squares algorithm in the 2-D inverse problem.

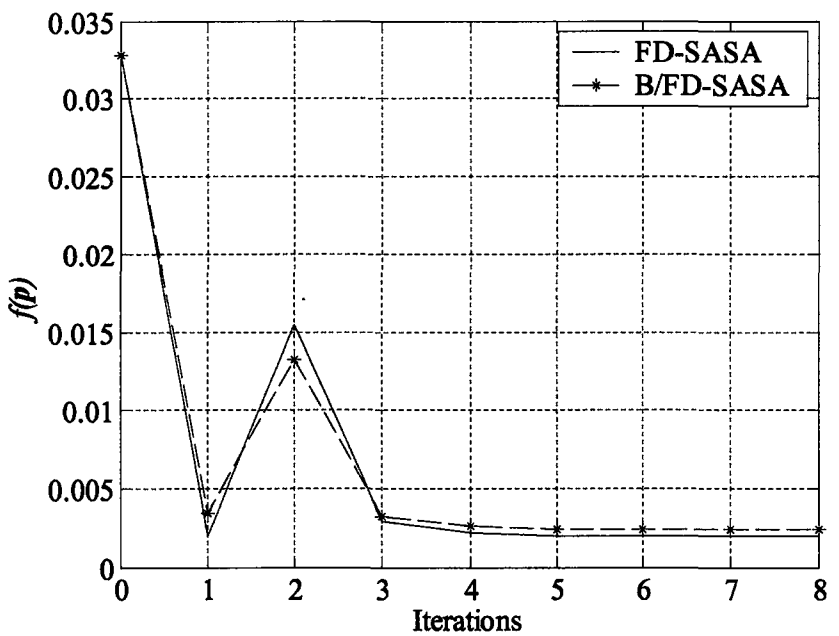

Figure 4.9 Objective function vs. optimization iterations using the least-squares algorithm in the 2-D inverse problem.

4.4 COMPARISON AND CONCLUSION

We compare the efficiency of the algorithms in terms of CPU time and iteration numbers in this section.

TABLE 4.2 shows the respective number of iterations and the overall time cost in the H-plane filter example. We notice that SQP-minimax takes longer than TR-minimax despite the smaller number of iterations. This is due to the fact that SQP-minimax may call for a system analysis several times per iteration in order to determine the next iterate. One FEM simulation (full frequency sweep) of this structure takes about 43 s. To obtain the *S*-parameters and their 7 derivatives with respect to the design parameters, 8 FEM simulations are necessary using the FD-based sensitivity analysis (approximately 344 s). As shown in TABLE 4.2, TR-minimax with FD sensitivities takes 11 iterations to converge, each iteration involving one system analysis. Thus, the simulation time accounts for almost all of the time reported in the first column of TABLE 4.2. Also, the overhead associated with the FD estimation of the 7 derivatives is about 301 s per system analysis. This overhead with the FD-SASA is reduced to 168 s, while with B/FD-SASA it is 17 s. The reduction of the overhead of the sensitivity calculation as well as the overall time of the optimization process with B/FD-SASA becomes increasingly pronounced as the size of the system matrix increases.

TABLE 4.2

NUMBER OF ITERATIONS AND TIME COMPARISON BETWEEN THE DIFFERENT OPTIMIZATION METHODS

Methods	TR-Minimax			SQP-Minimax		
	FD	FD-SASA	B/FD-SASA	FD	FD-SASA	B/FD-SASA
Iterations	11	11	9	7	7	5
Time (s)	3825.7	2402.6	949.2	13561	8523.1	5479.2

In the 2-D inverse problem example using TR-minimax, the CPU time required by the optimization using FD-SASA and B/FD-SASA is 7328 s and 4330 s, respectively. Thus, the B/FD-SASA optimization is about 1.7 times faster than the FD-SASA one. The computational gain increases as the number of optimizable parameters increases and the size of the FEM system matrix increases [4].

We conclude from these examples that the time cost reduction of the B/FD-SASA method is significant when compared with the optimization exploiting response-level sensitivities as well as the optimization exploiting our original FD-SASA approach, where finite differences are used to compute the system matrix derivatives. We also observe that often the optimization algorithms exploiting the B/FD-SASA require fewer iterations to converge to the optimal point. At the same time, the optimization results are nearly the same as those obtained by the optimization algorithms based on either the FD-SASA or the finite-difference response level approximation. The time savings depend on the

optimization algorithms, as well as the numerical size of the problem. For electrically large 3-D problems with many design parameters, the time savings may be significant.

REFERENCES

[1] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering Optimization*. Upper Saddle River, NJ: Prentice Hall, 1999, pp. 56-91, 259-278.

[2] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. 52, pp. 397-407, Feb. 2004.

[3] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: the state-of-the-art," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 420-435, Jan. 2004.

[4] N. K. Nikolova, J. Zhu, D. Li, M. H. Bakr and J. W. Bandler, "Sensitivity analysis of network parameters with electromagnetic frequency-domain simulators," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 670-681, Feb. 2006.

[5] D. Li and N. K. Nikolova, "*S*-parameter sensitivity analysis of waveguide structures with FEMLAB," *COMSOL Multiphysics Conf. 2005*, Cambridge, MA, pp. 267-271, Oct. 2005.

[6] J. Zhu, N.K. Nikolova and J.W. Bandler, "Self-adjoint sensitivity analysis of high-frequency structures with FEKO," *22nd Annual Review of Progress in Applied Computational Electromagnetics ACES 06*, Miami, FL, Mar. 2006.

[7] N. K. Nikolova, R. Safian, E. A. Soliman, M. H. Bakr and J. W. Bandler, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antennas Propagat.* vol. 52, pp. 2147-2157, Aug. 2004.

[8] J. W. Bandler, S. H. Chen, S. Daijavad and K. Madsen, "Efficient optimization with integrated gradient approximations," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 444-455, Feb. 1988.

[9] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, pp. 577-593, 1965.

[10] FEMLAB®3.1 *User's Guide* 2004, COMSOL, Inc., 8 New England Executive Park, Suite 310, Burlington, MA 01803, USA, http://www.comsol.com.

[11] G. Matthaei, L. Young and E. M. T. Jones, *Microwave Filters, Impedance–Matching Networks, and Coupling Structures*. Norwood, MA: Artech House, 1980, pp. 545-547.

[12] J. W. Bandler, W. Kellerman and K. Madsen, "A superlinearly convergent minimax algorithm for microwave circuit design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, pp. 1519-1530, Dec. 1985.

CHAPTER 5

CONCLUSIONS

In this thesis, we propose a general approach to the sensitivity analysis of the network parameters based on full-wave electromagnetic solvers. The approach is referred to as the self-adjoint sensitivity analysis (SASA). Compared with the traditional finite-difference (FD) approximation and adjoint-variable method (AVM), our method is significantly more efficient. It requires only one full wave simulation to compute the responses and the sensitivities with respect to all design parameters. The SASA computational overhead is negligible in comparison with a full-wave 3-D simulation. We also investigate the application of the SASA method in gradient based optimization.

In Chapter 2, we review the theory of the traditional AVM, both with real and complex linear systems. We compare the required computational resources of the AVM and the FD method. Also, we point out the difficulties in the implementation of the AVM with commercial electromagnetic simulators.

We introduce the SASA method in Chapter 3, starting from the basic concepts of the AVM. We validate the SASA method using several numerical examples and compare its efficiency with the FD method in terms of CPU time.

In this implementation, our SASA algorithm uses finite differences in order to compute the system matrix derivatives. We refer to this algorithm as FD-SASA. The results indicate that the FD-SASA is at least 2 to 4 times faster than the FD method, depending on the size of the problem. We also conclude from the error estimation that the FD-SASA has comparable accuracy with the FD estimations.

In Chapter 4, we investigate the application of the SASA in gradient based optimization where it provides the Jacobian to the optimizer. The Broyden update is used to accelerate the computation of the system matrix derivatives instead of the forward finite differences. We refer to this algorithm as B-SASA. To guarantee the robustness of the convergence while preserving the sensitivity computation efficiency, we propose a hybrid algorithm, which switches between the FD-SASA and the B-SASA. The method is named Broyden/finite-difference SASA (B/FD-SASA). We test the method under different optimization algorithms. The numerical results suggest significant time saving.

There are difficulties in the implementation of the SASA with commercial electromagnetic solvers. The most significant problem is the manipulation of the mesh when we compute the system matrix derivatives. As discussed in Section 4.3, the commercial solver should have certain features regarding the mesh and the system matrix, which are not always available. Further research should focus on eliminating these constraints. A feasible way to solve this problem is to compute a system matrix using a mesh generated independently of the solver.

Possible methodology is the one based on the finite differences in the frequency domain.

Also, further investigation regarding the implementation of the SASA in gradient based optimization should be conducted. The switching criteria should be tested under a wider range of optimization algorithms to verify its robustness.

COMPLETE REFERENCE LIST

Chapter 1

[1] D. G. Cacuci, *Sensitivity & Uncertainty Analysis, Volume 1: Theory*. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[2] A. D. Belegundu and T.R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice Hall, 1999.

[3] E. J. Haug, K.K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando: Academic Press Inc., 1986.

[4] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, pp.318-323, Aug. 1969.

[5] J. W. Bandler and R. E. Seviora, "Current trends in network optimization," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp.1159-1170, Dec. 1970.

[6] V. A. Monaco and P. Tiberio, "Computer-aided analysis of microwave circuits," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp.249-263, Mar. 1974.

[7] J. W. Bandler, "Computer-aided circuit optimization," in *Modern Filter Theory and Design*, G. C. Temes and S. K. Mitra, Eds. New York: Wiley, 1973, ch. 6.

[8] K. C. Gupta, R. Garg, and R. Chadha, *Computer-Aided Design of Microwave Circuits*. Norwood, MA: Artech House, 1981.

[9] J. Vlach and K. Singhal, *Computer Methods for Circuit Analysis and Design*. New York: Van Nostrand, 1983.

[10] J. W. Bandler and R. E. Seviora, "Wave sensitivities of networks," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 138-147, Feb. 1972.

[11] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for time-domain TLM with wide-band Johns matrix boundaries," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 678-685, Feb. 2004.

[12] N. K. Nikolova, H. W. Tam, and M. H. Bakr, "Sensitivity analysis with the FDTD method on structured grids," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 1207-1216, Apr. 2004.

[13] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for frequency domain TLM problems with conducting boundaries," *IEEE Microwave and Wireless Component Letters*, vol. 13, pp. 408-410, Sept. 2003.

[14] S. M. Ali, N. K. Nikolova, and M. H. Bakr, "Central adjoint variable method for sensitivity analysis with structured grid electromagnetic solvers," *IEEE Trans. Magnetics*, vol. 40, pp. 1969-1971, Jul. 2004.

[15] N. K. Georgieva, S. Glavic, M. H. Bakr and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751-2758, Dec. 2002.

[16] N. K. Nikolova, J. Zhu, D. Li, M. Bakr, and J. Bandler, "Sensitivities of network parameters with electromagnetic frequency domain simulator," *IEEE Trans. Microwave Theory Tech.* vol. 54, pp. 670-681, Feb, 2006.

[17] N. K. Nikolova, J. Zhu, D. Li, et al., "Extracting the derivatives of network parameters from frequency domain electromagnetic solutions," *General Assembly of Intl. Union of Radio Science CDROM*, Oct. 2005.

[18] D. Li and N. K. Nikolova, "*S*-parameter sensitivity analysis of waveguide structures with FEMLAB," *COMSOL Multiphysics Conf. 2005 Proceedings*, Cambridge, MA, Oct. 2005, pp. 267-271.

[19] H. Akel and J. P. Webb, "Design sensitivities for scattering-matrix calculation with tetrahedral edge elements," *IEEE Trans. Magnetics*, vol. 36, pp. 1043-1046, Jul. 2000.

[20] N. K. Nikolova, *et al.*, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antenna Propagat.*, vol. 52, pp. 2147-2157, Aug. 2004.

[21] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, pp. 577-593, 1965.

[22] D. Li, N. K. Nikolova, and M. H. Bakr, "Optimization using Broyden-update self-adjoint sensitivities," *IEEE Antenna Propagat. Symposium 2006*, accepted.

Chapter 2

[1] N. K. Nikolova, *et al.*, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antenna Propagat.*, vol. 52, pp. 2147-2157, Aug. 2004.

[2] D. G. Cacuci, *Sensitivity & Uncertainty Analysis, Volume 1: Theory*. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[3] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice Hall, 1999.

[4] E. J. Haug, K. K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando: Academic Press Inc., 1986.

[5] J. W. Bandler, "Computer-aided circuit optimization," in *Modern Filter Theory and Design*, G. C. Temes and S. K. Mitra, Eds. New York: Wiley, 1973, ch. 6.

[6] S. M. Ali, N. K. Nikolova, and M. H. Bakr, "Recent advances in sensitivity analysis with frequency-domain full-wave EM solvers," *Applied Computational Electromagnetics Society Journal*, vol. 19, pp. 147-154, Nov. 2004.

[7] N. K. Nikolova, J. W. Bandler, and M. H. Bakr, "Adjoint techniques for sensitivity analysis in high-frequency structure CAD," *IEEE Trans. Microwave Theory Tech.*, vol. 52, No. 1, pp. 403-419, Jan. 2004.

[8] M. D. Greenberg, *Advanced Engineering Mathematics*. Upper Saddle River, NJ: Prentice Hall, 1998, pp. 605.

[9] J. Ureel and D. De Zutter, "Shape sensitivities of capacitances of planar conducting surfaces using the method of moments," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 198-207, Feb. 1996.

[10] J. Ureel and D. De Zutter, "A new method for obtaining the shape sensitivities of planar microstrip structures by a full-wave analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 249-260, Feb. 1996.

[11] N. K. Georgieva, S. Glavic, M. H. Bakr, and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751-2758, Dec. 2002.

[12] E. B. Saff and A. D. Snider, *Fundamentals of Complex Analysis*. Engle wood Cliffs, NJ: Prentice Hall, Inc., 1976.

Chapter 3

[1] D. G. Cacuci, *Sensitivity & Uncertainty Analysis, Volume 1: Theory*. Boca Raton, FL: Chapman & Hall/CRC, 2003.

[2] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice Hall, 1999.

[3] E. J. Haug, K. K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando: Academic Press Inc., 1986.

[4] P. Neittaanmäki, M. Rudnicki, and A. Savini, *Inverse Problems and Optimal Design in Electricity and Magnetism*. New York: Oxford University Press, 1996, Chapter 4.

[5] Hong-bae Lee and T. Itoh, "A systematic optimum design of waveguide-to-microstrip transition," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 803-809, May 1997.

[6] H. Akel and J. P. Webb, "Design sensitivities for scattering-matrix calculation with tetrahedral edge elements," *IEEE Trans. Magnetics*, vol. 36, pp. 1043-1046, July 2000.

[7] J. P. Webb, "Design sensitivity of frequency response in 3-D finite-element analysis of microwave devices," *IEEE Trans. Magnetics*, vol. 38, pp. 1109-1112, Mar. 2002.

[8] Y. S. Chung, C. Cheon, I. H. Park and S. Y. Hahn, "Optimal design method for microwave device using time domain method and design sensitivity analysis—part I: FETD case," *IEEE Trans. Magnetics*, vol. 37, pp. 3289-3293, Sep. 2001.

[9] N. K. Nikolova, J. W. Bandler, and M. H. Bakr, "Adjoint techniques for sensitivity analysis in high-frequency structure CAD," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 403-419, Jan. 2004.

[10] N. K. Georgieva, S. Glavic, M. H. Bakr and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751-2758, Dec. 2002.

[11] S. M. Ali, N. K. Nikolova, and M. H. Bakr, "Recent advances in sensitivity analysis with frequency-domain full-wave EM solvers," *Applied Computational Electromagnetics Society Journal*, vol. 19, pp. 147-154, Nov. 2004.

[12] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for frequency domain TLM problems with conducting boundaries," *IEEE Microwave and Wireless Components Letters*, vol. 13, pp. 408-410, Sept. 2003.

[13] O. C. Zienkiewicz and Y. K. Cheung, "Finite elements in the solution of field problems," *The Engineer*, vol. 200, pp. 507-510, 1965.

[14] J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. New York: John Wiley & Sons, 2002.

[15] N. K. Nikolova, J. Zhu, D. Li, M. H. Bakr, and J. W. Bandler, "Sensitivity analysis of network parameter with electromagnetic frequency-domain simulators," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 670-681, Feb. 2006.

[16] M. Salazar-Palma, T. K. Sarkar, L.-E. García-Castillo, T. Roy, A. Djordjević, *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modeling*. Norwood, MA: Artech, 1998, pp. 465-466.

[17] FEMLAB®3.1 User's Guide 2004, COMSOL, Inc., 8 New England Executive Park, Suite 310, Burlington, MA 01803, USA, http://www.comsol.com.

[18] L. Minakova and L. Rud, "Spectral approach to the synthesis of the waveguide bandstop filters based on dielectric rectangular posts," *Int. Conference on Mathematical Methods in Electromagnetic Theory MMET*, Sept. 2000, vol. 2, pp. 479-481.

[19] A. Abdelmonem, J. -F. Liang, H. -W. Yao, and K. A. Zaki, "Spurious free D.L. TE mode band pass filter," *IEEE MTT-S Int. Microwave Symposium Digest*, vol. 2, May 1994, pp. 735–738.

[20] N. K. Nikolova, R. Safian, E. A. Soliman, M.H. Bakr, and J.W. Bandler, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antennas Propagat.*, vol. 52, pp. 2147 – 2157, Aug. 2004.

[21] E. A. Soliman, M. H. Bakr, and N. K. Nikolova, "Accelerated gradient-based optimization of planar circuits," *IEEE Trans. Antennas Propagat.*, vol. 53, pp. 880-883, Feb. 2005.

Chapter 4

[1] A. D. Belegundu and T. R. Chandrupatla, *Optimization Concepts and Applications in Engineering Optimization*. Upper Saddle River, NJ: Prentice Hall, 1999, pp. 56-91, 259-278.

[2] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Trans. Antennas Propagat.* vol. 52, pp. 397-407, Feb. 2004.

[3] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: the state-of-the-art," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 420-435, Jan. 2004.

[4] N. K. Nikolova, J. Zhu, D. Li, M. H. Bakr and J. W. Bandler, "Sensitivity analysis of network parameters with electromagnetic frequency-domain simulators," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 670-681, Feb. 2006.

[5] D. Li and N. K. Nikolova, "*S*-parameter sensitivity analysis of waveguide structures with FEMLAB," *COMSOL Multiphysics Conf. 2005*, Cambridge, MA, pp. 267-271, Oct. 2005.

[6] J. Zhu, N.K. Nikolova and J.W. Bandler, "Self-adjoint sensitivity analysis of high-frequency structures with FEKO," *22nd Annual Review of Progress in Applied Computational Electromagnetics ACES 06*, Miami, FL, Mar. 2006.

[7] N. K. Nikolova, R. Safian, E. A. Soliman, M. H. Bakr and J. W. Bandler, "Accelerated gradient based optimization using adjoint sensitivities," *IEEE Trans. Antennas Propagat.* vol. 52, pp. 2147-2157, Aug. 2004.

[8] J. W. Bandler, S. H. Chen, S. Daijavad and K. Madsen, "Efficient optimization with integrated gradient approximations," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 444-455, Feb. 1988.

[9] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, pp. 577-593, 1965.

[10] FEMLAB®3.1 *User's Guide* 2004, COMSOL, Inc., 8 New England Executive Park, Suite 310, Burlington, MA 01803, USA, http://www.comsol.com.

[11] G. Matthaei, L. Young and E. M. T. Jones, *Microwave Filters, Impedance–Matching Networks, and Coupling Structures*. Norwood, MA: Artech House, 1980, pp. 545-547.

[12] J. W. Bandler, W. Kellerman and K. Madsen, "A superlinearly convergent minimax algorithm for microwave circuit design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-33, pp. 1519-1530, Dec. 1985.