

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

**ACCURACY CONTROL IN
THE OPTIMIZATION OF MICROWAVE DEVICES BY
FINITE ELEMENT METHODS**

by

Milorad Minya Gavrilovic, B. Eng.

Department of Electrical and Computer Engineering
McGill University, Montreal

A thesis submitted to the Faculty of Graduate Studies and
Research in partial fulfillment of the requirements
of the degree of Doctor of Philosophy

June 2000

Copyright © 2000 by Milorad Minya Gavrilovic



**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-69880-7

Canada

Abstract

Computational optimization methods have long been used to aid the design of microwave devices. Cost functions are commonly defined to be functions of S -parameters that are often unavailable analytically. With the ever-increasing speed of computer processors, it is becoming standard practice for designers to use field-based computer-aided design tools as “black boxes” to calculate as accurately as possible the cost functions for such devices. However, demanding high accuracy from every cost function evaluation (CFE) of the optimization can lead to excessive computational costs. These costs can be reduced by two means: use a black box that calculates the gradient of the cost function very cheaply and vary the accuracy of CFEs throughout the optimization.

This thesis presents a system for automatically controlling the accuracy throughout an optimization. A gradient-based, constrained optimizer is combined with a 2-D finite element p -adaptive scheme that calculates the gradient at a low cost. The accuracy of a CFE is controlled through a link from the optimization to adaption. The accuracy link is based on the last computed gradient of the cost function and serves as an error tolerance used to terminate the adaption. A new error estimator is developed to assess accurately the error in the gradient, used for the termination.

Three H -plane, rectangular waveguide test cases are used for validation: a waveguide T-junction with inductive post; a miter bend with dielectric column; a two-cavity, iris coupled filter. Numerical tests show that the system is very effective in reducing the computational costs of an optimization without sacrificing the accuracy of the final answer. Compared to an optimization system with no link to the adaption, there is a speed-up in computation by at least an order of magnitude for each problem.

Sommaire

Les méthodes d'optimisation computationnelles ont longtemps été utilisées pour faciliter la conception des dispositifs à micro-ondes. Des fonctions de coût sont généralement définies pour être des fonctions des paramètres de diffusion qui sont souvent peu disponibles sur le plan analytique. Avec la vitesse toujours croissante des processeurs informatiques, il est devenu fréquent pour les concepteurs de recourir à des outils de conception assistée par ordinateur, fondés sur les champs électromagnétiques, comme “boîtes noires” pour calculer aussi exactement que possible les fonctions de coût pour de tels dispositifs. Cependant, la recherche d'un haut degré de précision dans chaque évaluation de fonction de coût (EFC) de l'optimisation peut mener à des coûts de calcul excessifs. Ces coûts peuvent être réduits par deux moyens: (1) utiliser une “boîte noire” permettant de calculer le gradient de la fonction de coût à très bon marché ou (2) changer l'exactitude des EFC à travers le processus d'optimisation.

Cette thèse présente un système permettant de contrôler automatiquement l'exactitude à travers le processus d'optimisation. Un optimiseur avec contraintes, basé sur les méthodes gradientes, est combiné avec une méthode p -adaptif du 2-D éléments finis qui calcule le gradient à un bas coût. L'exactitude d'une EFC est contrôlée par un lien de l'optimisation à l'adaption. Le lien d'exactitude est basé sur le dernier gradient calculé de la fonction de coût servant de tolérance à l'erreur utilisée pour terminer l'adaption. Un nouvel estimateur d'erreurs, utilisé pour la terminaison, est développé afin d'évaluer l'erreur dans le gradient avec exactitude.

Trois situations de jonctions de guide d'ondes parallélépipédique sur le plan H sont utilisées pour fins de validation: un branchement en T avec poteau inductif; un coude (onglet) avec colonne diélectrique; un filtre à deux-cavités et à diaphragme couplée. Les essais numériques démontrent que le système est très efficace à réduire les coûts de calcul d'une optimisation sans sacrifier l'exactitude de la réponse finale. Comparativement à un système d'optimisation sans lien à l'adaption, il y a une accélération dans le calcul d'au moins un ordre de grandeur pour chaque problème.

Acknowledgements

I would like to express my sincere respect and gratitude to my supervisor, Professor Jon P. Webb for his guidance throughout my research. This work would not have been possible (or as enjoyable) without his keen insights and ability to remain focussed on our vision of the project. I am also grateful to Professors Steve McFee and David Lowther for their interesting discussions and helpful comments as well as the facilities they provided.

Most of all, I would like to thank my wife, Charmaine Gavrilovic, for her loving support, help and infinite patience throughout my studies. Her constant encouragement helped to lift my spirit at crucial times.

To my parents, Drs. Momcilo and Mirjana Gavrilovic, I owe a huge debt of gratitude. I thank them for having constant faith in me and for being the amazing role models that they are. I would especially like to thank my father for our frequent discussions and his useful advice. I am also grateful to my sister, Ivana, and her family for their support throughout.

Finally, the financial support of the *Natural Sciences and Engineering Research Council of Canada (NSERC)* and *les Fonds pour la Formation des Chercheurs et l'Aide à la Recherche (FCAR)* is gratefully acknowledged.

Table of Contents

| | |
|--|------------|
| Abstract | i |
| Sommaire | ii |
| Acknowledgements | iii |
| Table of Contents | iv |
| Acronyms and Conventions | vii |
| Chapter 1 – Introduction | 1 |
| Chapter 2 – FE Method for Computing the Cost Function and its Gradient | 7 |
| 2.1 – <i>Scattering Parameters</i> | 8 |
| 2.2 – <i>Computing Scattering Parameters by Electromagnetic Analysis</i> | 12 |
| 2.2.1 – <i>S-Parameter Calculation</i> | 12 |
| 2.2.2 – <i>Boundary Value Problem for $E_0^{(p)}$</i> | 13 |
| 2.2.2.1 – <i>Maxwell’s Equations</i> | 13 |
| 2.2.2.2 – <i>Interface Conditions</i> | 13 |
| 2.2.2.3 – <i>Boundary Conditions</i> | 14 |
| 2.2.2.4 – <i>Helmholtz Equation</i> | 17 |
| 2.2.3 – <i>Finite Element Solution of $E_0^{(p)}$</i> | 17 |
| 2.2.3.1 – <i>Weighted Residual Formulation</i> | 18 |
| 2.2.3.2 – <i>Discretization</i> | 20 |
| 2.2.4 – <i>H-Plane Waveguide Case</i> | 21 |
| 2.2.4.1 – <i>Weighted Residual Formulation</i> | 23 |
| 2.2.4.2 – <i>Discretization</i> | 24 |
| 2.3 – <i>Computing the Gradient of the S-parameters from Field Solutions</i> | 24 |
| 2.3.1 – <i>The Adjoint Variable Method</i> | 25 |
| 2.3.2 – <i>Scattering Parameter Sensitivities</i> | 26 |
| 2.4 – <i>The Cost Function</i> | 27 |
| 2.5 – <i>The Gradient of the Cost Function</i> | 28 |

| | |
|---|-----------|
| Chapter 3 – Hierarchal Elements and <i>P</i>-Adaption..... | 31 |
| 3.1 – <i>Hierarchal Elements</i> | 32 |
| 3.1.1 – Universal Matrix Assembly..... | 34 |
| 3.1.2 – Derivative of the Discretized <i>S</i> -Parameter | 36 |
| 3.1.3 – Orthogonal Polynomial Basis Functions..... | 38 |
| 3.1.4 – Continuity and Boundary Conditions..... | 39 |
| 3.2 – <i>Port Element</i> | 40 |
| 3.3 – <i>P-Adaption</i> | 43 |
| 3.4 – <i>Error Indication</i> | 47 |
| 3.4.1 – A General Framework for Targeted Error Indicators..... | 47 |
| 3.4.2 – A Targeted Error Indicator for Microwave Devices | 48 |
| Chapter 4 – Optimization..... | 52 |
| 4.1 – <i>The Constrained Optimization Problem</i> | 54 |
| 4.1.1 – Sequential Quadratic Programming | 55 |
| 4.1.2 – QP Sub-Problem..... | 58 |
| 4.1.3 – Line Search..... | 61 |
| 4.2 – <i>Termination Criteria</i> | 63 |
| 4.2.1 – General Termination Criteria | 63 |
| 4.2.2 – Goal Oriented Termination Criteria | 64 |
| 4.2.3 – Gradient Based Termination Criteria | 65 |
| Chapter 5 – Controlling Accuracy during Optimization | 67 |
| 5.1 – <i>Accuracy Control of the CFE</i> | 68 |
| 5.2 – <i>Accuracy Link between Optimization and FE Adaption</i> | 70 |
| 5.3 – <i>Estimating the Error in the Gradient of the Cost Function</i> | 74 |
| 5.4 – <i>Issues Arising from the Combination of Optimization and Adaption</i> | 77 |
| 5.4.1 – Implications of the Choice of α | 77 |
| 5.4.2 – Continuity of <i>C</i> and the Effects of Re-Meshing..... | 78 |
| 5.4.3 – Generalized Termination Criterion | 79 |
| 5.4.4 – Accuracy of CFEs during Line Searches | 80 |
| 5.4.5 – Smoothing out Sharp Changes in Accuracy..... | 80 |

| | |
|---|------------|
| 5.4.6 – Multiple Local Optima | 81 |
| 5.4.7 – Generating Suitable Meshes for the Error Estimator | 81 |
| Chapter 6 – Validation | 83 |
| 6.1 – <i>Choice of Measures of Computational Costs</i> | 85 |
| 6.2 – <i>Test Cases</i> | 86 |
| 6.2.1 – Length of Uniform Rectangular Waveguide | 86 |
| 6.2.2 – Waveguide T-Junction with Inductive Post | 89 |
| 6.2.3 – Miter Bend with Dielectric Column..... | 93 |
| 6.2.4 – Two-Cavity, Iris Coupled, Waveguide Filter | 97 |
| 6.3 – <i>Numerical Results</i> | 102 |
| 6.3.1 – Error Indication and Ports Elements for <i>P</i> -Adaptive Cost Function Calculation | 102 |
| 6.3.2 – Derivatives of Cost Functions using Hierarchal, <i>P</i> -Type Elements..... | 106 |
| 6.3.3 – Estimates of Error in Cost Function Derivatives..... | 114 |
| 6.3.4 – Accuracy Controlled Optimization | 121 |
| Chapter 7 – Conclusions..... | 135 |
| 7.1 – <i>Original Contributions</i> | 136 |
| 7.2 – <i>Suggestions for Further Work</i> | 137 |
| References | 139 |
| Appendix A – Derivative of Matrix [K] with respect to r_k^l | 150 |
| Appendix B – Local Optima of the Miter Bend Problem..... | 155 |

Acronyms and Conventions

Acronyms

| | | |
|--------|---|----------------------------------|
| CAD | – | Computer-Aided Design |
| CCC | – | Cumulative Computational Cost |
| CFE(s) | – | Cost Function Evaluation(s) |
| DOF(s) | – | Degree(s) of Freedom |
| EM | – | Electromagnetic |
| FE | – | Finite Element |
| LP | – | Linear Programming |
| N-D | – | N-Dimensional |
| PD | – | Positive Definite |
| QP | – | Quadratic Programming |
| SQP | – | Sequential Quadratic Programming |
| TEI | – | Targeted Error Indicator |

Conventions

$[V]$ is a matrix where V_{ij} is the ij^{th} entry

\underline{V} is a column vector where V_i is the i^{th} entry

V is a 2-D or 3-D vector

$[S^r]$ and $[S^i]$ are the real and imaginary parts of complex scattering matrix $[S]$

\underline{g}^k is a vector of geometric parameters at the k^{th} iteration of the optimization

Chapter 1

Introduction

As the turn of the millennium highlights the rapid advancement of integrated circuit technology, the demands for quick and accurate analysis and design of passive microwave and millimeter wave devices increase.

A passive microwave device is a microwave component that does not increase the power level of an electromagnetic wave passing through it. Examples of such components are filters, power dividers, directional couplers and resonators. All these components are junctions between some form of microwave guiding structures – e.g. coaxial transmission line, strip line, microstrip, waveguides of various cross-sections.

As in any engineering design, cost, time and precision are key elements in the use of computer-aided design (CAD) for passive microwave devices. The CAD of such a component involves the simulation of the device to determine performance and reduces and even eliminates the need to fabricate and test.

CAD gives the microwave designer great flexibility. Any minor design changes stemming from the need to improve device performance or simple design flaws can be easily re-worked and re-analyzed with little effort. Instead of fabrication and experimental tests on devices, designers can make many changes and afford to re-analyze the performance of a device repeatedly until they are satisfied with the results.

There are several choices to be made in the analysis of a passive microwave device. Figure 1.1 shows a tree of possible options in the analysis of a typical device. A passive microwave device can be approximated by a circuit model – consisting of a number of lumped circuit components and lengths of transmission line. The values for the lumped circuit components and the transmission line parameters are a result of electromagnetic (EM) field analysis of parts of the microwave device. For standard geometries, these values can be found in design manuals that have a wide range of published results based on EM field solutions [1].

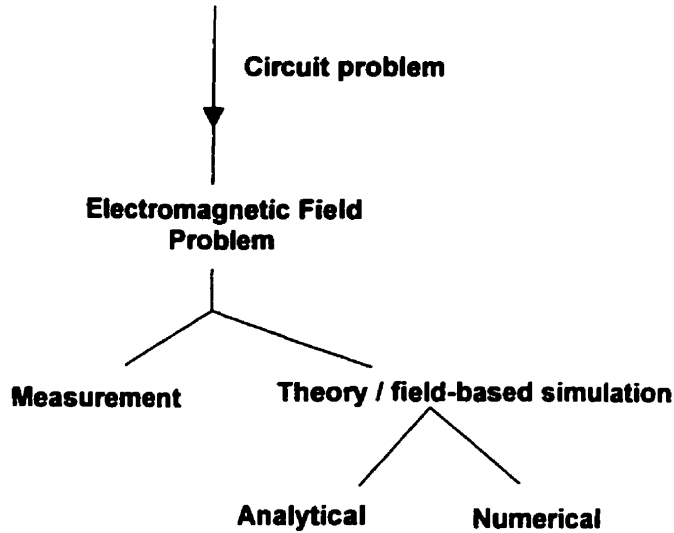


Figure 1.1: Microwave device analysis process.

The EM field problem can be solved by direct measurement of device components or theoretically, by solving Maxwell's equations [2,3]. Theoretically, the key parameters can be determined by the analytical or numerical solution. Closed form (analytic or semi-analytic) expressions for the parameters would yield the quickest and most accurate CAD tool. However, for microwave junctions of arbitrary shape there exist no closed-form expressions for the desired parameters and thus numerical CAD tools must be relied upon for accurate computation.

When the equivalent circuit components for different sections of the device are available, circuit-based simulators are extremely quick. However, there are certain inherent approximations in circuit models, such as undesired reactances and higher order mode effects [4]. In problems requiring very high accuracy, a field-based simulation of the *entire* microwave device, or at least substantial parts of it, is required – even though this may be computationally expensive (especially in three dimensions).

Some examples of microwave device field-based numerical simulation methods are finite difference time domain, finite element, transmission-line matrix (TLM), integral equation, method of moments, mode matching and spectral domain [5-8]. In this thesis, the focus is on field-based numerical analysis.

A powerful approach to the design of a passive microwave device is to first design an approximate model, then to “tune”^{*} it until the response is satisfactory. The goal is to create a device over a single or range of frequencies that yields a certain performance or response. The initial design can be obtained from formulae in a design manual, a circuit approach or any other approximation that yields a reasonable enough design [9]. This initial design is often inadequate. To improve it, a designer has two options: iteratively re-fabricate with altered geometric parameters, or include an element in the device that will allow post-production mechanical tuning. Re-fabrication is labour intensive and slows down the design process. Mechanical tuning could be effective but designers without high degrees of skill and experience may have difficulties tweaking a very frequency sensitive device – especially when dealing with multiple resonances. In addition, the inclusion of the tuning element to the device adds to the manufacturing costs and each individual device needs to be tuned after production. A more automated and refined approach is to tune the device by performing a computational optimization of the device for the given parameters.

While choices in optimizers are great, certain methods are more natural to this tuning process. The two main types of optimizers are deterministic and stochastic. Stochastic optimization deals with problems where the “parameters of the optimization problem are described by stochastic variables rather than by deterministic quantities” [10]. Typically, stochastic programming excels at finding a global optimum in a design space where the values of optimization parameters can vary greatly and which is laden with local minima (a bumpy search space). Effective optimization methods that can yield global minima have been developed, such as genetic algorithms [11] and simulated annealing [12]. These types of methods typically require 10^3 – 10^4 cost function analyses.

The initial design of a microwave device generally ensures that the starting point of the optimization is already reasonably close to the optimum. Since stochastic optimization methods require many cost function analyses (which are expensive if simulated with field-based CAD tools), it seems natural to turn towards deterministic based optimizers.

^{*} The reference to “tuning” in this thesis refers to the tuning of a device to improve design performance – not the mechanical tuning to correct manufacturing inaccuracies.

The geometric parameters of the optimization cannot vary freely and can be fabricated only to a realistic limit and tolerance of physical variation. The optimizer should be one that is not only deterministic in nature but finds a solution to the problem in a *constrained* design space. Within the spectra of deterministic optimizers, gradient-based methods tend to be the fastest.

To achieve the best design, the analysis method should be flexible so that a device of arbitrary geometry might be used. Typical optimization problems of the type considered in this thesis have fewer than 10 geometric parameters. A simple block diagram of the typical optimization (or “tuning”) of a microwave device is given in Figure 1.2 below.

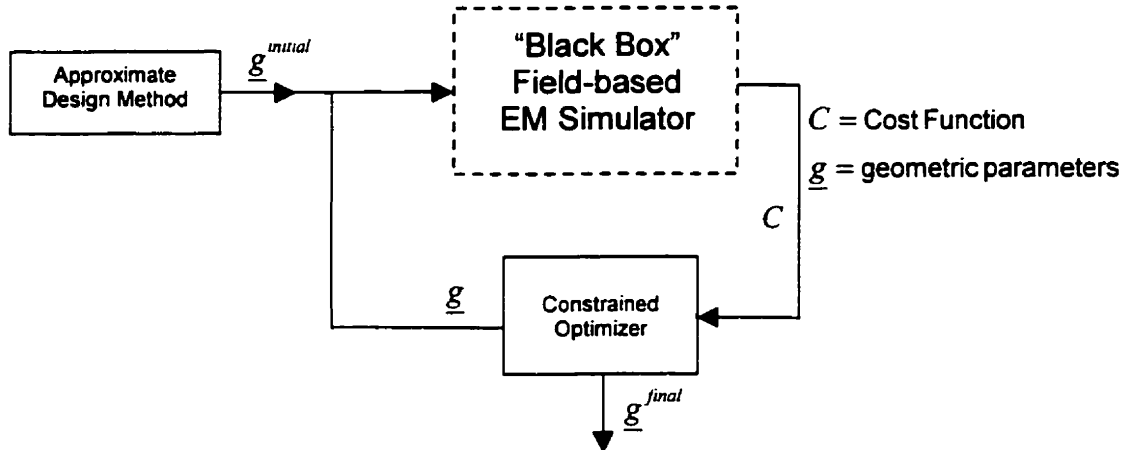


Figure 1.2: Block diagram of a typical optimization scheme.

The choice of optimizer and “black box” simulator remain quite broad as long as each tool performs its required job. In such a scheme, the ability of the EM simulator to achieve high accuracy is critical to the optimization. However, the optimization process is indifferent to the choice of simulator and is only interested in the value of the cost function (C). The optimizer should have access to sufficiently accurate values of the cost function in order to converge to the correct solution.

The thesis will address the issue of improving the efficiency and reducing the cost of the “tuning” portion of the design. Many optimization schemes that use “black box”

simulators to calculate the cost function and its gradient have to account for very long run times due to the lengthy run times of the EM simulator. A three-dimensional (3-D) field-based EM simulator can amount to hours of computation time per geometry per frequency. For example, a vector finite element simulation of a microwave component consisting of 100,000 tetrahedra can roughly have a run-time of 1 hour (using a direct solver) on a 400 MHz Pentium II processor [13]. This represents the run-time per frequency point. In reality, calculating a cost function may require a sweep over a range of frequencies – thus requiring simulations for multiple points, perhaps hundreds. The total time it takes for the frequency sweep and cost function evaluation is basically one step of the optimization. There may be hundreds of calls to such a simulator.

How can the expense of such a system be reduced? There are two major types of improvement that this thesis will deal with. The first is to develop a black box that calculates the gradient of the cost function very cheaply. This would mean calculating the partial derivatives of the cost function within the black box and not forcing the optimizer to use a finite difference approach to calculating the gradient. The second improvement is to develop a level of accuracy control for the optimization. While the easiest approach would be to ignore the level of accuracy and simply ask for the best answer the black box simulator can supply, a more sophisticated approach would be to open the black box and control the level of accuracy throughout the optimization. While at certain points of the optimization a high level of accuracy may be needed, other steps may only require a lower accuracy level for the cost function and its gradient to build a reliable design surface that is easily optimized.

The simulation method chosen for this system is the p -adaptive finite element (FE) method. P -adaption does not require re-meshing and allows accuracy control by varying the orders of different elements. The analysis for purposes of this thesis is performed in 2-D in order to reduce run times, but all theory can easily be applied to 3-D. Chapter 2 develops the finite element method for computing the cost function and more importantly the inexpensive calculation of its gradient.

Chapter 3 addresses the two main criteria that affect finite element solution accuracy – discretization and basis functions. The elements involving p -adaption (such as error indication) are addressed in this chapter.

Chapter 4 deals with the choice of optimizer and the developments of the control of the optimization system, such as the choice of termination criteria.

The link between the optimization and black box simulation is the accuracy control. Controlling accuracy during optimization with finite element p -adaption is covered in Chapter 5.

Chapter 6 validates the optimization system as well as certain of its key components on a variety of test cases. The key components are: error indication and port elements for p -adaption, derivatives of cost functions using hierarchal, p -type elements and estimates of error in cost function derivatives. Tests on the key components will precede the numerical results on accuracy controlled optimization.

Chapter 2

FE Method for Computing the Cost Function and its Gradient

In order to tune a microwave device with an optimization method, one needs the cost function analysis to be very accurate. The cost function is commonly a function of the network parameters of the device (scattering, admittance or impedance parameters). Since the devices designed in this thesis will be of arbitrary shape, a good field-based CAD tool is required to compute the network parameters (which are derived from field solutions).

Some of the best-established and most effective field-based simulators are based on the finite element method [14–17]. Finite element methods numerically solve a set of partial differential equations over a problem domain of arbitrary shape. Initially, the advantages of such methods were proven most effective for structural, fluid, and other civil/mechanical engineering applications [18]. Over the years, finite elements were pioneered to suit the needs of electromagnetic problems [19–22]. Today, finite elements are a mainstay in electrical engineering and are a well-developed electromagnetic analysis CAD tool [23–25].

An accurate cost function evaluation can be quite costly with a field-based CAD tool. Computing the partial derivatives of the cost function with respect to its optimization parameters (i.e. the gradient) would be very expensive if a traditional finite difference formula were used. The computational effort would be greatly reduced if the gradient of the cost function could be computed with no extra cost function evaluations. This chapter develops the finite element method for computing the cost function and its gradient, for a microwave device.

2.1 Scattering Parameters

Since microwave design generally involves analysis of a circuit problem, designers are seldom interested in the actual electromagnetic fields in a device. Rather, the performance of a particular component in a network is the desired quantity. A component can be viewed as a junction between N uniform waveguides (or transmission lines), commonly referred to as an N -port device. The performance of an N -port device can be characterized through its network parameters, given in the form of impedance, admittance or scattering matrices.

Consider the N -port junction in Figure 2.1. In general, the waveguides connected to each port can support multiple modes of propagation for a particular excitation frequency. For each mode k , incident and reflected waves move towards and away from the ports. However, practical devices are usually designed to allow only dominant mode propagation and will be treated as such in the formulations in this thesis. Electromagnetic simulators usually model each terminal plane, or port (P_i), far enough from the junction in order to allow the decay of evanescent modes.

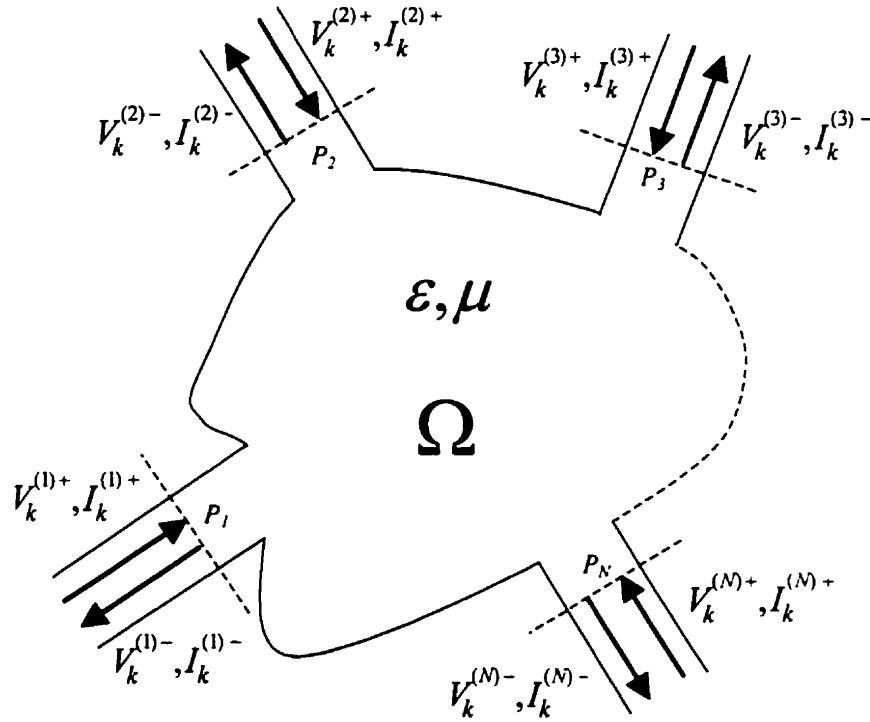


Figure 2.1: A general N -port waveguide junction.

Let $\mathbf{e}_k^{(i)}$ and $\mathbf{h}_k^{(i)}$ be the transverse electric and magnetic fields of mode k (where the dominant mode corresponds to $k = 0$) of the waveguide at port i , normalized such that

$$\int_{P_i} \mathbf{e}_k^{(i)} \times \mathbf{h}_k^{(i)} \cdot \mathbf{n} ds = -1 \quad (2.1)$$

where \mathbf{n} is a unit normal outward from the surface of the junction. Note that the normalization of $\mathbf{e}_k^{(i)}$ and $\mathbf{h}_k^{(i)}$ corresponds to unit power into the device [26]. This relation allows definition of the normalized voltage $(V_k^{(i)})$ and current $(I_k^{(i)})$ at a terminal plane. For an incident wave at port P_i , in general,

$$\mathbf{E}_t = V_k^{(i)+} \mathbf{e}_k^{(i)} \quad (2.2)$$

$$\mathbf{H}_t = I_k^{(i)+} \mathbf{h}_k^{(i)} \quad (2.3)$$

where \mathbf{E}_t and \mathbf{H}_t are the electric and magnetic tangential fields at port P_i and $V_k^{(i)+}$ and $I_k^{(i)+}$ are by definition, the voltage and current of the wave travelling towards the port. It follows that

$$\frac{V_k^{(i)+}}{I_k^{(i)+}} = 1 \quad (2.4)$$

thus making the real power flow into the port:

$$Power = \frac{1}{2} |V_k^{(i)+}|^2 \quad (2.5)$$

$V_k^{(i)+}$ is also called the incident “wave amplitude”. Similarly, for a reflected wave,

$$\mathbf{E}_t = V_k^{(i)-} \mathbf{e}_k^{(i)} \quad (2.6)$$

$$\mathbf{H}_t = I_k^{(i)-} \mathbf{h}_k^{(i)} \quad (2.7)$$

where $V_k^{(i)-}$ and $I_k^{(i)-}$ are voltage and current amplitudes travelling away from the port.

In this case

$$\frac{V_k^{(i)-}}{I_k^{(i)-}} = -1 \quad (2.8)$$

In general, when both incident and reflected waves are present, the total tangential fields at port i are thus

$$\mathbf{E}_t = V_k^{(i)} \mathbf{e}_k^{(i)} \quad (2.9)$$

$$\mathbf{H}_t = I_k^{(i)} \mathbf{h}_k^{(i)} \quad (2.10)$$

where the normalized voltage and current at this terminal are

$$V_k^{(i)} = V_k^{(i)+} + V_k^{(i)-} \quad (2.11)$$

$$I_k^{(i)} = I_k^{(i)+} + I_k^{(i)-} \quad (2.12)$$

It is now possible to characterize a microwave N -port junction by an admittance matrix $[Y]$, which is an N by N complex matrix that relates the port currents to the port voltages, such that

$$\underline{I}_k = [Y] \underline{V}_k \quad (2.13)$$

Any entry of matrix $[Y]$, Y_{ij} , can be found by applying unit voltage at port j , short-circuiting all other ports, and measuring the short circuit current at port i .

An equally effective method of characterizing a microwave device is the generalized scattering matrix, $[S]$, given by the relationship between incident and reflected wave or voltage amplitudes:

$$\underline{V}_k^- = [S]\underline{V}_k^+ \quad (2.14)$$

A scattering parameter S_{ij} can be obtained by exciting port j with an incident wave of voltage $V_k^{(j)+}$ and measuring the reflected wave amplitude at port i . All other ports of the junction are matched (all incident waves on ports other than the j^{th} port are set to zero) in order to avoid reflections. This generalized scattering matrix definition assumes that the voltage and current at all ports are normalized so that (2.4) and (2.5) hold [26].

Both admittance and scattering matrices are symmetric if the media of the microwave device are reciprocal. Furthermore, if the network is lossless, the admittance matrix becomes purely imaginary. The admittance and scattering matrices are equally good in describing the behavior at a microwave N -port junction. Furthermore, one matrix can easily be derived from the other. The scattering matrix can be found by

$$[S] = ([U] + [Y])^{-1}([U] - [Y]) \quad (2.15)$$

where $[U]$ is the N by N identity matrix. Similar calculation can yield the admittance matrix from the scattering parameters. The convenience in using scattering parameters to describe a waveguide junction is that each diagonal entry of the matrix is a reflection coefficient while the off-diagonal entries represent the transmission coefficients. Due to the normalization of voltage and current, the maximum value of the magnitude of any scattering parameter is conveniently 1 (for passive devices).

2.2 Computing Scattering Parameters by Electromagnetic Analysis

The definition of a scattering parameter (S -parameter) refers to the measurement of a reflected wave at a particular port while another port is excited with an incident wave, and all other ports are matched. Optimizing a microwave device by measuring the S -parameters for each geometry is time-consuming and expensive. In order to avoid measurement and obtain the S -parameters with field-based CAD tools, a relationship is needed between the scattering parameters and field solutions.

2.2.1 S -Parameter Calculation

Consider an N -port microwave device that is excited by its dominant mode at port p . The resulting electric field in the device is denoted $\mathbf{E}_0^{(p)}$. By mode orthogonality [27], an S -parameter can be calculated using [28]:

$$S_{rp} = \gamma_0^{(r)}(\mathbf{E}_0^{(p)}) - \delta_{rp} \quad (2.16)$$

where δ_{ij} is the Kroenecker delta and

$$\gamma_k^{(r)}(\mathbf{E}) = \int_{P_r} \mathbf{E} \times \mathbf{h}_k^{(r)} \cdot \mathbf{n} ds \quad (2.17)$$

i.e. $\gamma_k^{(r)}(\mathbf{E})$ is a linear operator extracting the k^{th} mode voltage at port r from electric field, \mathbf{E} .

To compute every entry of the scattering matrix then, the field solution $\mathbf{E}_0^{(p)}$ must be calculated for $p=1, \dots, N$.

2.2.2 Boundary Value Problem for $E_0^{(p)}$

Computing $E_0^{(p)}$ amounts to calculating the electric field intensity in an N -port microwave device where a port p is excited with its dominant mode and all other ports are matched. All methods (analytical or numerical) used to calculate $E_0^{(p)}$ stem from the theory that characterizes the classical electromagnetics problem. Note that in the subsections below, the theory describes the physical nature of general electromagnetic fields and all equations refer to the electric field as E rather than $E_0^{(p)}$ except in the section on port boundary conditions.

2.2.2.1 Maxwell's Equations

The physical laws that describe the behavior of electromagnetic fields are given by Maxwell's equations. For time harmonic fields in three dimensions with no sources,

$$\nabla \times \mathbf{E} = -j\omega\mu_o\mu_r\mathbf{H} \quad (2.18)$$

$$\nabla \times \mathbf{H} = j\omega\varepsilon_o\varepsilon_r\mathbf{E} \quad (2.19)$$

$$\nabla \cdot \varepsilon_o\varepsilon_r\mathbf{E} = 0 \quad (2.20)$$

$$\nabla \cdot \mu_o\mu_r\mathbf{H} = 0 \quad (2.21)$$

where E and H are the electric and magnetic field intensities, ω is the angular frequency, ε_o and ε_r are the permittivity in free space and the relative permittivity, μ_o and μ_r are the permeability in free space and relative permeability and $j = \sqrt{-1}$.

2.2.2.2 Interface Conditions

At the interface of different media, E and H fields and their derivatives may be discontinuous. The interface conditions describe the behavior of the fields at either side of an interface and are:

$$\mathbf{n} \times (\mathbf{E}_1 - \mathbf{E}_2) = 0 \quad (2.22)$$

$$\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_s \quad (2.23)$$

$$\mathbf{n} \cdot (\mathbf{D}_1 - \mathbf{D}_2) = \sigma \quad (2.24)$$

$$\mathbf{n} \cdot (\mathbf{B}_1 - \mathbf{B}_2) = 0 \quad (2.25)$$

where \mathbf{D} and \mathbf{B} are the electric and magnetic flux densities, σ and \mathbf{J}_s are the electric surface charge and current densities and \mathbf{n} is a unit vector normal to the interface.

2.2.2.3 Boundary Conditions

For any differential equation, a boundary condition is necessary for a complete and unique solution. Boundary conditions apply a known value to the tangential field components on a surface. There are three types of boundary conditions that can be applied:

- (a) *Dirichlet* boundary condition constrains the tangential component of the field on the surface to be

$$\mathbf{n} \times \mathbf{E} = \mathbf{E}_o \quad (2.26)$$

If \mathbf{E}_o is a non-zero value, the surface (or boundary) excites the problem. Such a boundary condition could be used to excite a mode at a port; however, a different approach is used in this thesis. If \mathbf{E}_o is zero, the surface is a homogeneous Dirichlet boundary condition which corresponds to a short circuit or electric wall.

- (b) *Neumann* boundary condition is

$$\mathbf{n} \times \nabla \times \mathbf{E} = 0 \quad (2.27)$$

(2.27) can be re-written as

$$\mathbf{n} \times \mathbf{H} = 0 \quad (2.28)$$

and physically corresponds to a magnetic wall.

(c) *Port boundary condition*

Consider the following problem. Let $\mathbf{e}_k^{(q)}$ and $\mathbf{h}_k^{(q)}$ be the transverse electric and magnetic waves of mode k at port q of a waveguide junction (Figure 2.2) that satisfy (2.1).

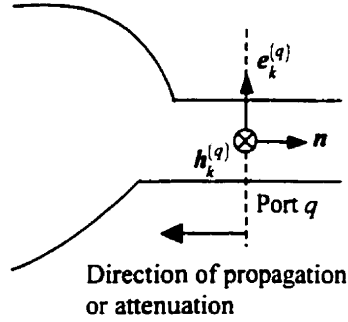


Figure 2.2: Port q .

Using the normalized relations for voltage and current in (2.4) and (2.8), the total current at port q , from (2.12), can be re-written in terms of incident and reflected voltages:

$$I_k^{(q)} = V_k^{(q)+} - V_k^{(q)-} \quad (2.29)$$

From (2.11), we can write the reflected voltage in terms of total and incident voltages where

$$V_k^{(q)-} = V_k^{(q)} - V_k^{(q)+} \quad (2.30)$$

so (2.29) can now be re-expressed:

$$I_k^{(q)} = 2V_k^{(q)+} - V_k^{(q)} \quad (2.31)$$

Suppose that we impose a unit incident wave, on port p for the dominant mode, such that

$$V_0^{(p)+} = \delta_{pq} \delta_{k0} \quad (2.32)$$

(2.31) now becomes

$$I_k^{(q)} = 2\delta_{pq} \delta_{k0} - V_k^{(q)} \quad (2.33)$$

Assuming dominant mode excitation at port p , the total tangential magnetic field on port q , $H_{0r}^{(p)}$, is the sum of all modal tangential fields (from (2.10)). Substituting the expression in (2.33) for the current, the total tangential magnetic field is

$$H_{0r}^{(p)} = \sum_{l=0}^{\infty} (2\delta_{pq} \delta_{k0} - V_l^{(q)}) h_l^{(q)} \quad (2.34)$$

where l is an infinite number of modes and $V_l^{(q)}$ are the modal voltages on port q . As shown in (2.32), for $p = q$, $V_k^{(q)+}$ is a unit excitation for the dominant mode and is unity for mode 0 and zero for all other modes. If port q is not the excited port ($p \neq q$), the incident voltage is at port p , thus the first term of (2.34) becomes zero for all modes. A new boundary condition can be written for the tangential magnetic field at port q due to dominant mode excitation at port p with modal voltages extracted using (2.17) and is written as

$$\mathbf{H}_{0r}^{(p)} = 2\delta_{pq}\mathbf{h}_0^{(q)} - \sum_{l=0}^{\infty} \gamma_l^{(q)} (\mathbf{E}_0^{(p)}) \mathbf{h}_l^{(q)} \quad (2.35)$$

Non-propagating or decaying (evanescent) modes travelling outwards from the device will be absorbed, as if the boundary were not there. This new specification on the port is essentially an absorbing boundary condition for incident waves, which will be called the port boundary condition. Similar formulation for both incident and reflected waves can be derived [29]; however, assuming only incident waves has proven to be an adequately accurate approximation. Properly formulating a “port element”, along with the boundary condition of (2.35), will allow ports to be placed as close as one desires to the junction. The advantages and details of this flexible modeling capability will be described in a later chapter on the “port element”.

2.2.2.4 Helmholtz equation

Taking the curl of (2.18) and substituting in that the expression for the curl of \mathbf{H} from (2.19), we obtain a second order differential equation known as the vector Helmholtz or curl-curl equation:

$$\nabla \times \frac{1}{\mu_r} \nabla \times \mathbf{E} - k_o^2 \epsilon_r \mathbf{E} = 0 \quad (2.36)$$

where k_o is the free-space wavenumber, and $k_o = \omega \sqrt{\epsilon_o \mu_o}$. Similar substitution could yield a curl-curl expression for only \mathbf{H} .

2.2.3 Finite Element Solution of $\mathbf{E}_0^{(p)}$

Solving the wave equation (2.36) for an N -port device of arbitrary shape may not be possible analytically. A numerical CAD tool must then be used to calculate $\mathbf{E}_0^{(p)}$.

The problem is to find $E_0^{(p)}$ when port p is excited with a unit incident wave of dominant mode $k = 0$. The remaining ports are matched at all other modes. The partial differential equation and the corresponding boundary conditions are as follows. The equation that governs the behavior of $E_0^{(p)}$ in the waveguide device is the curl-curl equation, and can be written as

$$\nabla \times \frac{1}{\mu_r} \nabla \times E_0^{(p)} - k_o^2 \epsilon_r E_0^{(p)} = 0 \quad \text{in } \Omega \quad (2.37)$$

The boundary conditions are

$$E_0^{(p)} \times \mathbf{n} = 0 \quad \text{on electric walls } \partial\Omega_e \quad (2.38)$$

$$\nabla \times E_0^{(p)} \times \mathbf{n} = 0 \quad \text{on magnetic walls } \partial\Omega_m \quad (2.39)$$

and the boundary condition on port q is

$$H_{0l}^{(p)} = 2\delta_{pq} h_0^{(q)} - \sum_{l=0}^{\infty} \gamma_l^{(q)} (E_0^{(p)}) h_l^{(q)} \quad \text{on port } q \quad (2.40)$$

where $H_0^{(p)}$ is the magnetic field due to the dominant mode excitation from port p , δ_{pq} is the Kroenecker delta and l are all modes on this the port. The boundaries for the microwave device are shown in Figure 2.3.

2.2.3.1 Weighted Residual Formulation

To calculate the electric field in Ω , with the boundary condition of the form (2.40) at a particular port, additional terms must be added to the familiar weighted-residual formulation of (2.37) [23]. The modified weighted-residual equation for the problem defined in section 2.2.3 is

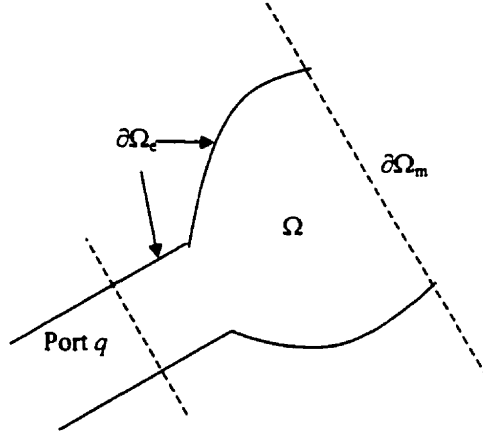


Figure 2.3: Boundary conditions.

$$B(\mathbf{E}_0^{(p)}, \mathbf{w}) = R(\mathbf{w}) \quad (2.41)$$

for all weight functions \mathbf{w} , where B is the bilinear form given by

$$B(\mathbf{E}_0^{(p)}, \mathbf{w}) = \frac{1}{jk_o\eta_o} \int_{\Omega} \left\{ \nabla \times \mathbf{E}_0^{(p)} \cdot \frac{1}{\mu_r} \nabla \times \mathbf{w} - k_o^2 \mathbf{E}_0^{(p)} \cdot \epsilon_r \mathbf{w} \right\} d\Omega \\ + \sum_{p=1}^N \sum_{l=0}^{\infty} \gamma_l^{(p)}(\mathbf{E}_0^{(p)}) \gamma_l^{(p)}(\mathbf{w}) \quad (2.42)$$

and R is the linear function given by

$$R(\mathbf{w}) = 2\gamma_0^{(p)}(\mathbf{w}) \quad (2.43)$$

η_o is the intrinsic impedance of free space.

Equations (2.42) and (2.43) allow the electric field to be found under dominant mode excitation at one or more ports.

2.2.3.2 Discretization

To solve the defined problem numerically, the unknown field must be represented by a finite number of degrees of freedom. The domain or volume of interest, Ω , is broken up into non-overlapping finite elements [18–25]. The grid of nodes and elements is referred to as the finite element mesh. Basis or trial functions approximate the field distribution inside the element [23] so that for a particular element,

$$\underline{E}_0^{(p)}(x, y, z) = \sum_i N_i(x, y, z) E_{0i}^{(p)} \quad (2.44)$$

where N_i is a vector basis function for an element and $E_{0i}^{(p)}$ denotes the corresponding unknown coefficient.

If the weighting function of the weighted residual expression is taken to be a basis function, (2.41) can be re-written as

$$B\left(\sum_j N_j E_{0j}^{(p)}, N_i\right) = R(N_i) \quad (2.45)$$

which becomes

$$\sum_j E_{0j}^{(p)} B(N_j, N_i) = R(N_i) \quad (2.46)$$

which leads to a system of equations of form $[K]\underline{E}_0^{(p)} = \underline{b}$, where $[K]$ is a square, sparse, symmetric matrix, \underline{b} is a known column vector and $\underline{E}_0^{(p)}$ is the vector of unknowns for a particular element. The elemental local matrix $[K]$ and righthand side vector \underline{b} are given by

$$K_{ij} = B(N_i, N_j) = \frac{1}{jk_o \eta_o} \int_{\text{element}} \left(\frac{1}{\mu_r} \nabla \times N_i \cdot \nabla \times N_j - k_o^2 \epsilon_r N_i \cdot N_j \right) d\Omega$$

$$+ \sum_{q=1}^N \sum_{l=0}^{\infty} \gamma_l^{(q)}(N_i) \gamma_l^{(q)}(N_j) \quad (2.47)$$

$$b_i = R(N_i) = 2\gamma_0^{(p)}(N_i) \quad (2.48)$$

Through use of the (2.47), (2.48) and a global numbering scheme, a global matrix $[\tilde{K}]$ and righthand side vector \tilde{b} can be assembled to form a global matrix equation $[\tilde{K}]\tilde{E}_0^{(p)} = \tilde{b}$. Matrix $[\tilde{K}]$ is also sparse and $\tilde{E}_0^{(p)}$ can be now computed using a direct or iterative matrix solver. From $\tilde{E}_0^{(p)}$ and the basis functions, a good approximation of the field $E_0^{(p)}$ can be found at any point in the problem domain.

2.2.4 *H*-plane Waveguide Case

Up to this point, the theory has allowed for microwave junctions of arbitrary shape in three dimensions. However, FE analysis could take hours in computation time had the problems required solution by 3-D vector FE. In order to facilitate testing and optimization, *H*-plane rectangular waveguide junctions will be examined in this thesis as they can be analyzed in 2-D. This kind of analysis is still of great importance to many applications – particularly space technology [30–32].

If each port of an *H*-plane rectangular waveguide junction (Figure 2.4) is excited by its dominant TE₁₀ mode, there is no field variation in the *y* direction and the electric field varies sinusoidally along each port plane [33]. This reduces the curl-curl equation of (2.36) to a scalar Helmholtz equation governing $E_0^{(p)}$, the *y* component of the electric field $E_0^{(p)}$ of this planar device.

Since $\partial E_0^{(p)} / \partial y$ is zero when port *p* is excited with a unit incident wave of dominant mode TE₁₀, the Helmholtz equation and the boundary conditions simplify to a scalar partial differential problem. The scalar Helmholtz equation that governs the behavior of $E_0^{(p)}$ (for the TE₁₀ mode) in the waveguide device is

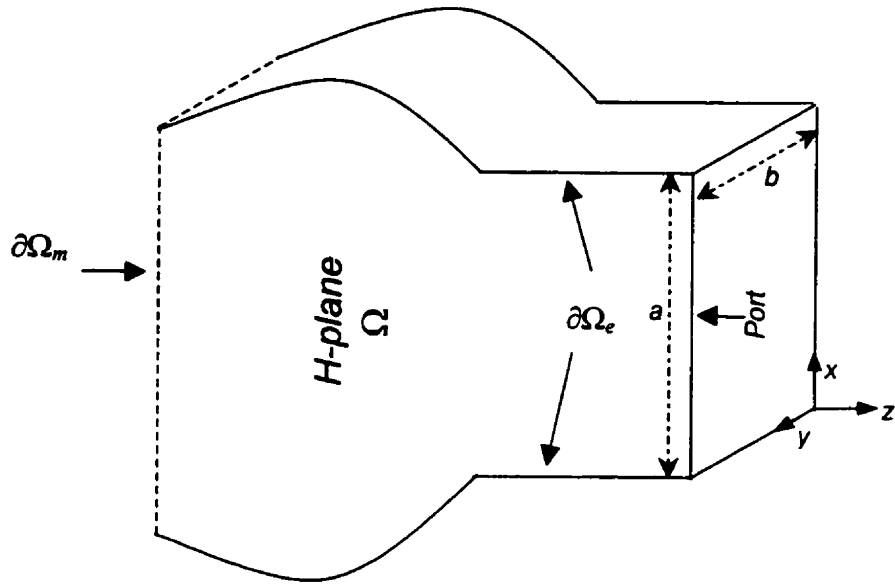


Figure 2.4: *H*-plane rectangular waveguide junction.

$$\nabla \cdot \frac{1}{\mu_r} \nabla E_0^{(p)} + k_o^2 \varepsilon_r E_0^{(p)} = 0 \quad \text{on surface } \Omega \quad (2.49)$$

The corresponding boundary conditions become:

$$E_0^{(p)} = 0 \quad \text{on electric walls } \partial\Omega_e \quad (2.50)$$

$$\frac{\partial E_0^{(p)}}{\partial n} = 0 \quad \text{on magnetic walls } \partial\Omega_m \quad (2.51)$$

and the boundary condition on port q is

$$\mathbf{H}_{0i}^{(p)} = 2\delta_{pq} \mathbf{h}_0^{(q)} - \sum_{l=0}^{\infty} \gamma_l^{(q)} (E_0^{(p)}) \mathbf{h}_l^{(q)} \quad \text{on port } q \quad (2.52)$$

where the boundary locations are shown in Figure 2.4 and

$$\gamma_k^{(q)}(E) = \int_{P_r} E h_k^{(q)} dl \quad (2.53)$$

where $\gamma_k^{(q)}(E)$ is a linear operator extracting the k^{th} mode voltage from a scalar electric field, E . P_r is a port length of the port in the x' -direction (local co-ordinate in the x - z plane) and (2.53) reduces to

$$\gamma_k^{(q)}(E) = \sqrt{\frac{2\beta_k^{(q)}}{ak_o\eta_o}} \int_{x'=0}^a E \sin\left(\frac{k\pi x'}{a}\right) dx' \quad (2.54)$$

for propagating modes k and

$$\gamma_k^{(q)}(E) = (1-j) \sqrt{\frac{\alpha_k^{(q)}}{ak_o\eta_o}} \int_{x'=0}^a E \sin\left(\frac{k\pi x'}{a}\right) dx' \quad (2.55)$$

for evanescent modes k where $\beta_k^{(q)}$ and $\alpha_k^{(q)}$ are the propagation and attenuation constants for TE_{m0} modes and a is actual the length of the 1-D port.

2.2.4.1 Weighted Residual Formulation

The weighted residual for the H -plane rectangular waveguide junction becomes

$$B(E_0^{(p)}, w) = R(w) \quad (2.56)$$

for all scalar weight functions w , where B is the bilinear form given by

$$B(E_0^{(p)}, w) = \frac{1}{jk_o\eta_o} \int_S \left\{ \frac{1}{\mu_r} \nabla E_0^{(p)} \cdot \nabla w - k_o^2 E_0^{(p)} \epsilon_r w \right\} dS + \sum_{p=1}^N \sum_{l=0}^{\infty} \gamma_l^{(p)}(E_l^{(p)}) \gamma_l^{(p)}(w) \quad (2.57)$$

and R is the linear function given by

$$R(w) = 2\gamma_0^{(p)}(w) \quad (2.58)$$

where $\gamma_0^{(p)}(w)$ can be easily evaluated using (2.54) or (2.55).

2.2.4.2 Discretization

Entries of $[K]$ and \underline{b} now become

$$K_{ij} = B(N_i, N_j) = \frac{1}{jk_o\eta_o} \int_{\text{element}} \left(\frac{1}{\mu_r} \nabla N_i \cdot \nabla N_j - k_o^2 \epsilon_r N_i N_j \right) dx dz \quad (2.59)$$

$$+ \sum_{q=1}^N \sum_{l=0}^{\infty} \gamma_l^{(q)}(N_i) \gamma_l^{(q)}(N_j)$$

$$b_i = R(N_i) = 2\gamma_0^{(p)}(N_i) \quad (2.60)$$

2.3 Computing the Gradient of the S -parameters from Field Solutions

To compute the entire scattering matrix for an N -port device, N field solutions are required. Calculating the sensitivity of the scattering matrix by a finite difference formula can be very costly for field-based CAD tools – especially in 3-D. A typical finite forward difference equation for calculating the sensitivities of the scattering matrix with respect to a design parameter g is

$$\frac{\partial[S]}{\partial g} \cong \frac{[S(g + \Delta g)] - [S(g)]}{\Delta g} \quad (2.61)$$

Calculating the gradient (for problems involving more than one design variable) would require N additional field solutions for each design parameter, thus adding greatly to the

computational cost. This section presents a formulation for calculating the sensitivity of an S -parameter where no additional fields need to be calculated.

For problems involving V design parameters, the gradient vector of a scattering parameter is

$$\nabla S_r = \left[\frac{\partial S_r}{\partial g_1} \quad \frac{\partial S_r}{\partial g_2} \quad \dots \quad \frac{\partial S_r}{\partial g_{i'}} \right]^T \quad (2.62)$$

where each entry of vector ∇S_r is just the design sensitivity with respect to each different design parameter, g_k .

Expression (2.16) gives the relation between scattering parameter and field solution. The goal of this section is to obtain a similar relation between the design sensitivity of a scattering parameter, dS_r/dg , and the same field solutions obtained by the 3-D vector finite element formulations from section 2.2.3. Once a formulation for a design sensitivity is derived, assembling the gradient of the S -parameter becomes trivial with the use of (2.62).

2.3.1 The Adjoint Variable Method

An efficient way of calculating design sensitivities of microwave devices is with the adjoint variable method [34]. Assume that \underline{E} is a column vector of unknown field values in a simulated microwave device. Solving for the electric field with the finite element method results in the computation of $[K]\underline{E} = \underline{b}$, as explained in 2.2.3, where $[K]$ is a sparse, square matrix and \underline{b} is a known column vector. Assume g to be a geometric parameter of the modeled device. Any perturbation in g would result in an altered field solution of \underline{E} , thus $[K]$ and \underline{b} will depend on g . If F is a quantity of interest and a function of the field, \underline{E} , it can be shown that [35]

$$\frac{dF}{dg} = \underline{\hat{E}}^T \frac{d\underline{b}}{dg} - \underline{\hat{E}}^T \frac{d[K]}{dg} \underline{E} \quad (2.63)$$

where $\hat{\underline{E}}$ is an adjoint variable, the solution to the adjoint matrix problem, $[\underline{K}]\hat{\underline{E}} = dF/d\underline{E}$. Once the adjoint problem is solved, any of the derivatives of F can be found at a very low computational cost.

2.3.2 Scattering Parameter Sensitivities

The quantity of interest, F , is in our case any scattering parameter S_{rp} . Assume the problem definition for the 3-D finite element formulation in section 2.2.3. Given the equation for S_{rp} in (2.16) and the result of the finite element formulation for matrix $[\underline{K}]$ and vector \underline{b} in (2.47) and (2.48), it can be shown [28] that if port p does not vary with parameter g , the design sensitivity becomes

$$\frac{dS_{rp}}{dg} = -\frac{1}{2} \underline{E}_0^{(r)\top} \frac{d[\underline{K}]}{dg} \underline{E}_0^{(p)} \quad (2.64)$$

The derivative of local matrix $[\underline{K}]$ can be calculated from the derivatives of the vertices of the mesh with respect to the Cartesian co-ordinates [28]:

$$\frac{d[\underline{K}]}{dg} = \sum_l \sum_k \frac{\partial[\underline{K}]}{\partial r_k^l} \frac{dr_k^l}{dg} \quad (2.65)$$

where r_k^l is co-ordinate l of the k^{th} vertex – which is dependent on design variable g .

From (2.64), the design sensitivity can be computed directly from the N field solutions needed to compute the scattering matrix. No additional system of equations need solving – thus providing cheap and efficient calculation of ∇S_{rp} .

2.4 The Cost Function

The previous sections of this chapter described a method for calculating the scattering parameters and their gradients for a microwave device at a single frequency. The tuning of the microwave device by optimizing a cost function with respect to its geometric parameters can be performed for a single frequency or over a range of frequencies*. The computation of any cost function must be repeated at discrete frequency points or some other method found that builds a frequency response of the cost function for a desired range [36–40].

The choice of the cost function is left to the designer. Cost functions are commonly taken to be functions of magnitudes and/or phases of scattering parameters because of their versatility in describing device performance. For the purposes of this thesis, the cost function is found at a number of discrete points over the range of interest and is given by

$$C = \sum_{i=1}^{N_f} w_i c_i \quad (2.66)$$

where w_i is a weighting function, N_f is the total number of frequency points and c_i is a function evaluated at discrete frequency point f_i . Function c_i is a function of the magnitudes and phases of scattering parameters and is in general,

$$c_i = c_i(|S_{11}(f_i)|, \angle S_{11}(f_i), |S_{12}(f_i)|, \angle S_{12}(f_i), \dots, |S_{NN}(f_i)|, \angle S_{NN}(f_i)) \quad (2.67)$$

The choice of S -parameters for a particular c_i is dependent on the frequency. Microwave filters, for e.g., typically allow wave transmission for only a band of frequencies and therefore require different definitions of c_i in the pass and stop bands. Function c_i (for each frequency or range of frequencies) is chosen by the designer to best suit the particular design of the microwave device.

* Note that over the range of frequencies for all microwave devices in this thesis, only the dominant mode can propagate.

2.5 The Gradient of the Cost Function

The gradient of the cost function is a vector given by

$$\nabla C = \left[\frac{\partial C}{\partial g_1} \quad \frac{\partial C}{\partial g_2} \quad \dots \quad \frac{\partial C}{\partial g_v} \right]^T \quad (2.68)$$

where $\partial C / \partial g_k$ is given by

$$\frac{\partial C}{\partial g_k} = \frac{\partial}{\partial g_k} \sum_i c_i = \sum_i \frac{\partial c_i}{\partial g_k} \quad (2.69)$$

c_i is a function of either the magnitude or phase (or both) of S -parameters and the complex value $\partial S_{rp} / \partial g_k$ is readily available from (2.64). Some practical expressions for the sensitivities are given below.

a) *Derivative of $|S_{rp}|^2$ with respect to g_k .*

$$\frac{\partial |S_{rp}|^2}{\partial g_k} = \frac{\partial}{\partial g_k} \left((S'_{rp})^2 + (S''_{rp})^2 \right) \quad (2.70)$$

which becomes

$$\frac{\partial |S_{rp}|^2}{\partial g_k} = \left(2S'_{rp} \frac{\partial S'_{rp}}{\partial g_k} + 2S''_{rp} \frac{\partial S''_{rp}}{\partial g_k} \right) \quad (2.71)$$

or

$$\frac{\partial |S_{rp}|^2}{\partial g_k} = \left(2S_{rp}^r \left(\frac{\partial S_{rp}}{\partial g_k} \right)^r + 2S_{rp}' \left(\frac{\partial S_{rp}}{\partial g_k} \right)' \right) \quad (2.72)$$

b) *Derivative of $|S_{rp}|$ with respect to g_k .*

$$\frac{\partial |S_{rp}|}{\partial g_k} = \frac{\partial}{\partial g_k} \left((S_{rp}^r)^2 + (S_{rp}')^2 \right)^{1/2} = \frac{1}{2|S_{rp}|} \frac{\partial}{\partial g_k} \left((S_{rp}^r)^2 + (S_{rp}')^2 \right) \quad (2.73)$$

and similar to the derivation in a),

$$\frac{\partial |S_{rp}|}{\partial g_k} = \frac{1}{|S_{rp}|} \left(S_{rp}^r \left(\frac{\partial S_{rp}}{\partial g_k} \right)^r + S_{rp}' \left(\frac{\partial S_{rp}}{\partial g_k} \right)' \right) \quad (2.74)$$

c) *Derivative of $|S_{rp}|_{dB}$ with respect to g_k .*

$$\frac{\partial |S_{rp}|_{dB}}{\partial g_k} = \frac{\partial}{\partial g_k} \left(20 \log_{10} |S_{rp}| \right) = \frac{20 \log_{10}(e)}{|S_{rp}|} \frac{\partial |S_{rp}|}{\partial g_k} \quad (2.75)$$

which becomes

$$\frac{\partial |S_{rp}|_{dB}}{\partial g_k} = \frac{20 \log_{10}(e)}{|S_{rp}|^2} \left(S_{rp}^r \left(\frac{\partial S_{rp}}{\partial g_k} \right)^r + S_{rp}' \left(\frac{\partial S_{rp}}{\partial g_k} \right)' \right) \quad (2.76)$$

d) *Derivative of $\angle S_r$ with respect to g_k .*

$$\frac{\partial \angle S_r}{\partial g_k} = \frac{\partial}{\partial g_k} \arctan\left(\frac{S'_r}{S^r_r}\right) = \frac{1}{1 + \left(\frac{S'_r}{S^r_r}\right)^2} \frac{\partial}{\partial g_k} \left(\frac{S'_r}{S^r_r}\right) \quad (2.77)$$

which reduces to

$$\frac{\partial \angle S_r}{\partial g_k} = \frac{S^r_r \left(\frac{\partial S_r}{\partial g_k}\right)' - S'_r \left(\frac{\partial S_r}{\partial g_k}\right)^r}{(S^r_r)^2 + (S'_r)^2} \quad (2.78)$$

Chapter 3

Hierarchical Elements and *P*-Adaption

H-plane waveguide junctions are an important class of microwave devices that allow 2-D analysis. Chapter 2 develops the FE formulation for an *N*-port waveguide junction in the *H*-plane with the new port boundary condition. This chapter focuses on the two main criteria that affect solution accuracy: discretization and basis functions.

Since any finite element solution is a numerical approximation to the actual solution, it will contain a certain error. The solution accuracy is a function of the number and placement of degrees of freedom (DOFs) in the problem. Increasing the number of DOFs can reduce the error in the solution. This can be achieved either by generating a mesh with a larger number of elements or approximating the field solution in each element with higher order basis functions. Increasing the number of DOFs is obviously costly. Although the final matrix equation of the FE method is sparse, the computational cost per solution can still be quite high.

The electric field in a microwave device will have areas where it varies rapidly and other areas where the variation is slow. There is no question that uniform distribution of a large number of DOFs throughout the problem domain yields accurate results. However, a great reduction in computational cost (while not sacrificing accuracy) can be achieved by concentrating the distribution of DOFs to local areas of the problem that require better field approximation.

This chapter addresses the type of discretization for the *H*-plane waveguide junctions, how to control and add the DOFs to the problem and their effect on the FE formulations from Chapter 2.

3.1 Hierarchal Elements

The order of an element refers to the polynomial degree of the basis functions used to approximate the field solution in that element. High-order polynomial basis functions for finite elements are well established for general electromagnetic problems and waveguide analysis [41,23,20]. Lagrangian elements allow high-order polynomial interpolation, but it is not possible for Lagrangian elements of different order to be used together in the same mesh. Since mixing the orders of elements is highly desirable, a different kind of high-order element was devised – the hierarchal element [42]. Hierarchal elements allow adjacent triangles to be of different order without causing any discontinuities in the approximated quantity – thus allowing p -adaption [43].

Hierarchal finite elements have been widely used and were described for the first time almost 30 years ago by Zienkiewicz et al. [44]. Since then, they have been successfully used for finite elements in low frequency electromagnetics [45,46] and for microwave devices [47].

In hierarchal elements, the basis functions for the order $n-1$ element are a subset of the basis functions of the order n element. With this relationship, any edge shared by two elements of different order will contain a set of basis functions common to both elements. For the element of higher order, the coefficients of the basis functions not in the common set are set to zero – thus ensuring continuity across edges. This is not possible with Lagrangian elements. The hierarchal triangular elements used in this thesis make use of orthogonal, Jacobi polynomials, described by Webb and Abouchacra [48]. The basis functions of these elements not only preserve continuity between adjacent elements, but maintain a strong degree of linear independence – thus avoiding potentially ill-conditioned matrices.

The region of interest can be subdivided into non-overlapping finite elements such as rectangles, triangles or curvilinear shapes. The 2-D problems examined in this thesis will be discretized using simple, linear triangular elements.

In dealing with high-order basis functions on triangles, it is usual to introduce simplex (or area) co-ordinates (see Figure 3.1), each ranging from 0 to 1, defined by

$$\zeta_1 = \frac{\Delta_{P23}}{\Delta_{123}}; \quad \zeta_2 = \frac{\Delta_{1P3}}{\Delta_{123}}; \quad \zeta_3 = \frac{\Delta_{12P}}{\Delta_{123}} \quad (3.1)$$

where Δ_{ijk} is the area of triangle ijk . Simplex co-ordinates have the property that $\zeta_1 + \zeta_2 + \zeta_3 = 1$.

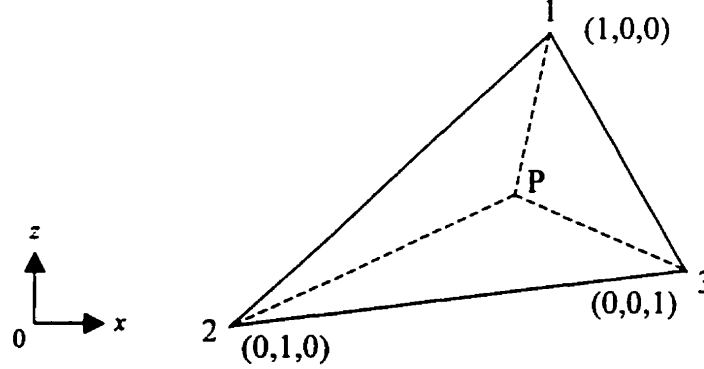


Figure 3.1: Simplex coordinates in a triangular element.

For an element complete to order n , (2.44) can be re-written as

$$E_0^{(p)} = \sum_{i=1}^m a_i N_i(\zeta_1, \zeta_2, \zeta_3) \quad (3.2)$$

where $m = (n+1)(n+2)/2$ is the number of DOFs, N_i are linearly independent polynomial basis functions of order n or less and $\zeta_1, \zeta_2, \zeta_3$ are simplex coordinates, each a function of two global Cartesian co-ordinates (x and z). In a high-order element with Lagrange basis functions [23], the coefficients a_i are actually values of the electric field at the nodes of a triangle, as mentioned in section 2.2.3. For a hierarchal element, the a_i are not necessarily field values. Note that since $\zeta_3 = 1 - \zeta_1 - \zeta_2$, the basis functions N_i may be expressed entirely in terms of ζ_1 and ζ_2 . They will be treated as functions of just ζ_1 and ζ_2 henceforth.

3.1.1 Universal Matrix Assembly

Formulation of the FE matrix was shown in the previous chapter to require integration of the basis functions and their derivatives over the problem domain. These integrals can be evaluated numerically but this procedure can become lengthy for elements of high-order. An alternate approach is to use pre-computed, universal matrices [43,49,50]. The integrals can be expressed in a way independent of the size and shape of the triangular element. The integrals are evaluated only one time and subsequently stored in a universal matrix – used by the FE program for local matrix assembly.

If the basis functions are generated symbolically (by a symbolic algebra package such as Maple or Mathematica) and are assumed to be polynomials, the integration for the universal matrix can be performed exactly [48].

Universal matrices are used for the local matrix assembly for the hierarchal, high order scalar elements of this thesis. The computation of K_{ij} in (2.59) consists of three terms: I_1 , I_2 and I_3 , where

$$I_1 = \frac{1}{jk_o n_o \mu_r} \int_{\Omega^e} \nabla N_i \cdot \nabla N_j d\Omega \quad (3.3)$$

$$I_2 = \frac{k_o \epsilon_r}{jn_o} \int_{\Omega^e} N_i N_j d\Omega \quad (3.4)$$

I_3 is the term that takes the ports into account and $K_{ij} = I_1 - I_2 + I_3$. The integrals in (3.3) and (3.4) are surface integrals over the triangular element where $d\Omega = dx dz$. Since x and z are mapped into normalized simplex co-ordinates and Ω^e is the surface of the triangle, (3.3) becomes

$$I_1 = \frac{2A}{jk_o n_o \mu_r} \int_{\Delta} \nabla N_i \cdot \nabla N_j d\zeta_1 d\zeta_2 \quad (3.5)$$

where A is the area of the triangle, factor $2A$ is just the Jacobian of the mapping from xz to $\zeta_1\zeta_2$ co-ordinates and Δ is the mapped triangle in the $\zeta_1\zeta_2$ plane, i.e. $\zeta_2 = 0 \dots 1 - \zeta_1$ and $\zeta_1 = 0 \dots 1$. Using the chain rule:

$$\nabla N_i \cdot \nabla N_j = \left(\sum_{p=1}^2 \frac{\partial N_i}{\partial \zeta_p} \nabla \zeta_p \right) \cdot \left(\sum_{q=1}^2 \frac{\partial N_j}{\partial \zeta_q} \nabla \zeta_q \right) \quad (3.6)$$

and (3.5) becomes

$$I_1 = \frac{2A}{jk_o n_o \mu_r} \sum_{p=1}^2 \sum_{q=1}^2 \nabla \zeta_p \cdot \nabla \zeta_q \int_{\Delta} \frac{\partial N_i}{\partial \zeta_p} \frac{\partial N_j}{\partial \zeta_q} d\Delta \quad (3.7)$$

where $d\Delta = d\zeta_1 d\zeta_2$. We define three matrices:

$$S_y^{(1)} = \int_{\Delta} \frac{\partial N_i}{\partial \zeta_1} \frac{\partial N_j}{\partial \zeta_1} d\Delta \quad (3.8)$$

$$S_y^{(2)} = \int_{\Delta} \left(\frac{\partial N_i}{\partial \zeta_1} \frac{\partial N_j}{\partial \zeta_2} + \frac{\partial N_i}{\partial \zeta_2} \frac{\partial N_j}{\partial \zeta_1} \right) d\Delta \quad (3.9)$$

$$S_y^{(3)} = \int_{\Delta} \frac{\partial N_i}{\partial \zeta_2} \frac{\partial N_j}{\partial \zeta_2} d\Delta \quad (3.10)$$

The $[S^{(r)}]$ matrices (not to be confused with the scattering matrix) can be pre-computed, are independent of element size or shape, and are actually square universal matrices analogous to the Dirichlet matrix described by Silvester [22]. Using these definitions, I_1 becomes

$$I_1 = \frac{2A}{jk_o n_o \mu_r} \left(\nabla \zeta_1 \cdot \nabla \zeta_1 S_y^{(1)} + 2 \nabla \zeta_1 \cdot \nabla \zeta_2 S_y^{(2)} + \nabla \zeta_2 \cdot \nabla \zeta_2 S_y^{(3)} \right) \quad (3.11)$$

It is easy to see that when the co-ordinate system is changed, the integral in (3.4) is in itself a universal matrix of the form $2A[T]$, where

$$T_y = \int_{\Delta} N_i N_j d\Delta \quad (3.12)$$

The $[T]$ matrix is exactly Silvester's metric [22].

The FE local matrix $[K]$ can now be expressed in terms of the pre-computed matrices:

$$K_{ij} = \frac{2A}{jk_o n_o} \left(\frac{1}{\mu_r} \sum_{p=1}^3 b_p S_{ij}^{(p)} - k_o^2 \epsilon_r T_{ij} \right) + \sum_{q=1}^N \sum_{l=0}^{\infty} \gamma_l^{(q)}(N_i) \gamma_l^{(q)}(N_j) \quad (3.13)$$

where b_p is given by:

$$b_1 = \nabla \zeta_1^2; \quad b_2 = 2 \nabla \zeta_1 \cdot \nabla \zeta_2; \quad b_3 = \nabla \zeta_2^2 \quad (3.14)$$

Pre-computing the $[T]$ matrix and $[S^{(p)}]$ matrices eliminates the need to evaluate the surface integrals in (2.59) for every element (local matrix). For elements with higher order basis functions, this results in a large reduction in computational cost for matrix assembly.

3.1.2 Derivative of the Discretized S-Parameters

It was shown in section 2.3.2 that given the FE solutions for a device and $d[\tilde{K}]/dg$, calculating the sensitivity of any S-parameter becomes trivial. As in the assembly of global matrix $[\tilde{K}]$, the local matrix $d[K]/dg$ must be first computed for each element. The derivative is based on the idea of changing the position of vertex k , while keeping all other vertices fixed (see Figure 3.2).

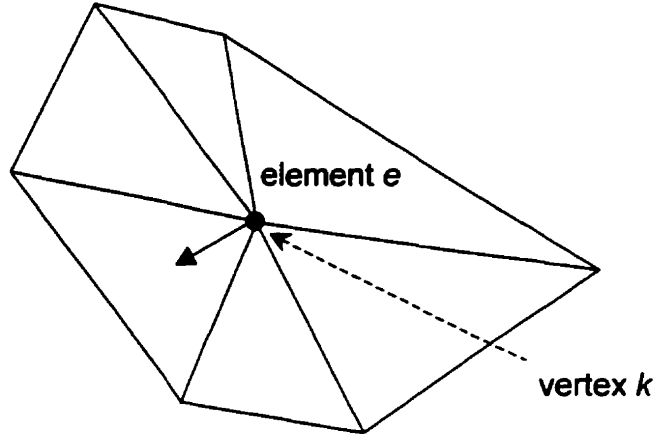


Figure 3.2: Changing position of vertex k .

As shown by (2.65), matrix $d[K]/dg$ can be found from $\partial[K]/\partial r_k^l$, where r_k^l is coordinate l of the k^{th} vertex. (Evaluating $\partial r_k^l/\partial g$ requires little effort, as it is available from a user input data file – depending on the chosen parameter g .) Node k is shared among N_e elements. An entry in matrix $\partial[K]/\partial r_k^l$ is found by summing the total contribution to the derivative at node k from all adjacent elements:

$$\frac{\partial K_{ij}}{\partial r_k^l} = \sum_{e=1}^{N_e} \frac{\partial K_{ij}^e}{\partial r_k^l} \quad (3.15)$$

where $\partial K_{ij}^e/\partial r_k^l$ is the contribution to the ij^{th} entry from the e^{th} element. It is shown in Appendix A that given any set of basis functions for a triangular element:

$$\frac{\partial K_{ij}}{\partial r_k^l} = \frac{2A}{jk_o n_o} \left(\frac{1}{\mu_r} \sum_{p=1}^3 d_{lkp} S_{ij}^{(p)} - k_o^2 \epsilon_r \nabla \zeta_k \cdot \mathbf{a}_l T_{ij} \right) \quad (3.16)$$

A is the area of the triangle and d_{lkp} is given by:

$$\begin{aligned}
d_{lk1} &= \nabla \zeta_1^2 \nabla \zeta_k \cdot \mathbf{a}_l - 2(\nabla \zeta_1 \cdot \mathbf{a}_l)(\nabla \zeta_1 \cdot \mathbf{a}_k) \\
d_{lk2} &= (\nabla \zeta_1 \cdot \nabla \zeta_2) \nabla \zeta_k \cdot \mathbf{a}_l - (\nabla \zeta_1 \cdot \mathbf{a}_l)(\nabla \zeta_2 \cdot \mathbf{a}_k) - (\nabla \zeta_2 \cdot \mathbf{a}_l)(\nabla \zeta_1 \cdot \mathbf{a}_k) \\
d_{lk3} &= \nabla \zeta_2^2 \nabla \zeta_k \cdot \mathbf{a}_l - 2(\nabla \zeta_2 \cdot \mathbf{a}_l)(\nabla \zeta_2 \cdot \mathbf{a}_k)
\end{aligned} \tag{3.17}$$

where \mathbf{a}_l is a unit vector parallel to the axis of co-ordinate l . Co-ordinate $l=1$ is x , $l=2$ is z .

The elegance of the result of the formulation of Appendix A is in the similarity of (3.16) to (3.13). Expression (3.16) uses the same universal $[T]$ and $[S^{(p)}]$ matrices needed for the regular matrix assembly of the FE problem – eliminating the need for evaluating additional integrals for computation of $\partial K_{ij} / \partial r_k^l$. Since the d_{lkp} terms are purely geometric and the same universal matrices are used, evaluation of (3.16) requires very little computational effort.

The formulation for (3.16) is applicable to any set of linearly independent scalar basis functions. It is more general and more efficient than the early formulation in [51].

3.1.3 Orthogonal Polynomial Basis Functions

The basis functions of Webb and Abouchacra [48] for a hierarchal triangle of order n are

$$N_1 = \zeta_1 \tag{3.18}$$

$$N_2 = \zeta_2 \tag{3.19}$$

$$N_3 = \zeta_3 \tag{3.20}$$

For $p = 2, 3, \dots, n$ where $a = p(p+1)/2$,

$$N_{a+1} = E_{p-1}(\zeta_1, \zeta_2) \tag{3.21}$$

$$N_{a+2} = E_{p-1}(\zeta_2, \zeta_3) \tag{3.22}$$

$$N_{a+3} = E_{p-1}(\zeta_3, \zeta_1) \tag{3.23}$$

where E_i are *edge functions* such that

$$E_i = \zeta_1 \zeta_2 P_{i-1}^{(2,2)}(\zeta_2 - \zeta_1) \quad (3.24)$$

$P_i^{(\alpha,\beta)}$ are Jacobi polynomials of order i , are explicitly available in literature [52]. They can be calculated recursively.

For $p = 3, 4, \dots, n$ and $i = 0, 1, \dots, p-3$,

$$N_{a+4+i} = F_{p_i}(\zeta_1, \zeta_2, \zeta_3) \quad (3.25)$$

where F_{p_i} are *face functions* given by

$$F_{p_i} = \zeta_1 \zeta_2 \zeta_3 (1 - \zeta_3)^v P_i^{(2,2v+5)}(1 - 2\zeta_3) P_v^{(2,2)}\left[\frac{\zeta_2 - \zeta_1}{1 - \zeta_3}\right] \quad (3.26)$$

where $v = p-3-i$. It can be shown that all edge functions and face functions are not only linearly independent, but orthogonal. The first four edge and face functions are plotted in [48].

3.1.4 Continuity and Boundary Conditions

Hierarchical basis functions can be used to guarantee field continuity between elements of different orders. For two elements of different orders, the basis functions along their shared edge must be matched. The first three basis functions (N_1, N_2, N_3) are Lagrangian and their coefficients correspond to the nodal electric field values of the element, so continuity of this part of the field is achieved simply by having a single value of field at each vertex in the mesh. Face functions are not of concern because they vanish along any edge of the triangle and thus are not implicated in enforcing continuity.

Assume an element is of order n and has an edge supporting basis functions E_1, E_2, \dots, E_{n-1} and its neighbouring element of higher order p supporting functions

$E_1, E_2, \dots, E_{n-1}, E_n, E_{p-1}$. The coefficients of the edge functions of the neighbouring element (up to $n-1$) are matched to those of the first element while the coefficients of its higher order functions $E_n, E_{n+1}, \dots, E_{p-1}$ are set to zero. The polynomial order along that edge for the neighbour element has dropped to n with matching coefficients – thus ensuring C_0 continuity.

A small problem with matching coefficients arises when edge functions are odd functions of distance about the center of an edge. An edge function of this kind may match the *negative* of the corresponding edge function from a neighbouring triangle. This problem is corrected by numbering the vertices in all triangles of the mesh in a counterclockwise fashion and actually always including a negative value when matching coefficients of edges [48].

Edge functions must be able to approximate the field variation at Dirichlet boundaries. For piecewise linear boundary conditions, all the coefficients of the edge functions are zeroed and vertex values are constrained to those specified for the boundary. For boundaries such as ports, where the electric field varies sinusoidally, the order of the polynomial edge functions must be sufficiently high to accurately approximate that variation. Alternatively, as in this work, the port boundaries can be handled by a special port element.

3.2 Port Element

At a port of a microwave device, there are in general an infinite number of waveguide modes, most of which are non-propagating, or evanescent. Several authors have included in their finite element formulation an expansion of the field at the port in terms of the modal fields in order to model the ports accurately [53–56]. A similar approach was used to derive the boundary condition on a port and resulted in (2.40).

The port terms in the weighted residual formulation of (2.42) and (2.43) are handled with the use of a new “port element” [29]. For 2-D problems, such as the H -plane waveguide junction, this element becomes a line element approximating the same scalar

quantity, $E_0^{(p)}$ (electric field in the y -direction), as inside the problem domain. In order to match the triangular elements along the port, the port element is discretized into segments corresponding to the edges of the triangles. Figure 3.3 shows an example of three elements on the port of different order. The order of the field approximation on each segment must match that of the higher order triangles within. Each segment thus contains high-order, hierarchal, 1-D polynomials that match the basis functions on the edges of the hierarchal triangular elements.

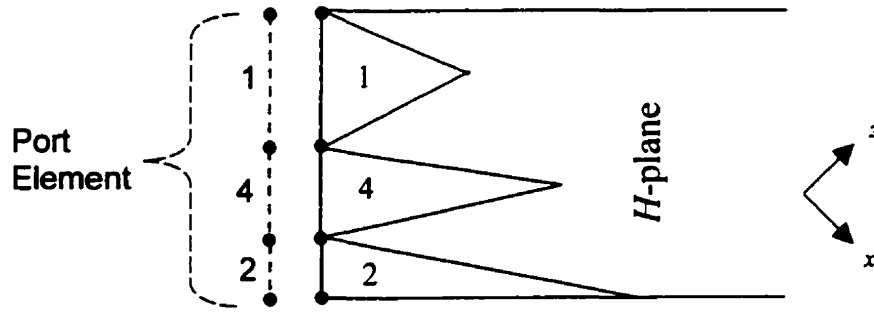


Figure 3.3: A port element. The numbers are the orders of the triangular elements or the polynomial approximation on port element segments.

The field on a segment s of a port element is

$$E = \sum_{i=1}^{order+1} a_i^{(s)} N_i(\zeta) \quad (3.27)$$

where E is the dominant mode scalar electric field in the y -direction, $a_i^{(s)}$ are unknown coefficients, ζ is a normalized length coordinate that varies between 0 and 1 in the segment and N_i are the 1-D polynomial basis functions. To match the hierarchal basis functions described in the previous section, N_i must be defined as:

$$N_1 = \zeta \quad (3.28)$$

$$N_2 = 1 - \zeta \quad (3.29)$$

and for $i=3,4,\dots$

$$N_i = \zeta(1 - \zeta)P_{i-3}^{(2,2)}(1 - 2\zeta) \quad (3.30)$$

where $P_i^{(\alpha,\beta)}(x)$ are Jacobi polynomials [52].

The port terms in (2.59) and (2.60) require the evaluation of the integral in the voltage extraction operator, (2.53). This integral can be decomposed into a sum of integrals over each segment of the port. For each segment, the electric field is replaced by (3.27). The port term in (2.59) is expressed in terms of the unknowns $a_i^{(s)}$ and a stiffness matrix local to the port. The size of the square matrix is just the number of DOFs (D_f) on the port. A similar procedure can be done for the port terms in (2.60), creating a right-hand side vector, D_f entries in length, that represent the excitation of the port in question.

Until now, all port terms involved summations up to an infinite number of modes. The series of summations must be truncated to a finite number of terms (M_p). For modes greater than M_p , the field will be reflected back (and not ideally absorbed) from the port – thus introducing some level of error. An ideal choice for M_p would be to ensure that the amount of error from truncating the series is equal to the discretization error from the FE problem inside the device. If M_p were any higher than this value, instabilities may occur due to the finite element's inadequate representation of these higher modes. It would also account for an unnecessary increase in computational cost for the overall solution. Choosing M_p to match the error in port representation to discretization error is not trivial. A practical choice of M_p is to relate it to the number of DOFs on the port. This method of choosing M_p is very convenient in that the better the discretization of the finite elements at a port, the more modes are taken into account. Both the mesh errors for the regions local to the port as well as the error of the truncation of the modes for the port decrease with higher order basis functions in those elements. This relationship lends itself perfectly to the concept of p -adaption, where the port element is now adaptive because as triangular elements at ports are increased in order, so is the number of modes taken into account. The choice $M_p = D_f/2$ was found to work well in practice [29].

3.3 *P*-Adaption

Adaptive refinement methods in numerical electromagnetics are well-established. They have been used in both FE methods [57–59] and integral methods [60,61]. The concept of adaption is simple, yet extremely effective: analyze a device with the FE method, estimate the error in the computed field, use the error to refine the mesh (i.e. distribute and add DOFs) in key areas, and re-analyze. This process is continued until a given accuracy is achieved. The most important feature in adaption is that given some error tolerances, the process is automated with minimal user interaction. A typical block diagram of an automatic adaptive algorithm is given in Figure 3.4 [62].

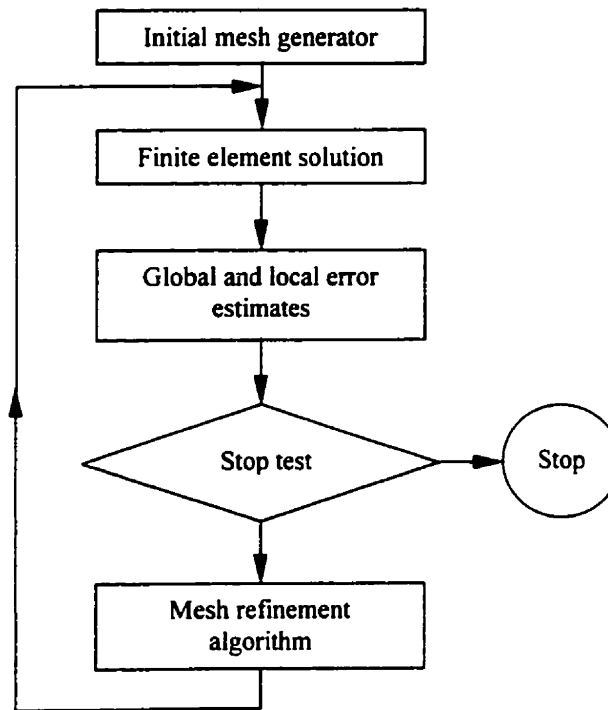


Figure 3.4: Automatic adaptive algorithm.

The concept of adaption is very general and there are many possibilities in the choice of major steps of the algorithm. The three key components are error estimation, error indication and refinement.

An error estimator estimates the error in the quantity of interest in the problem. A good estimate of the error gives the accuracy of the problem solution for a particular FE mesh. The estimate in error can be either local or global to the problem domain. A local error estimator tries to calculate the error in a local quantity of interest (electric field, for example) for a particular element. A global error estimator attempts to calculate how much a global quantity over the problem domain (such as energy) is in error. The estimator is vital to the adaptive algorithm because it governs the termination of the process. Two common termination criteria (the “stop test” block in Figure 3.4) are to adapt until the estimated solution error in each element has been reduced to a certain tolerance [63] or to adapt until the estimated error in a global quantity of interest has reached a certain tolerance. Since the quantity of interest to a microwave designer is a cost function that is a function of S -parameters of a device, any error estimations referred to in this thesis will be global to the problem domain.

An error indicator [64] assesses an error of a particular element, for use in the refinement algorithm’s choice of elements for adding DOFs. While the indicator could be an error estimate in an element, it need not be. It only has to assess the error *relative* to the rest of the elements in the mesh. Note also that an indicator may be an assessment of the local contribution to a *global* quantity [65].

Refinement techniques are methods of adding DOFs to one or many elements in a mesh in order to increase the accuracy of the overall solution. An example of an adaptive strategy is to refine a fixed number of elements at each step (e.g. 25% of all elements). What is common to most methods is that the elements refined are chosen based on their error indication. As mentioned at the beginning of this chapter, to increase the accuracy of the solution, one must increase the DOFs by either adding more elements or increasing the orders of existing elements in the mesh. The process of subdividing an element is referred to as *h*-type adaption [63]. *P*-type adaption (or *p*-adaption) is a method based on increasing the orders (*p*) of elements to increase the accuracy of the solution [66,67]. A combination of both *h* and *p* refinement schemes has also been shown to be very effective [68,69]. Adaptive improvement of the solution without adding DOFs is also possible. The general idea of this scheme is to change the position of existing nodes until they are in near optimal placement (*r*-type adaption) [70].

Both h and p adaption techniques for the analysis of microwave devices are effective [58,59]. Re-meshing can be an expensive procedure – especially in three dimensions. The advantage of p -adaptive refinement is that it needs only a single, initial mesh. A further advantage of p -type methods is specific to wave problems. In quasi-static field problems, a large number of DOFs are needed near singularities. Lower densities of DOFs are sufficient far away from the singularities, where the field becomes increasingly uniform. For wave problems, areas away from singularities can still have large, wave-like field variations and may require a large number of DOFs. P -adaptive techniques typically give a better rate of convergence away from singularities than methods that use uniform, low order elements [58].

In this thesis, p -adaption, using the hierarchal elements of Webb and Abouchacra, is used for refinement. The p -adaptive method is used to compute a cost function (at one frequency) and its gradient, for a microwave device, to a given accuracy. A block diagram specific to the adaptive method that is used in this work is given in Figure 3.5. Variable r is a fixed percent value common to all adaptive steps, f_a is the chosen frequency for the adaptive analysis, and c_a and ∇c_a are defined in section 2.4 and 2.5 to be the cost function and its gradient for a particular frequency. For problems involving analysis over a range of frequencies, the adaption is performed for a particular single frequency f_a (chosen arbitrarily to be the center frequency). A “post-adaptive” frequency sweep is then performed for a discrete number of frequency points in the desired range with the same distribution of DOFs as yielded by the adaptive procedure for the center frequency. For each analysis in the frequency sweep, c_i and ∇c_i are calculated and summed to give the overall cost function and its gradient (defined in (2.66) and (2.68)). For an analysis for a single frequency, no sweep is necessary and the cost function is just c_a .

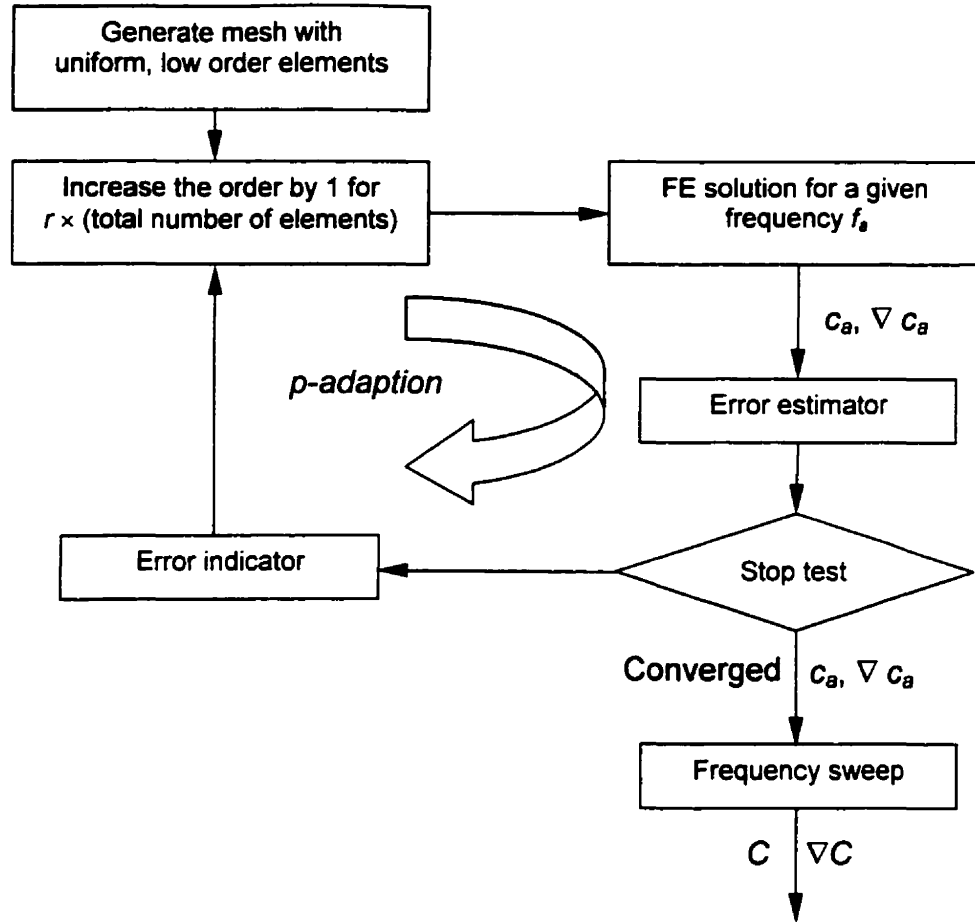


Figure 3.5: P -type adaptive process for microwave devices.

The stop test from Figure 3.5 assesses whether the error estimated in global quantities c_a or ∇c_a is low enough. The error estimate and stop test are crucial to the accuracy control of the cost function and its gradient for the overall optimization, and are discussed in a later chapter. The error indicator provides a local indication of the relative error in c_a for each element. The elements are ranked in order of highest to lowest errors and a percentage $r \times$ (total number of elements) are increased in order by one for that particular adaptive step.

3.4 Error Indication

The goal of the p -adaptive finite element simulation (Figure 3.5) is to reach a required accuracy of c_a in the smallest number of adaptive steps for a particular strategy (e.g. 25% adaption). While the estimate of the error in c_a dictates the stopping point of the algorithm, the error indicator directs the speed of convergence. To achieve a discretization that yields the desired error in cost function, a poor error indicator may require many more adaptive steps than a good one.

The definition of error indicators can vary greatly [71–77]. Many error indicators assess the general quality of the field in an element [74–77]. It was shown that in problems where a global quantity is desired rather than the field itself, a targeted error indicator (TEI) is a better choice [65]. Since the quantities of interest in the analysis of a microwave device are the scattering parameters, a TEI is the most natural choice.

3.4.1 A General Framework for Targeted Error Indicators

A TEI targets areas for mesh refinement based on local contributions to a global quantity of interest, a single real parameter Q . The general idea behind a TEI is the following: let U_1, \dots, U_N be N FE solutions to the problem where $U_i = U_i(x, y, z)$ (for e.g., U_i might be the electric field in a microwave device when port i is excited). Solution U_i^{new} is different from U_i^{old} in some respect, due to a change in the field in a specific region (e.g. element) for which an indicator is needed. Let I_1, \dots, I_N be a set of solution-dependent parameters:

$$I_i = I_i(U_1, \dots, U_N) \quad (3.31)$$

and

$$I_i^{new} = I_i(U_1^{new}, \dots, U_N^{new}) \quad \text{and} \quad I_i^{old} = I_i(U_1^{old}, \dots, U_N^{old}) \quad (3.32)$$

For example, in a two port microwave device, $I_1=S_{11}$, $I_2=S_{12}$, $I_3=S_{22}$. An estimate of the error in I_i arising from the change in field is

$$\Delta I_i = I_i^{new} - I_i^{old} \quad (3.33)$$

Finally, suppose Q is a function of I_1, \dots, I_N . Then an error indicator, ε , for Q , arising from the specified region, is

$$\varepsilon = \left| Q(I_1^{old} + \Delta I_1, \dots, I_N^{old} + \Delta I_N) - Q(I_1^{old}, \dots, I_N^{old}) \right| \quad (3.34)$$

or, alternatively,

$$\varepsilon = \left| \sum_{i=1}^N \frac{\partial Q}{\partial I_i} \Delta I_i \right| \quad (3.35)$$

Varying the definition of U_i^{new} yields different TEIs for Q . The indicator developed in this chapter calculates the effect of each element towards the global quantity. An “ideal” error indicator of this type would be obtained by increasing the order of an element, re-solving the entire FE problem to obtain U_i^{new} and using the difference in global quantities before and after the increase as the indicator of error in that element [78]. While requiring an entire FE solution for each element may be computationally unfeasable, there are indicators that emulate this ideal case successfully [78].

3.4.2 A Targeted Error Indicator for Microwave Devices

For the p -adaptive process for microwave devices of Figure 3.5, a suitable TEI targets the cost function, c_a , and can be developed from the preceding general framework. The solution vector for N individually excited ports is:

$$(U_1^{old}, \dots, U_N^{old})^T = (E_0^{(1)}, E_0^{(2)}, \dots, E_0^{(N)})^T \quad (3.36)$$

where $E_0^{(q)}$ is the field solution when the q^{th} port is excited. Solution U_i^{new} is just U_i^{old} with the k^{th} degree of freedom (DOF) zeroed and all other DOFs remain unchanged. The quantities I_1, \dots, I_N are defined to be the real and imaginary parts of the complex scattering parameters, such that

$$I_1, \dots, I_N = S_{11}^r, S_{11}^i, S_{12}^r, S_{12}^i, \dots, S_{NN}^r, S_{NN}^i \quad (3.37)$$

For calculation of complex S -parameters, expression (2.16) can be alternately expressed in terms of the bilinear form, for the TE_{10} mode, such that

$$S_{rp} = \frac{1}{2jk_o n_o} B(E_0^{(r)}, E_0^{(p)}) - \delta_{rp} \quad (3.38)$$

The weighted residual formulation reduces (3.38) to

$$S_{rp} = \frac{1}{2jk_o n_o} \underline{E}_0^{(r)T} [\tilde{K}] \underline{E}_0^{(p)} - \delta_{rp} \quad (3.39)$$

where $[\tilde{K}]$ is the global FE matrix.

The estimated errors in I_1, \dots, I_N are:

$$\Delta I_1, \dots, \Delta I_N = \Delta S_{11}^r, \Delta S_{11}^i, \Delta S_{12}^r, \Delta S_{12}^i, \dots, \Delta S_{NN}^r, \Delta S_{NN}^i \quad (3.40)$$

where ΔS_{rp} is an estimate of the change in complex S -parameter when the k^{th} DOF is zeroed. The error indicator, based on global quantity $Q (= c_a)$ for the k^{th} DOF is:

$$\varepsilon^k = \left| \sum_i \frac{\partial c_a}{\partial I_i} \Delta I_i \right| = \left| \frac{\partial c_a}{\partial S_{11}^r} \Delta S_{11}^r + \frac{\partial c_a}{\partial S_{11}^i} \Delta S_{11}^i + \dots + \frac{\partial c_a}{\partial S_{NN}^i} \Delta S_{NN}^i \right| \quad (3.41)$$

where the terms of the sum depend on the the initial choice of cost function defined in (2.67).

An error indicator for an n^{th} order element can now be defined as a sum of the contributions of ε^k for all DOFs k that are of order n (i.e. not contained in the element of order $n-1$):

$$\varepsilon = \sum_{n^{\text{th}} \text{ order DOFs}} \varepsilon^k \quad (3.42)$$

While it may seem computationally expensive to compute ΔS_p a number of times for each element in order to evaluate the indicator of (3.42), the programming is quite simple and the computational costs are low. From (3.39), we see that the computational expense in calculating an S -parameter comes from the matrix calculation of the general form:

$$F = \underline{u}^T [\tilde{K}] \underline{v} \quad (3.43)$$

where \underline{u} and \underline{v} are a pair of FE solutions and $[\tilde{K}]$ is the $m \times m$ global matrix. Function F can be decomposed into two parts:

$$F = F_k^z + \Delta F_k \quad (3.44)$$

where F_k^z is F with its k^{th} DOF zeroed:

$$F_k^z = \sum_{\substack{i=1 \\ i \neq k}}^m \sum_{\substack{j=1 \\ j \neq k}}^m u_i \tilde{K}_{ij} v_j \quad (3.45)$$

and ΔF_k can be computed by:

$$\Delta F_k = \sum_{j=1}^m u_k \tilde{K}_{kj} v_j + \sum_{\substack{i=1 \\ i \neq k}}^m u_i \tilde{K}_{ik} v_k \quad (3.46)$$

An efficient algorithm which uses (3.46) to calculate all changes, $\Delta F_1, \dots, \Delta F_m$, is:

$i = 1, \dots, m$:

$$\Delta F_i = 0$$

$i = 1, \dots, m$:

$j = 1, \dots, m$:

If $\tilde{K}_{ij} = 0$, skip (to account for the sparsity of \tilde{K}_{ij})

$$\Delta F_i = \Delta F_i + u_i \tilde{K}_{ij} v_j$$

If $i \neq j$:

$$\Delta F_j = \Delta F_j + u_i \tilde{K}_{ij} v_j$$

This algorithm computes $\Delta F_1, \dots, \Delta F_m$ with no more computational cost than computing $\underline{u}^T [\tilde{K}] \underline{v}$. Using this approach, the estimate of changes of complex S -parameters for every DOF zeroed, $\Delta S_{11}, \dots, \Delta S_{NN}$, need only be computed *once*, globally, and stored for subsequent use. Any estimate of ΔS_{rp} , needed in (3.41) for DOF k , is readily available from global quantities, $\Delta S_{11}, \dots, \Delta S_{NN}$.

Since an estimate exists for every DOF in the FE solution, the change in global cost function need not be limited to only one element. Some variations of (3.42) can possibly provide an effective global error estimator for c_a .

Discussion of error estimation for the gradients of cost functions is postponed to Chapter 5, after a consideration of the optimization process.

Chapter 4

Optimization

Optimization methods have long been used to aid the design of electromagnetic devices [79]. Since microwave devices can be modeled as circuits after sufficient analysis or measurement, much of the practical optimization theory in microwaves originated from the circuit community [80,81].

There are effective optimization algorithms that use actual measurements of device parameters [82,83]. However, with the ever-increasing speed of computer processors, many methods take advantage of the advances in computational analysis [84,85,32,4,51]. Since it has become standard practice for designers to use commercial CAD tools for their own design [86], the CAD industry has noted the importance of incorporating optimizers into its own software packages [87].

While many new optimization techniques have appeared in the past decade [88–91], the debate between deterministic and stochastic algorithms remains. Stochastic optimizers are effective tools for finding global optima in electromagnetic devices [92–94,11,12], but require large numbers of field solutions – which is expensive. Deterministic methods are quick and efficient in finding local optima, but are hampered by their dependence on a well-chosen set of initial parameters to find a global optimum. However, it is not always necessary for a global optimum to be found. As Alotto and Magele point out, “there are many problems, mainly shape optimization problems, where any improvement of the current design can be considered as a success” [95].

Since gradient-based optimizers are particularly powerful, efficient sensitivity analysis is an important facet of the research [96–99]. Using gradient information to guide the design (especially the shape of the device) is highly desirable [100]. Avoiding the numerical computation of the gradient leads to a considerable reduction in computational cost.

In this thesis, which uses optimization to “tune” a microwave device, it is assumed that a reasonable choice has been made for an initial set of geometric parameters. The

block diagram in Figure 1.2 is re-drawn in Figure 4.1 to better resemble the block diagram of a gradient-based deterministic optimizer.

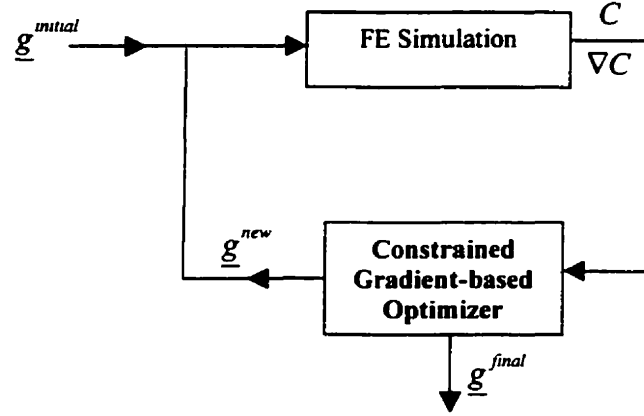


Figure 4.1: Block diagram for gradient-based optimization scheme.

The cost function (chosen by the designer) and its gradient are available from the FE simulation. Two blocks that have been omitted from Figure 4.1 are the parameterization and meshing blocks. The geometric parameters must be parameterized into Cartesian coordinates to generate a data input file for mesh generation. Automatic parameterization of arbitrary geometric structures is a complex area of research on its own [101], and for purposes of this thesis, parameterization specific to a particular problem is used. Parameterization and mesh generation are discussed in Chapter 6.

In the practical design of any device, the geometric parameters, \underline{g} , are in reality limited to certain ranges. In this thesis, the constrained gradient-based optimizer from the MATLAB toolbox [102] is used. The theory that underlies this algorithm is summarized in this chapter.

4.1 The Constrained Optimization Problem

The optimization problem begins with a set of parameters, \underline{g} , that are restricted (or constrained) to acceptable values. The objective or cost function, $C(\underline{g})$, is a real scalar measure of “goodness”, dependent on the parameters [103]. Finding an optimal value for the cost function translates mathematically to either minimizing or maximizing* the cost function. The problem of minimizing an objective or cost function, $C(\underline{g})$, subject to equality and inequality constraints can be characterized by:

$$\begin{aligned} & \underset{\underline{g} \in \mathcal{R}^n}{\text{minimize}} \quad C(\underline{g}) \\ & \text{subject to:} \quad h_i(\underline{g}) = 0, \quad i = 1, \dots, m' \\ & \quad \quad \quad h_i(\underline{g}) \leq 0, \quad i = m' + 1, \dots, m \end{aligned} \tag{4.1}$$

where \underline{g} is a vector of geometric design parameters ($\underline{g} \in \mathcal{R}^n$), C is the cost function ($C: \mathcal{R}^n \rightarrow \mathcal{R}$) and \underline{h} is a vector of constraints ($\underline{h}: \mathcal{R}^n \rightarrow \mathcal{R}^m$).

The Kuhn-Tucker equations are necessary conditions for local optimality of the constrained problem:

$$\begin{aligned} \nabla C(\hat{\underline{g}}) + \sum_{i=1}^m \hat{\lambda}_i \nabla h_i(\hat{\underline{g}}) &= 0 \\ \hat{\lambda}_i h_i(\hat{\underline{g}}) &= 0, \quad i = 1, \dots, m \\ \hat{\lambda}_i &\geq 0, \quad i = m' + 1, \dots, m \end{aligned} \tag{4.2}$$

* Since $\max_{\underline{g}} C(\underline{g}) = \min_{\underline{g}} [-C(\underline{g})]$, without loss of generality, optimal solutions for this chapter will be minima.

where $\underline{\hat{g}}$ is a local minimizer for (4.1) and $\hat{\lambda}_i (i=1, \dots, m)$ are Lagrange multipliers. In the special case where $C(\underline{g})$ and $h_i(\underline{g}) (i=1, \dots, m)$ are convex functions, the Kuhn-Tucker equations are necessary and sufficient conditions for the global optimality of $\underline{\hat{g}}$. The Lagrangian of the problem is defined as:

$$L(\underline{g}, \underline{\lambda}) = C(\underline{g}) + \sum_{i=1}^m \lambda_i h_i(\underline{g}) \quad (4.3)$$

It is easy to see from (4.2) that $\nabla L(\underline{\hat{g}}, \underline{\hat{\lambda}}) = 0$.

In order to calculate the Lagrange multipliers, a nonlinear programming algorithm must be used to solve the Kuhn-Tucker equations. An effective constrained method that uses second order information about the Kuhn-Tucker equations (obtained with a quasi-Newton updating scheme) is Sequential Quadratic Programming [102].

4.1.1 Sequential Quadratic Programming

The Sequential Quadratic Programming (SQP) algorithm [103] attempts to emulate the unconstrained Newton method [102, 104–107]. The general idea is the formulation of a quadratic programming (QP) sub-problem based on a quadratic approximation of the Lagrangian [102]. Given a latest estimate \underline{g}^k of parameters, solve

$$\begin{aligned} \underset{\underline{d} \in \mathcal{R}^n}{\text{minimize}} \quad & \frac{1}{2} \underline{d}^T [H]^k \underline{d} + \nabla C(\underline{g}^k)^T \underline{d} \\ & \nabla h_i(\underline{g}^k)^T \underline{d} + h_i(\underline{g}^k) = 0, \quad i = 1, \dots, m' \\ & \nabla h_i(\underline{g}^k)^T \underline{d} + h_i(\underline{g}^k) \leq 0, \quad i = m' + 1, \dots, m \end{aligned} \quad (4.4)$$

Each major iteration k of the SQP algorithm uses the search direction \underline{d}^k , obtained by solving (4.4) – which can be achieved by any QP algorithm. An appropriate line search

procedure then tries to minimize C along this direction. Figure 4.2 shows the block diagram for the SQP method used in the MATLAB toolbox.

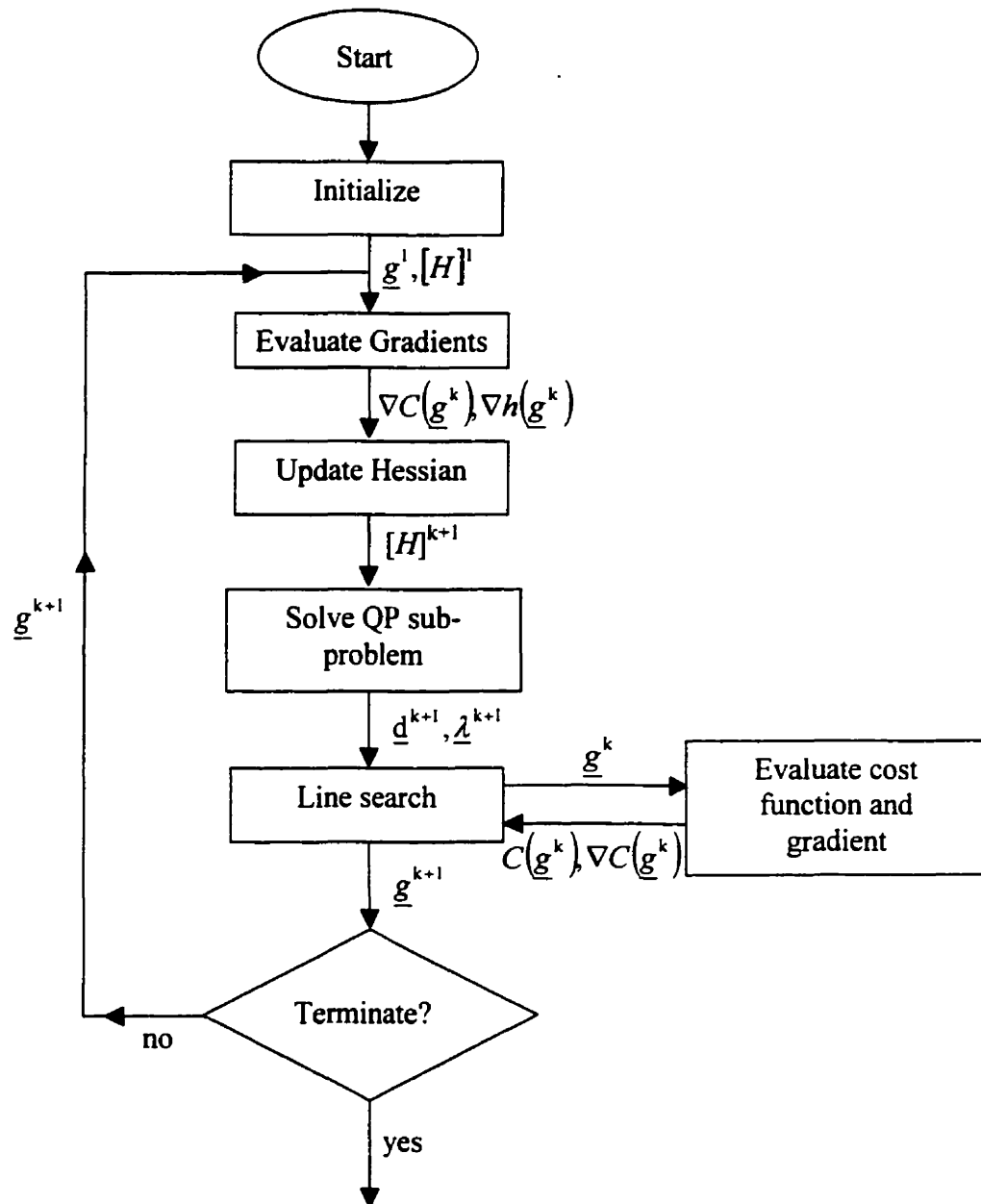


Figure 4.2: Block diagram for SQP in [102].

At each major step of the SQP method, a positive-definite approximation of the Hessian $[H]$ of the Lagrange function is made with a quasi-Newton updating scheme. At the $(k+1)^{\text{th}}$ iteration, an update can be made of $[H]$ using the Broyden-Fletcher-Golfarb-Shanno (BFGS) update, where

$$[H]^{k+1} = [H]^k + \frac{\underline{q}^k \underline{q}^{kT}}{\underline{q}^{kT} \Delta \underline{g}^k} - \frac{[H]^k [H]^k}{\Delta \underline{g}^{kT} [H]^k \Delta \underline{g}^k} \quad (4.5)$$

$$\text{where } \Delta \underline{g}^k = \underline{g}^{k+1} - \underline{g}^k$$

$$\underline{q}^k = \nabla C(\underline{g}^{k+1}) + \sum_{i=1}^m \lambda_i \nabla h_i(\underline{g}^{k+1}) - \left(\nabla C(\underline{g}^k) + \sum_{i=1}^m \lambda_i \nabla h_i(\underline{g}^k) \right)$$

λ_i ($i = 1, \dots, m$) are estimates of the Lagrange multipliers.

The Lagrange multiplier estimates are values that minimize the Lagrangian in a least-squares sense, omitting the multipliers for inactive constraints (because these are zero anyway). If \underline{g}^{k+1} is the latest feasible estimate of the parameters, the active Lagrange multipliers estimates, λ_i^{k+1} ($i = m' + 1, \dots, m$), are given by the minimization:

$$\sum_{i=m'+1}^m \lambda_i^{k+1} \nabla h_i(\underline{g}^{k+1}) = -\nabla C(\underline{g}^{k+1}) \quad (4.6)$$

Multiplying (4.6) by $\nabla h_j(\underline{g}^{k+1})$ for each active constraint gives a matrix equation of the form:

$$[G] \underline{\lambda}^{k+1} = \underline{f} \quad (4.7)$$

where

$$\begin{aligned} G_{ij} &= \nabla h_i(\underline{g}^{k+1}) \cdot \nabla h_j(\underline{g}^{k+1}) \\ f_i &= -\nabla h_i(\underline{g}^{k+1}) \cdot \nabla C(\underline{g}^{k+1}) \end{aligned} \quad (4.8)$$

The active multiplier estimates, $\underline{\lambda}^{k+1}$, can be found by solving the system of equations of (4.7) where $[G]$ is a square, symmetric and positive definite matrix.

A positive-definite (PD) Hessian can be maintained if the initial choice for $[H]$ is PD (e.g. $[H]^1 = [U]$, the identity matrix) and $\underline{q}^{kT} \Delta \underline{g}^k$ is positive at each update [102]. This quantity is kept positive by halving the most negative diagonal element of $\underline{q}^k \Delta \underline{g}^{kT}$ until it is greater than a certain tolerance (e.g. -10^{-5}). If this is not sufficient to make $\underline{q}^{kT} \Delta \underline{g}^k > 0$, \underline{q}^k is modified by:

$$\underline{q}^k = \underline{q}^k + w \underline{v} \quad (4.9)$$

where

$$v_i = \nabla h_i(\underline{g}^{k+1}) h_i(\underline{g}^{k+1}) - \nabla h_i(\underline{g}^k) h_i(\underline{g}^k) \quad (4.10)$$

if $q_i^k w_i < 0$ and $q_i^k \Delta g_i < 0$ and $v_i = 0$ otherwise. The scalar w is increased in size until $\underline{q}^{kT} \Delta \underline{g}^k$ becomes positive [102].

4.1.2 QP Sub-Problem

The MATLAB Optimization Toolbox [102] uses an active set strategy (similar to the projection method described in [108,109]) to solve the QP problem for each major iteration of the SQP algorithm. Expression (4.4) can be re-written as:

$$\begin{aligned} \underset{\underline{d} \in \mathbb{R}^n}{\text{minimize}} \quad & Q(\underline{d}) = \frac{1}{2} \underline{d}^T [H] \underline{d} + \underline{c}^T \underline{d} \\ & \underline{A}_i \underline{d} = b_i, \quad i = 1, \dots, m' \\ & \underline{A}_i \underline{d} \leq b_i, \quad i = m' + 1, \dots, m \end{aligned} \quad (4.11)$$

where \underline{A}_i^T is the i^{th} row of an $m \times n$ matrix $[A]$. The solution to (4.11) involves first calculating a feasible point (should one exist), then generating a set of feasible points that converge to the solution [102]. Let matrix $[\bar{A}]^k$ (size $l \times n$) be an estimate of l active constraints (on the boundaries) of (4.11) at the minimum [102]. $[\bar{A}]^k$ is updated for each iteration within the QP problem, k , and is used to form a search direction, \underline{d}^k (not to be confused with \underline{d}^k of the major iterations for the SQP algorithm). Search direction, \underline{d}^k , minimizes the quadratic function Q while remaining on the boundaries of the active constraints, and a new iterate is desired such that

$$\underline{d}^{k+1} = \underline{d}^k + \alpha \underline{d}^k \quad (4.12)$$

The feasible subspace for \underline{d}^k , guaranteed to remain on the active constraint boundaries, is created from a basis $[Z]^k$ (i.e. \underline{d}^k is a linear combination of the $m-l$ columns of $[Z]^k$, which is $n \times (m-l)$ in size). The columns of $[Z]^k$ are orthogonal to $[\bar{A}]^k$, thus $[\bar{A}]^k [Z]^k = 0$. Matrix $[Z]^k$ is formed from the last $m-l$ columns (l is the number of active constraints) of the QR decomposition of $[\bar{A}]^k$ [102].

Once $[Z]^k$ is calculated, it is desired to find \underline{d}^k that minimizes $Q(\underline{d})$. Since \underline{d}^k is a linear combination of the columns of $[Z]^k$, the search direction can be expressed as $\underline{d}^k = [Z]^k \underline{p}$, where \underline{p} is an unknown column vector. Substituting \underline{p} into (4.8) for the $(k+1)^{\text{st}}$ iteration yields:

$$Q(\underline{p}) = \frac{1}{2} \underline{p}^T [Z]^{kT} [H] [Z]^k \underline{p} + \underline{c}^T [Z]^k \underline{p} \quad (4.13)$$

Differentiating (4.13) with respect to \underline{p} gives:

$$\nabla Q(\underline{p})^{k+1} = [Z]^{kT} [H] [Z]^k \underline{p} + [Z]^{kT} \underline{c} \quad (4.14)$$

If the Hessian, $[H]$, is positive definite (it was computed to be so in 4.1.1), the minimum of $Q(\underline{p})$ can be found by setting its gradient to zero, thus giving a linear matrix equation for \underline{p} :

$$[Z]^{kT} [H] [Z]^k \underline{p} = -[Z]^k ([H] \underline{d}^k + \underline{c}) \quad (4.15)$$

The next search direction can be found from $\underline{d}^{k+1} = [Z]^k \underline{p}$ and step α is taken in direction \underline{d}^k to give the next iterate of parameters for the QP sub-problem, $\underline{d}^{k+1} = \underline{d}^k + \alpha \underline{d}^k$. Since (4.11) is quadratic, a step of unity along \underline{d}^k is exactly the minimum; and if there is no violation of constraints, the QP problem is solved. However, if a constraint is violated, the step α along \underline{d}^k is less than one and the new constraint is added to the active set for the next iteration. The distance to the constraint boundaries in direction \underline{d}^k (towards a constraint boundary) can be found for constraints not in the active set by [102]:

$$\alpha_i = \frac{-(\underline{A}_i \underline{g}^k - b_i)}{\underline{A}_i \underline{d}^k}, \quad i = 1, \dots, m \quad (4.16)$$

This is checked for each constraint not in the active set to make sure that α does not violate these constraints.

When there are n independent constraints in the active set and there is no minimum at that point, the Lagrange multipliers are calculated with a set of non-singular, linear equations [102]:

$$[\bar{A}] \underline{\lambda}^k = [H] \underline{g} + \underline{c} \quad (4.17)$$

When all terms of $\underline{\lambda}^k$ are positive, \underline{d}^k is the minimum of the QP sub-problem in (4.11). If one or more entries of $\underline{\lambda}^k$ are negative (and do not correspond to equality constraints), those elements are removed from the active set and a new iteration begins.

The QP algorithm requires a feasible starting point. In cases when the SQP method yields points that are not feasible for the QP problem, an initialization procedure for the starting point (for the QP sub-problem) must be performed. An initial point can be found by minimizing a linear programming (LP) problem for a slack variable, γ , subject to the same equality constraints as in (4.11) and inequality constraints of the form, $\underline{A}_i \underline{g} - \gamma \leq b_i$. A feasible point can be obtained from the solution of the set of linear equations formed from the equality constraints. If such a solution (or feasible point) exists, γ is set to the maximum value of inequality constraints for that point. The LP problem is iteratively solved using the gradient of C (steepest descent) as its search direction, such that

$$\underline{d}^k = [Z]^{kT} [Z]^k \underline{\nabla} C \quad (4.18)$$

If the LP problems yields a feasible point, the initial search direction for the QP problem is found by solving the matrix equation $[H] \underline{d}^1 = -\underline{\nabla} C$.

If the QP sub-problem does not yield a feasible solution, the search direction for the SQP algorithm is the solution that minimizes γ [102].

4.1.3 Line Search

For each main step of the SQP method, an improved value for the vector of parameters, \underline{g} , is obtained from

$$\underline{g}^{k+1} = \underline{g}^k + \alpha^k \underline{d}^k \quad (4.19)$$

The QP sub-problem provides the search direction, \underline{d}^k . The length of step that should be taken in this direction, α^k , is computed using an appropriate line search method [102]. The line search procedure generates α^k that reduces a merit function [110,102] of the form:

$$\psi(\underline{g}) = C(\underline{g}) + \sum_{i=1}^{m'} r_i |h_i(\underline{g})| + \sum_{i=m'+1}^m r_i \max\{0, h_i(\underline{g})\} \quad (4.20)$$

where r_i is a penalty function [110]:

$$r_i^k = \max\left\{\lambda_i, \frac{1}{2}(r_i^{(k-1)} + \lambda_i)\right\}, \quad i = 1, \dots, m \quad (4.21)$$

The initial penalty function is chosen to ensure larger contributions from constraints with small gradients (i.e. active constraints at the solution point) and is [102]:

$$r_i = \frac{\|\nabla C(\underline{g})\|_2}{\|\nabla h_i(\underline{g})\|_2}, \quad i = 1, \dots, m \quad (4.22)$$

There are many effective line search methods (e.g. Fibonacci, Golden Section) and polynomial methods (e.g. Quadratic, Cubic) that can be used to find α^k . The default method for calculating α^k in MATLAB's constrained optimizer, is bisection.

For every iteration k of the bisection method, a cost function evaluation is performed in order to find a new point, \underline{g}^{k+1} , that satisfies:

$$\Psi(\underline{g}^{k+1}) < \Psi(\underline{g}^k) \quad (4.23)$$

If (4.23) is not satisfied, α^k is halved (or bisected) to form a new step, α^{k+1} [102]. Step size, α^k , is reduced by half for every iteration k until there is a sufficient decrease in the merit function. For each line search, the initial step size is reset to unity.

4.2 Termination Criteria

An important part of the optimization process is knowing when to stop. Judging when the cost function has converged to a local minimum is vital to an efficient algorithm.

4.2.1 General Termination Criteria

For any general optimization problem, there are two common ways to terminate the process [102]:

a) *Termination for \underline{g}*

If the worst case precision required by an independent parameter is ε_g , the optimization can terminate when for every i in vector \underline{g} , at step $k+1$:

$$|g_i^{k+1} - g_i^k| < \varepsilon_g \quad (4.24)$$

b) *Termination for C*

The optimization can similarly be terminated with respect to the precision of the cost function, such that

$$|C^{k+1} - C^k| < \varepsilon_C \quad (4.25)$$

where ϵ_C is the worst case precision for the cost function.

Termination criteria a) and b) can be used independently or together. An overall effective strategy is to use both criteria to ensure proper convergence. For constrained optimization, the constraints must be taken into account and there are many variations of termination criteria that can be used [103].

In the construction of microwave devices, there is a physical limit to the precision of dimensions or material properties. For optimizers terminating on \underline{g} , an obvious choice for ϵ_g is the manufacturing tolerance (e.g. 10^{-5} m for geometric parameters). The constraint violation can also be chosen in the same way. The choice of ϵ_C can be based on the desired precision in the cost function.

4.2.2 Goal Oriented Termination Criteria

Since the cost function of a microwave device has physical meaning beyond its mathematical definition, the termination can be viewed as a goal in the design that is to be achieved by the optimization. While the goal of a design process is very subjective and open-ended, an example is given to illustrate the point.

Assume that the desired design of a band pass filter is an attenuation lower than 0.5 dB in the frequency band and higher than 30 dB outside the upper and lower band frequencies. For each discrete frequency point, the cost function c_i is given as $c_i = |S_{11}|^2$. The global cost function is a sum of c_i of the form (2.66). A possible termination for the optimization with respect to the cost function C is to ensure that for every frequency point i ,

$$|c_i^{k+1} - c_i^k| < 10^{-3} \quad (4.26)$$

or

$$\left| 10 \log \left(\frac{c_i^{k+1}}{c_i^k} \right) \right| < 0.5 \text{dB} \quad (4.27)$$

Expression (4.26) corresponds to smaller variations of $|S_{11}|$ in the pass band and (4.27) is more effective for larger variations.

4.2.3 Gradient Based Termination Criteria

The termination criteria in sections 4.2.1 and 4.2.2 are based on a relative change in cost function (or geometric parameters) between two iterations. A deficiency in such an approach is that while there may be only a small change in cost function for a particular step, the solution can still be far from the minimum. If the change is small enough to satisfy any of the criteria described previously, it will prematurely terminate the optimization without reaching the minimum. An alternative is to use a residual approach, based on the gradient of the cost function.

At a local minimum for an unconstrained problem, the gradient of the cost function is zero. In lieu of monitoring the relative change in cost function from step to step, an absolute termination with regard to the Euclidean norm of the gradient of that cost function can serve as an effective indicator of the optimizer's convergence to a local minimum. Such a termination criterion is:

$$\|\nabla C(\underline{g}^{k+1})\|_2 < \frac{1}{R} \|\nabla C(\underline{g}^{\text{initial}})\|_2 \quad (4.28)$$

where R is a constant, chosen by the designer. The concept is to reduce the magnitude of the gradient to one R^{th} of its initial value. $R=1000$ was found to work well.

For a constrained problem, in addition to the norm of the gradient, the constraint vector, \underline{h} , must be accounted for. From the Kuhn-Tucker equations (4.2), the local minimum of a constrained problem satisfies, by definition,

$$\|\nabla L(\underline{g}^{k+1})\|_2 = 0 \quad (4.29)$$

where

$$\|\nabla L(\underline{g}^{k+1})\|_2 = \left\| \nabla C(\underline{g}^{k+1}) + \sum_{i=1}^m \lambda_i^{k+1} \nabla h_i(\underline{g}^{k+1}) \right\|_2 \quad (4.30)$$

and λ_i^{k+1} is an estimate of the i^{th} Lagrange multiplier at the $(k+1)^{\text{th}}$ iteration. Expression (4.28) can be adjusted to:

$$\|\nabla L(\underline{g}^{k+1})\|_2 < \frac{1}{R} \|\nabla L(\underline{g}^{\text{initial}})\|_2 \quad (4.31)$$

This termination criterion is used in the problems solved in this thesis to avoid the problem of premature termination to which the other criteria are prone. In addition, no knowledge of the minimum of the cost function is needed and convergence is dependent solely on gradient information.

Chapter 5

Controlling Accuracy during Optimization

Many optimization problems deal with objective or cost functions that can be determined analytically. In such cases, the cost function is in effect “infinitely” accurate and accuracy of cost function evaluation (CFE) is not a concern. In problems where a CFE is calculated using a numerical approximation, as in the present work, the accuracy of the CFE becomes an issue.

The highest level of accuracy that can be achieved in a CFE by a FE simulation depends on the performance of the computer being used. High accuracy requires a large number of DOFs (whether it arises from dense meshes or high-order basis functions), leading to large matrices. Solving large matrix problems creates three major demands of a computer: large amounts of RAM, a lot of disk space and high processor speed. The machine must have a large amount of memory available to hold part or all of the large matrix. Lots of storage space may be required if matrix files are written and read from the hard disk (the case in direct matrix solvers). The processor’s ability to execute floating-point operations quickly greatly affects the overall time it will take to solve the matrix problem. Achieving high accuracy in a CFE can result in a long computation time.

In the optimization of microwave devices, where each CFE requires full-wave electromagnetic simulation, the computational cost can be excessive. Since computational cost is dependent on the accuracy required in a CFE, varying the accuracy of CFEs throughout an optimization design can save a lot of computation time. This is the central idea of this thesis, and is explored in detail in this chapter.

5.1 Accuracy Control of the CFE

Figure 5.1 shows what a plot of cost function versus a single design parameter may look like for three different levels of accuracy.

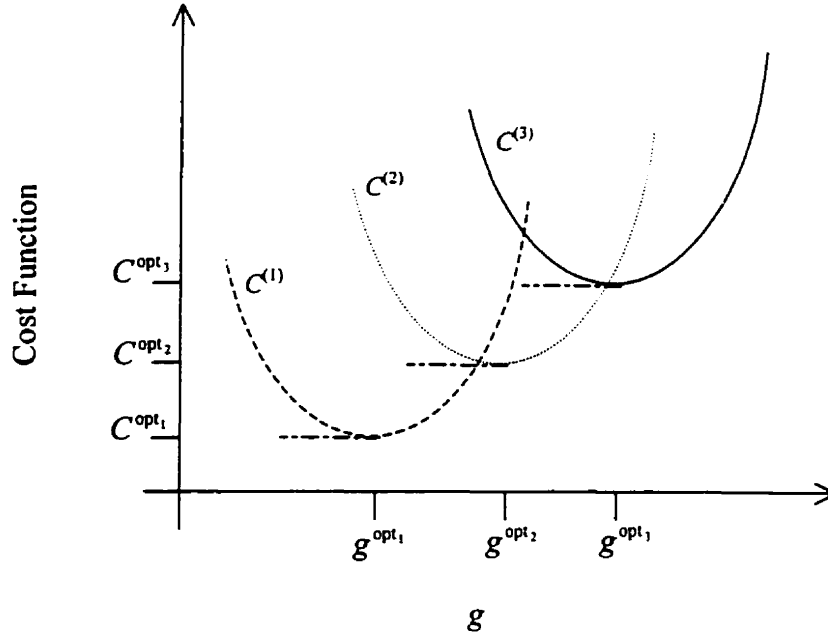


Figure 5.1: Graph of variation of cost function, $C(g)$, at three different accuracy levels.

Each cost function has its own variation with parameter g and there are thus three different optima: g^{opt_1} , g^{opt_2} and g^{opt_3} . If $C^{(3)}$ is the cost function of highest accuracy possible, and $C^{(1)}$ the least accurate, the minimum of $C^{(1)}$ is $\Delta g = g^{\text{opt}_2} - g^{\text{opt}_1}$ away from the “true” optimum of the problem (i.e. the optimum of the highest accuracy cost function).

The interest to a microwave designer in using a CAD tool to simulate and optimize the design of a microwave device is to quickly generate a solution that meets the design specifications. Two parameters that must be chosen are the accuracy of the optimization process itself (termination criteria) and the accuracy of the cost function at the optimum. As seen in Figure 5.1, it is important that once the optimizer finds an optimum, $\underline{g}^{\text{opt}}$, it is

based on an accurate representation of the cost function for those parameters. One way of ensuring that an optimizer finds such an accurate optimum is to perform every CFE as accurately as that of the final CFE. However, heavy computation can be avoided by using CFEs of lower accuracy for steps where \underline{g}^k is far from the optimum. The early steps of an optimization scheme may not require CFEs that are very accurate. As the process continues and the parameters near an optimum, every CFE can become increasingly accurate. The design space of the cost function in such a scenario is dynamic in nature and is allowed to change (rather than stay fixed) throughout the optimization until it gets close to an optimum, where the changes in the design space become very small.

An automated optimization scheme “links” the optimizer to each CFE in order to control the level of accuracy of the CFE throughout the optimization (see Figure 5.2), instead of demanding a fixed accuracy.

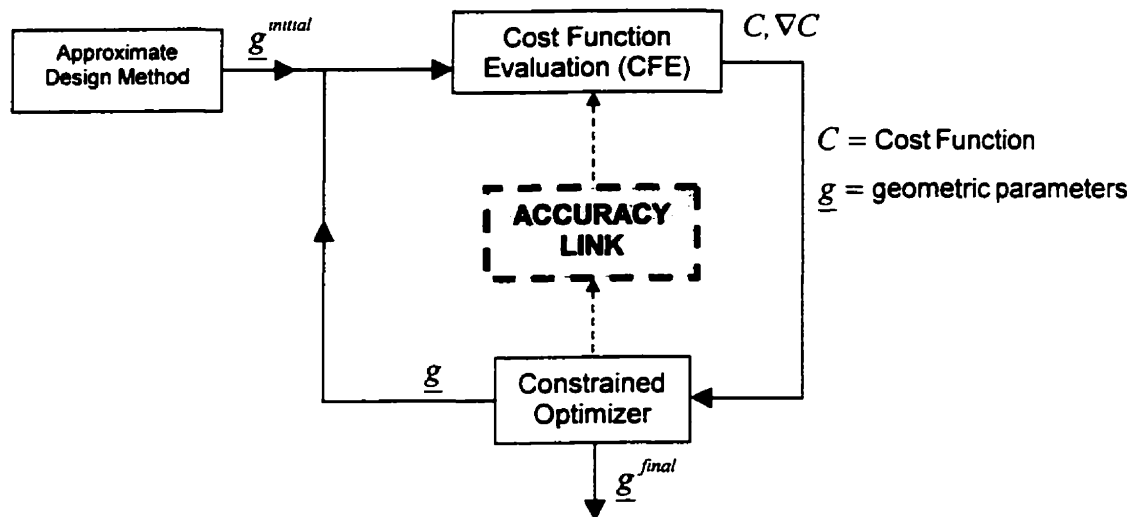


Figure 5.2: Block diagram an optimization scheme with an accuracy link.

5.2 Accuracy Link between Optimization and FE Adaption

The accuracy of an adaptive FE solution is dependent on an error estimate. When the error estimated in a FE solution at an adaptive step is below a certain tolerance, the adaptive process stops. The accuracy link can increase the accuracy of a CFE by reducing the error tolerance.

A quantity that decreases throughout an optimization is the gradient of the cost function (or the gradient of the Lagrangian, $\nabla L(\underline{g})$, for constrained problems). At an optimum, theoretically, $\|\nabla L(\underline{g})\|_2 = 0$. We could, then, require of the CFE that

$$\|\text{error in } \nabla L(\underline{g}^k)\|_2 < \alpha \|\nabla L(\underline{g}^{k-1})\|_2 \quad (5.1)$$

As the optimization progresses, this would hold the percentage error in $\|\nabla L(\underline{g})\|_2$ fixed at a level determined by α .

As $\|\nabla L(\underline{g}^k)\|_2$ reduces with an increase in k , the error in $\|\nabla L(\underline{g}^k)\|_2$ will reduce – thus increasing the accuracy of the FE solution as an optimum is approached. Since the CFE* computed with the adaptive FE method described in this thesis calculates both the cost function and its gradient, demanding high accuracy in the gradient ensures an increased number of DOFs and thus there will be also be high accuracy in the cost function. Computing the actual error in $\|\nabla L(\underline{g}^k)\|_2$ is of course not possible and must be replaced by an estimate of that quantity from the FE solution – discussed in section 5.3, below.

Applying this approach to the p -adaptive FE process from Figure 3.5 requires some modification. For problems involving analysis over a range of frequencies, the cost function and its gradient (and in turn, $\|\nabla L(\underline{g}^k)\|_2$) is calculated as a sum over a number of discrete frequency points. Since the adaptive process is performed for a single frequency, the estimate of the error in $\|\nabla L(\underline{g}^k)\|_2$ should also be calculated for a single frequency.

* Note that CFE will refer to the evaluation of both the cost function and its gradient from here on in.

The adaptive frequency (frequency at which the adaption is performed) is chosen to be the center point of the frequency range. Requirement (5.1) can be approximated by

$$N_f \left\| \text{estimate of error in } \nabla L_a(\underline{g}^k) \right\|_2 < \alpha \left\| \nabla L(\underline{g}^{k-1}) \right\|_2 \quad (5.2)$$

where a is the center point of the frequency range; $\nabla L_a(\underline{g}^k)$ is the gradient of the Lagrangian of the single frequency cost function, c_a , and N_f is the number of frequency points sampled in that range. Since $\nabla L_a(\underline{g}^k)$ is calculated by:

$$\nabla L_a(\underline{g}^k) = \nabla c_a(\underline{g}^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(\underline{g}^k) \quad (5.3)$$

we have from (5.2)

$$N_f \left\| \text{est. error in } (\nabla c_a(\underline{g}^k)) + \text{est. error in } \left(\sum_{i=1}^m \lambda_i^k \nabla h_i(\underline{g}^k) \right) \right\|_2 < \alpha \left\| \nabla L(\underline{g}^{k-1}) \right\|_2 \quad (5.4)$$

Since we do not have error estimates for the Lagrange multipliers, we drop this term and require simply that

$$N_f \left\| \text{estimate of error in } \nabla c_a(\underline{g}^k) \right\|_2 < \alpha \left\| \nabla L(\underline{g}^{k-1}) \right\|_2 \quad (5.5)$$

Any CFE for the optimization has two parts. The first part is the p -adaptive process where the FE solution adapts at the center frequency until (5.5) is satisfied. (The error indicator used for the adaption is (3.41), targeted towards the single frequency cost function, c_a .)

After the adaption is complete, the second part of the FE analysis is a post-adaptive frequency sweep over the N_f points performed with the same distribution of DOFs as the

final step of the adaption. The center frequency is available from the adaptive part and does need to be re-calculated, so the frequency sweep is actually performed over $N_f - 1$ points. The result of the frequency sweep is the cost function and its gradient – the desired quantities needed as input to the optimizer.

The adaption need not necessarily be performed over a single frequency. Alternatively, a frequency sweep could be performed at each adaptive step to calculate the full gradient, $\nabla C(\underline{g}^k)$. However, this would be very costly computationally and the approximation of $\|\nabla c_a(\underline{g}^k)\|_2 \cong \|\nabla C(\underline{g}^k)\|_2 / N_f$ works adequately in practice.

An optimization-adaption system is illustrated as a block diagram in Figure 5.3. In order to allow the first CFE of the optimization to be adaptive, a value for $\|\nabla L(\underline{g}^0)\|_2$ is required. To provide a rough reference value, a pre-optimization CFE is performed non-adaptively at a uniform low order (chosen to be 2nd order). The geometry for this pre-optimization step is the same as in the first optimization step. The approximation is made that $\|\nabla L(\underline{g}^0)\|_2$ is approximately equal to the pre-optimization computation of $\|\nabla L(\underline{g}^1)\|_2$. This quick, uniform order CFE adds little to the overall computational cost yet allows adaptive solving for the first CFE of the optimization.

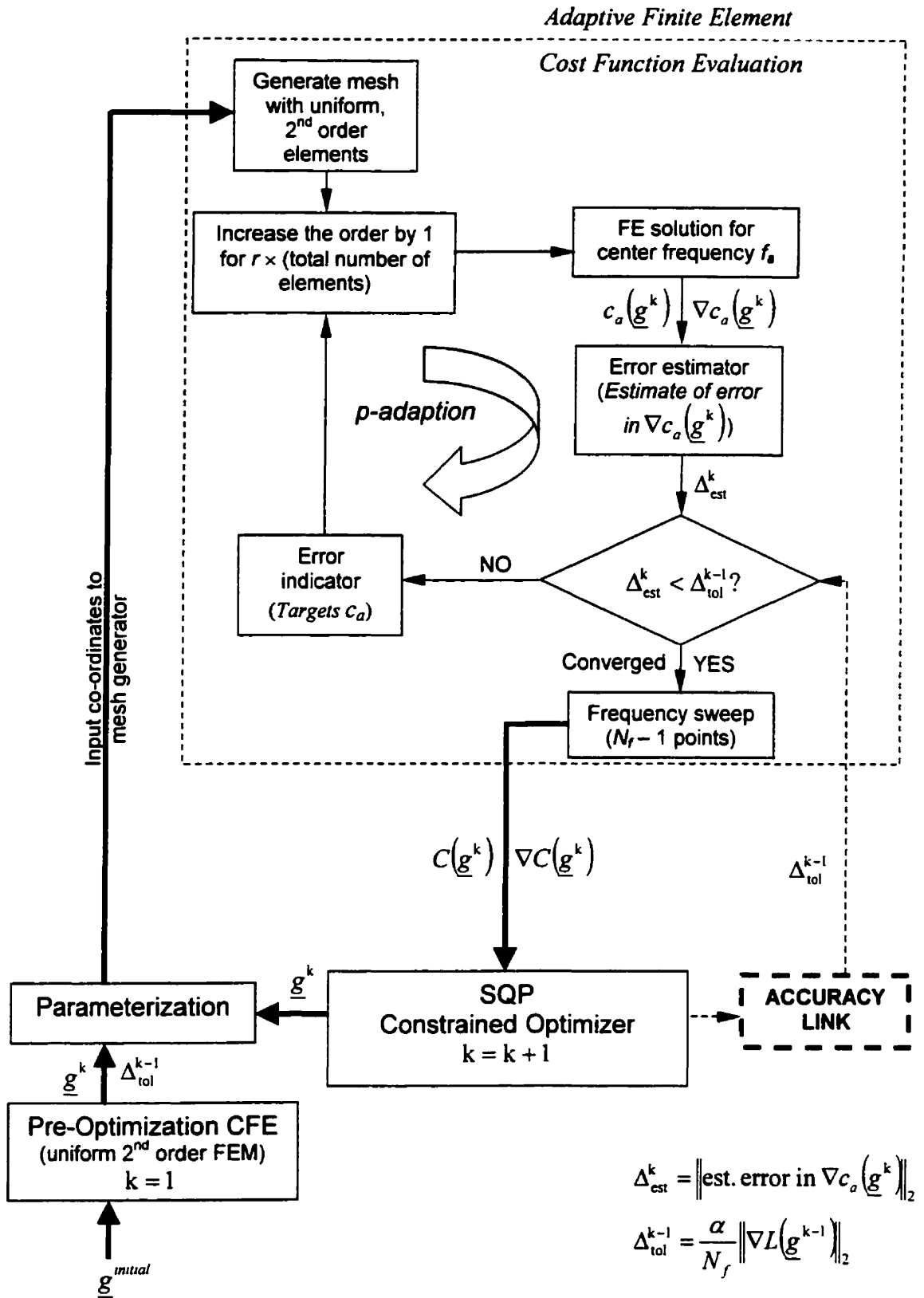


Figure 5.3: Optimization-adaption block diagram.

5.3 Estimating the Error in the Gradient of the Cost Function

Accurate error estimation in any quantity used for terminating an adaptive FE process is very important. Underestimating the error can lead to premature convergence, since the actual FE solution is not as accurate as the estimation predicts. Overestimating the quantity is generally safer than underestimation because the adaption will not be terminated too quickly. A good estimator computes an error not too different from the actual error.

The proposed error estimator for the gradient of the cost function works in the following way. Assume cost function c_a (which is any function of scattering parameters) varies with geometric parameter g_l . The derivative of c_a with respect to g_l can be expressed as:

$$\frac{\partial c_a}{\partial g_l} = \sum_k \frac{\partial c_a}{\partial x_k} \frac{\partial x_k}{\partial g_l} + \frac{\partial c_a}{\partial y_k} \frac{\partial y_k}{\partial g_l} + \frac{\partial c_a}{\partial z_k} \frac{\partial z_k}{\partial g_l} \quad (5.6)$$

where k are nodes of the finite elements on any boundary or interface parameterized by g_l . The error in the derivative of c_a at those nodes is

$$\Delta \frac{\partial c_a}{\partial g_l} \leq \sum_k \Delta \left| \frac{\partial c_a}{\partial x_k} \frac{\partial x_k}{\partial g_l} \right| + \Delta \left| \frac{\partial c_a}{\partial y_k} \frac{\partial y_k}{\partial g_l} \right| + \Delta \left| \frac{\partial c_a}{\partial z_k} \frac{\partial z_k}{\partial g_l} \right| \quad (5.7)$$

where $\Delta \frac{\partial c_a}{\partial g_l}$ is the magnitude of the error in $\frac{\partial c_a}{\partial g_l}$. To estimate $\Delta \frac{\partial c_a}{\partial x_k}, \Delta \frac{\partial c_a}{\partial y_k}, \Delta \frac{\partial c_a}{\partial z_k}$, we

consider the internal nodes of the mesh. An internal node is a node of a finite element whose movement does not alter the geometry of the problem being solved. In other words, these are nodes in a mesh that are not located on any boundary or interface in the problem. Since the geometry does not vary with movement of these internal nodes, neither should c_a . The fact that c_a does change is a result of discretization error. An

estimate of $\Delta \frac{\partial c_a}{\partial x_k}$ or $\Delta \frac{\partial c_a}{\partial y_k}$ or $\Delta \frac{\partial c_a}{\partial z_k}$ is

$$\frac{1}{N_i} \sum_{i=1}^{N_i} \max \left(\left| \frac{\partial c_a}{\partial x_i} \right|, \left| \frac{\partial c_a}{\partial y_i} \right|, \left| \frac{\partial c_a}{\partial z_i} \right| \right) \quad (5.8)$$

where i is an internal node and N_i is the total number of internal nodes in the mesh. If, further, $c_a = |S_{pq}|^2$,

$$\left| \frac{\partial c_a}{\partial x_i} \right| \leq 2(|S_{pq}^r| + |S_{pq}^i|) \max \left(\left| \frac{\partial S_{pq}^r}{\partial x_i} \right|, \left| \frac{\partial S_{pq}^i}{\partial x_i} \right| \right) \quad (5.9)$$

and an upper bound for (5.8) is

$$2(|S_{pq}^r| + |S_{pq}^i|) S_m \quad (5.10)$$

where

$$S_m = \frac{1}{N_i} \sum_{i=1}^{N_i} \max \left(\left| \frac{\partial S_{pq}^r}{\partial x_i} \right|, \left| \frac{\partial S_{pq}^i}{\partial x_i} \right|, \left| \frac{\partial S_{pq}^r}{\partial y_i} \right|, \left| \frac{\partial S_{pq}^i}{\partial y_i} \right|, \left| \frac{\partial S_{pq}^r}{\partial z_i} \right|, \left| \frac{\partial S_{pq}^i}{\partial z_i} \right| \right) \quad (5.11)$$

The RHS of (5.7), with (5.10) replacing $\Delta \frac{\partial c_a}{\partial x_k}$ etc., serves as an estimate for the error in

$$\frac{\partial c_a}{\partial g_i} :$$

$$\Delta \frac{\partial c_a^{est}}{\partial g_i} = 2(|S_{pq}^r| + |S_{pq}^i|) S_m \sum_k \left(\left| \frac{\partial x_k}{\partial g_i} \right| + \left| \frac{\partial y_k}{\partial g_i} \right| + \left| \frac{\partial z_k}{\partial g_i} \right| \right) \quad (5.12)$$

Note that all derivatives of scattering parameters are already available (at no extra cost) for every adaptive step from the adjoint variable method described in Chapters 2 and 3.

The derivative terms $\frac{\partial x_k}{\partial g_l}$, $\frac{\partial y_k}{\partial g_l}$ and $\frac{\partial z_k}{\partial g_l}$ are purely geometrical and again already available from the FE formulation previously described. The summation from (5.12) of the absolute values of these terms needs to be performed only once, at the beginning of the adaptive FE solution.

The cost function does not need to be of the form $c_a = |S_{pq}|^2$. However, (5.12) changes depending on the cost function. Using the practical expressions for sensitivities from section 2.5 and similar derivation as (5.9) through (5.12), two other practical estimates of error of cost function derivatives are for:

a) $c_a = |S_{pq}|$

$$\Delta \frac{dc_a^{est}}{dg_l} = \frac{|S_{pq}'^r| + |S_{pq}'^i|}{|S_{pq}|} S_m \sum_k \left(\left| \frac{\partial x_k}{\partial g_l} \right| + \left| \frac{\partial y_k}{\partial g_l} \right| + \left| \frac{\partial z_k}{\partial g_l} \right| \right) \quad (5.13)$$

b) $c_a = \angle S_{pq}$

$$\Delta \frac{dc_a^{est}}{dg_l} = \frac{|S_{pq}'^r| + |S_{pq}'^i|}{|S_{pq}|^2} S_m \sum_k \left(\left| \frac{\partial x_k}{\partial g_l} \right| + \left| \frac{\partial y_k}{\partial g_l} \right| + \left| \frac{\partial z_k}{\partial g_l} \right| \right) \quad (5.14)$$

In both (5.13) and (5.14), S_m is given by (5.11).

For error estimators of this type to work, the FE mesh must contain at least one internal node. All estimates of this type are based on errors in the derivatives of global quantities. For that reason, it is probably preferable to ensure that a mesh contains enough internal nodes to represent *all* areas of the geometry. A well discretized mesh for use with these types of estimators has internal nodes in the vicinity of each node located on a parameterized boundary or interface.

5.4 Issues Arising from the Combination of Optimization and Adaption

A number of important issues arose during the construction and testing of the combined optimization-adaption system.

5.4.1 Implications of the Choice of α

In general, α controls two aspects of the optimization: final accuracy of the answer and the way the accuracy is adjusted en route.

The accuracy of any cost function is set at a fraction α of the previous CFE's $\|\nabla L(\underline{g}^k)\|_2 / N_f$. Since the termination criterion for the optimization is that $\|\nabla L(\underline{g}^k)\|_2 < 0.1\% \|\nabla L(\underline{g}^1)\|_2$, the accuracy level of the final CFE must be less than α/N_f of $0.1\% \|\nabla L(\underline{g}^1)\|_2$. The choice of α in this case is therefore a control over the final accuracy of the answer.

The choice of α also determines the speed at which the accuracy levels of the CFEs change. A large α means that until \underline{g}^k is in the vicinity of an optimum, the CFEs throughout the optimization will have a low accuracy level. Alternatively, a small α means not only an accurate final answer, but also more accuracy (and cost) at *every* stage of the optimization.

The fact that α controls both these aspects is not necessarily ideal. Schemes could be developed to de-couple them, e.g. using two different α 's – one to control the final accuracy and the other controlling the path taken to get there.

A practical limitation is the possible inability to achieve the required accuracy of a CFE. A limitation to using only a p -adaptive process is saturation of all elements at a maximum of 10^{th} order. In the present work, while α theoretically controls both aspects of accuracy, in nearly all cases it actually had no effect on the final accuracy due to this saturation – so it just controlled efficiency.

5.4.2 Continuity of C and the Effects of Re-Meshing

Given any geometry \underline{g}^k , the parameterization block writes out the corresponding co-ordinates of the nodes as a data file – later used as input to the mesh generator. Because of the new parameterization and re-meshing, a change in the geometry (even a small one) can change the structure of a mesh. A subtlety that was encountered was the phenomenon of “diagonal flipping”. For very small changes in geometry, the structure of the mesh (around the change in geometry) might change. This change may be as subtle as a diagonal being flipped between two adjacent triangles (shown in Figure 5.4).

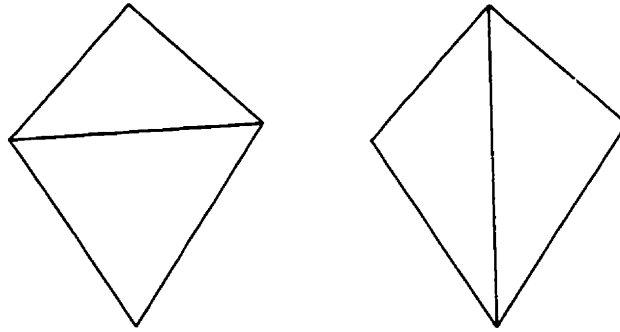


Figure 5.4: Diagonal flipping between adjacent triangular elements.

Different variations in mesh discretization will give different FE solutions to C . The numerical cost function is thus discontinuous (see Figure 5.5). Small discontinuities in C throughout the optimization are usually acceptable, except near termination when \underline{g}^k is changing by small amounts and the gradients are small. Here there can be an adverse effect on convergence.

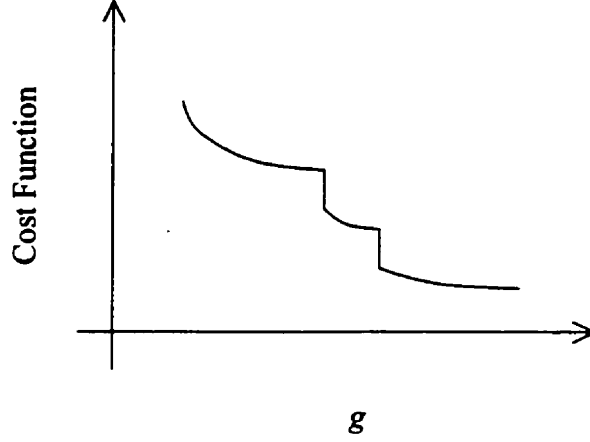


Figure 5.5: Discontinuous cost function close to an optimum.

A certain level of continuity near local optima can be preserved by avoiding re-meshing; instead all nodes affected by that change are dragged in the appropriate direction by the amount, $\Delta \underline{g}^{k+1}$.

Forcing the same mesh configuration for large changes in geometry can create poorly shaped elements. For that reason, dragging the nodes replaces re-meshing only when the maximum change in a parameter is less than 10%, i.e. when

$$100 \times \max \left(\frac{g_1^{k+1} - g_1^k}{g_1^k}, \frac{g_2^{k+1} - g_2^k}{g_2^k}, \dots, \frac{g_V^{k+1} - g_V^k}{g_V^k} \right) < 10 \quad (5.15)$$

where V is the total number of geometric parameters, \underline{g}^k is the geometry from the last major step (which generates the search direction) and \underline{g}^{k+1} is the geometry updated at either the next line search or major step.

5.4.3 Generalized Termination Criterion

An interesting observation from optimizing microwave devices is that the optimizer sometimes tends to perform many line search CFEs close to the optimum – thus increasing the computational cost of the overall optimization considerably and reducing

the cost function very little. For the test cases solved in this thesis, the problem of excessive line search CFEs was resolved by generalizing the termination criterion to allow the optimization process to terminate *within* a line search. While this may mean an actual increase in C , the increase in cost function is usually negligible and the process successfully converges.

In order to terminate the optimization during a line search, values of $\|\nabla L(\underline{g}^k)\|_2$ are needed for every CFE of the k^{th} line search step. While each CFE provides the gradients of the cost function, the Lagrange multiplier estimates, $\underline{\lambda}^k$, are needed to fully compute $\|\nabla L(\underline{g}^k)\|_2$. The MATLAB optimizer provides $\underline{\lambda}^k$ for only main steps of the optimization. Expression (4.7) was implemented to calculate $\underline{\lambda}^k$ for any CFE. The programming implications were trivial and added little to the computational expense.

5.4.4 Accuracy of CFEs during Line Searches

The accuracy of a CFE was only changed at main optimization steps. Line search CFEs were kept at a constant level of accuracy as far as possible. If the CFEs from the line search were of different accuracy than that of the main step, the MATLAB optimizer had trouble converging to an optimum.

For every main step of the optimization, the Hessian is updated and a search direction is generated. The subsequent line search depends on that search direction to point roughly towards the optimum of the cost function. Varying the accuracy of line search CFEs changes the target of the bisection algorithm (the line search method used). Attempts to reduce the cost function in such cases can lead to a large number of line search steps and possible non-convergence.

5.4.5 Smoothing out Sharp Changes in Accuracy

It is possible for the value of $\|\nabla L(\underline{g}^k)\|_2$ to reduce in magnitude sharply as an optimum is approached. Since this can result in a change of accuracy levels by a great

amount, a weighted-average is taken to smooth out (to an extent) the potential discontinuity between design spaces. Expression (5.5) becomes

$$N_f \left\| \text{estimate of error in } \nabla c_a(\underline{g}^k) \right\|_2 < \alpha \left(w \left\| \nabla L(\underline{g}^{k+1}) \right\|_2 + (1-w) \left\| \nabla L(\underline{g}^{k+2}) \right\|_2 \right) \quad (5.16)$$

where the weight w is intended as a “safety” measure to avoid drastic increases in accuracy and was chosen to be 0.9.

5.4.6 Multiple Local Optima

The graph in Figure 5.1 plots a hypothetical cost function at three different accuracy levels versus a geometric parameter. Although g^{opt_1} (the minimum of the least accurate of the cost functions) is not really a minimum at all for curve $C^{(3)}$, if an optimization were started at this point at an accuracy level equivalent to the third cost function, g^{opt_1} would eventually be found as the local minimum. This is only the case if the curve $C^{(3)}$ is convex and quadratic, thus containing only one, global, minimum. In problems with multiple optima, it was observed that varying the accuracy levels during optimization might lead to finding different local optima. Since using a gradient based optimizer can only guarantee quick convergence to a local optimum, this characteristic is not of great concern – due to the equal likelihood of finding a better optimum than a worse one.

5.4.7 Generating Suitable Meshes for the Error Estimator

At each step, the mesh generator creates an initial mesh based on \underline{g}^k . The mesh generator used does not have the flexibility to guarantee an initial mesh discretized with a specified number of internal nodes. For this reason, an initial mesh may contain few or no internal nodes – thus hampering the effectiveness of the error estimator.

To overcome this problem, in cases where fewer than $N_i = 5$ internal nodes are generated, a uniform refinement of the mesh is applied, dividing each element into four new triangles – thus creating four times the number of total elements and more internal

nodes. Uniform refinements are performed until $N_i \geq 5$. A uniform refinement of the initial mesh is sometimes overkill – creating more elements than are needed for p -adaption. However, the error estimator worked relatively well in ensuring that in CFEs with big differences in numbers of elements, if the accuracy demanded was the same in both, a similar number of DOFs was used to solve the problem. The p -adaption in the denser mesh requires fewer adaptive steps (i.e., a lower overall order for each element) to reach the desired error tolerance than in the mesh with fewer elements.

Chapter 6

Validation

All computer programs used to implement the theory described in this thesis were written in either Fortran 90 [111] or in MATLAB [112]. All codes were implemented on a Windows NT 4.0 [113] platform with an Intel Celeron 400 MHz processor, 128 Mb RAM, and 6.4 Gb hard disk. Implementation of the optimization-adaption system of Figure 5.3 requires four independent blocks of computer code necessary for testing the theory from Chapters 2 to 5: optimization, parameterization, mesh generation and adaptive FE solution.

The commercial constrained optimizer from the MATLAB Optimization Toolbox [102], *constr.m*, was slightly modified to include features needed to implement the accuracy link for the optimization-adaption system. The optimization is executed from a MATLAB command window while the rest of the blocks used in the optimization-adaption system are run as executable programs outside the MATLAB environment. For this reason, all data between programs is transferred by ASCII files that can be read by both MATLAB programs and Fortran executables.

The termination criterion given in (4.31) was added to the optimization code and used instead of the optimizer's default termination criteria. The accuracy link was implemented within the actual optimization program by computing $\frac{\alpha}{N_f} \|\nabla L(\underline{g}^k)\|_2$ and writing it to an output file – later to be used as the error tolerance for a CFE. A small MATLAB subroutine was written in order to allow calculation of $\|\nabla L(\underline{g}^k)\|_2$ at any optimization step (including line searches). The constant value of α is hard-coded at the beginning of *constr.m*. The optimization code (in general) requires two user-defined subroutines necessary for calculating the cost function and its gradient: *fun.m* and *grad.m*. Subroutine *fun.m* generally evaluates expressions for $C(\underline{g}^k)$ and each constraint, $h_i(\underline{g}^k)$,

while *grad.m* computes $\nabla C(\underline{g}^k)$ and partial derivatives of the constraints (a matrix whose ij^{th} entry is $\frac{dh_j}{dg_i}$). All constraints are problem dependent and their partial derivatives can be performed analytically. A CFE is performed from *fun.m* by invoking a Windows batch file that executes programs that parameterize the input geometry \underline{g}^k , generate the mesh and solve the FE problem adaptively. The geometric parameters are transferred to the CFE via data file, as are the return values of $C(\underline{g}^k)$ and $\nabla C(\underline{g}^k)$. Both the cost function and gradient are not actually calculated by MATLAB code, but are read in as data – written by the FE program.

The parameterization of \underline{g}^k varies from problem to problem and is performed through a separate, Fortran-written executable. The parameterization program takes as input \underline{g}^k and outputs a file full of co-ordinate information, which is used as input to the mesh generator. Each optimization problem requires its own parameterization program.

A mesh is represented in a data file with information about elements, nodes and their x-y co-ordinates. An additional, separate file contains the derivatives of the nodes with respect to the user-defined parameters, dx/dg_i and dy/dg_i . The mesh generator (written in Fortran) creates an initial mesh, which can be manually or automatically refined by subdividing each element into four new triangles, and writes all pertinent information to the two output data files.

A small Fortran program was created to drag the nodes (as an alternative to parameterization and re-meshing, described in section 5.4.2) of an existing mesh by an amount $\Delta \underline{g}$ in the correct direction. The dragging program requires information about the mesh, its derivatives and the amount by which the nodes should be dragged. The program only modifies the co-ordinates of the mesh.

The entire adaptive FE solution was coded in Fortran and available as an executable file (for use in the batch file, called from MATLAB). The FE code must be slightly altered for different types of cost functions, but is general for computing the scattering matrix of any H -plane, N -port, rectangular waveguide problems. For each adaptive step, the sparse matrix equation is solved by the frontal method [114].

6.1 Choice of Measures of Computational Costs

In a typical FE method, most of the computational effort lies in solving the matrix equation. The size of the matrix equation depends on the number of DOFs (n) for that FE solution. While the number of DOFs is a good measure of the size of a particular FE problem, a more interesting quantity is a measure of the cost of solving the problem, i.e. the computational cost. In 2-D, the computational cost of solving a matrix equation by the frontal method is $O(n^2)$, so this is taken as approximately the total computational cost of an FE solution.

In comparing the costs of optimizing different microwave devices using the system from Figure 5.3, cumulative computational costs are used rather than clock speeds. Ideally, both the optimization routine and the adaptive FE CFE would be written in the same language. Combining the two through batch files in the Windows environment caused some difficulty that was resolved through the use of “pause” statements in the MATLAB code. The computational cost of the optimization program, parameterization, and meshing is negligible compared to the cost of performing a CFE. For that reason, the cumulative computational cost (CCC) of any CFE is based on the cumulative cost of solving matrix equations. In addition, the overall cost of an optimization is taken to be the sum of the CCCs of each CFE.

For measuring the computational cost of a CFE (which involves multiple adaptive steps and a frequency sweep), a cumulative measure of cost must be defined. An N -port device requires N FE solutions per frequency. In the p -adaptive FE solution (given in Figure 3.5), for N_a adaptive steps taken to converge the solution at the center frequency and a frequency sweep (requiring an additional $N_f - 1$ FE solutions) for an N -port device, the measure of CCC is assigned to be:

$$\text{Cumulative Computational Cost} = N \left(\sum_{i=1}^{N_a} n^2(i) + (N_f - 1) n^2(N_a) \right) \quad (6.1)$$

6.2 Test Cases

Four different test cases are used to validate the theory developed in the previous chapters. Each test case is an H -plane rectangular waveguide component excited by its dominant TE_{10} mode. All operating frequency ranges and guide dimensions are selected to propagate only the dominant TE_{10} mode. Conductor losses due to induced current in the walls of the waveguides are ignored and all test cases are assumed to be lossless and source free. All guide walls are subject to the Dirichlet boundary conditions given by (2.50) while planes of symmetry are handled by the Neumann boundary condition, given in (2.51). Electromagnetic fields at ports are modeled using the port boundary condition in (2.52). The port element allows various lengths of waveguide to be attached to the ports. Except for the results that compare the convergence of p -adaption for different lengths, a length of $l = 0.1\lambda$ is used.

All test cases contain materials that are isotropic – thus the scattering matrix is symmetric in each case, i.e. $S_{ji} = S_{ij}$.

While the MATLAB optimizer can handle many different forms of constrained problems, all optimizations in this chapter minimize a chosen cost function subject to inequality constraints only.

6.2.1 Length of Uniform Rectangular Waveguide

The first test case is a simple length of air-filled uniform rectangular waveguide. It is used in only one section of the results – to validate gradient computations of the cost function (because they can be calculated analytically).

The propagation constant β for the TE_{10} mode is:

$$\beta = \sqrt{k_0^2 - \left(\frac{\pi}{a}\right)^2} \quad (6.2)$$

where k_0 is the wavenumber and a is the width of the broader wall of the waveguide (or port width). For a length of waveguide with a port at each end, $S_{11} = 0$ and S_{12} is given by

$$S_{12} = e^{-j\beta l} \quad (6.3)$$

where l is the length of the line. The phase of S_{12} is:

$$\angle S_{12} = -\beta l \quad (6.4)$$

and thus the derivative of the phase of S_{12} with respect to length is simply:

$$\frac{d\angle S_{12}}{dl} = -\beta \quad (6.5)$$

The problem is normalized to have a port width of $a = 1$ m and the length of the device is 2 m. The wavenumber range is $1.1\pi < k_0 < 1.9\pi$ (where the cutoff wavenumber is $k_c = \pi$), safely within propagating range of the dominant TE_{10} mode. The geometry and dimensions of the length of waveguide in the H -plane and the mesh generated for the problem are given below, in Figure 6.1.

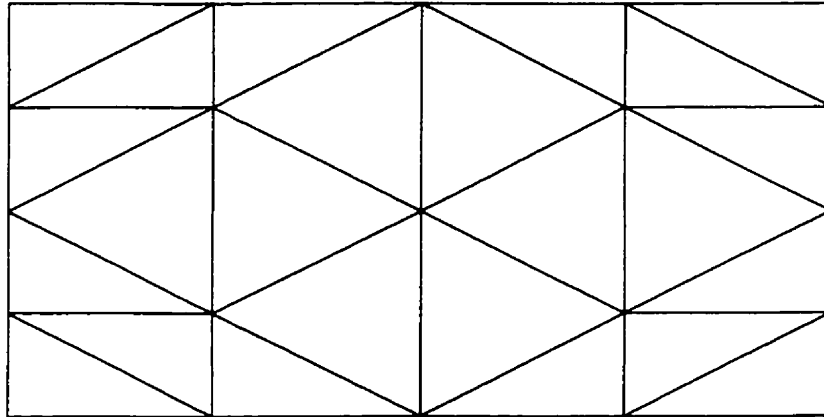
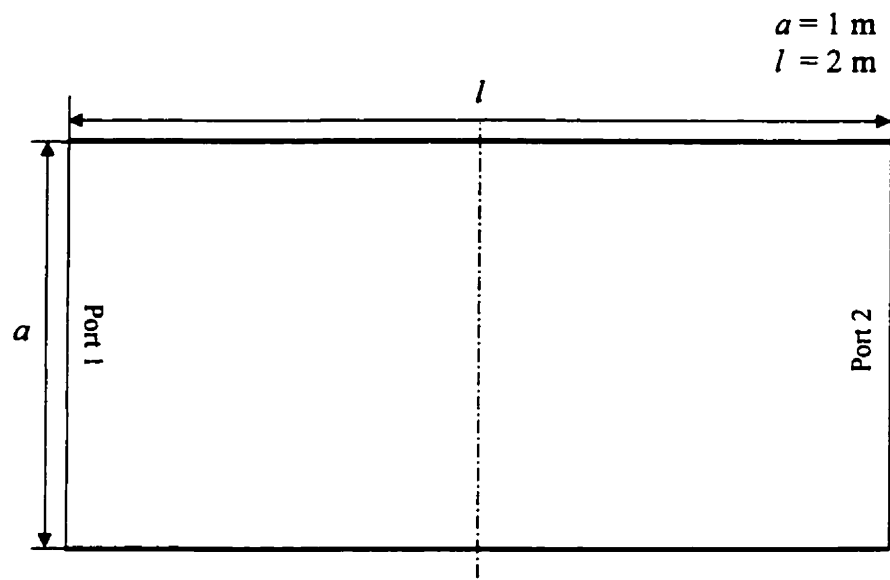


Figure 6.1: Length of uniform rectangular waveguide and its mesh.

6.2.2 Waveguide T-Junction with Inductive Post

Consider the problem of maximizing transmission in a waveguide T-junction with the use of a symmetrical, inductive septum (or post), shown in Figure 6.2. A waveguide T-junction has two functions: equal power division of the incoming wave at port 1 and a 90° change of direction. The interest in such a device is to minimize the loss of power transmitted from port 1 to ports 2 and 3. Since losses due to heat are minimal and ignored, the main reason for loss is the reflection of the wave. Loss can be minimized by inserting an inductive post to help the incident wave split and change direction by 90° (i.e. maximizing transmission to ports 2 and 3) rather than reflect back to port 1. The inductive post is an effective feature in dividing power and optimizing transmission [115,87]. The post chosen for the problem is rectangular and symmetric – guaranteeing an even split of power to ports 2 and 3.

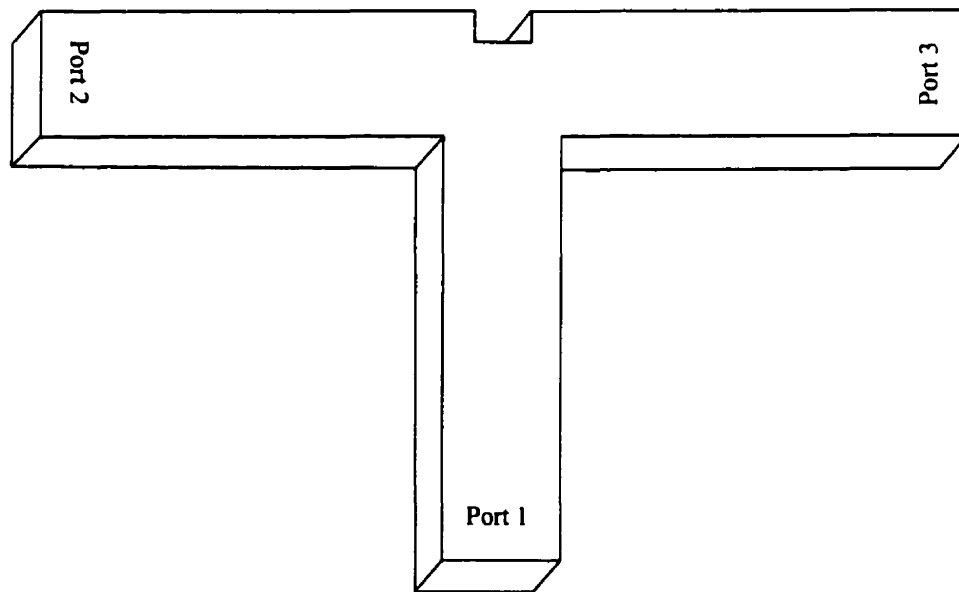


Figure 6.2: Waveguide T-Junction with inductive post.

In order to maximize transmission to port 2 (and/or 3) over a range of frequencies, the reflection at port 1 must be minimized. The cost function for the problem is defined as:

$$C(\underline{g}) = \sum_{i=1}^{N_f} c_i \quad (6.6)$$

where the optimization parameters \underline{g} are the length and width of the post (see Figure 6.3).

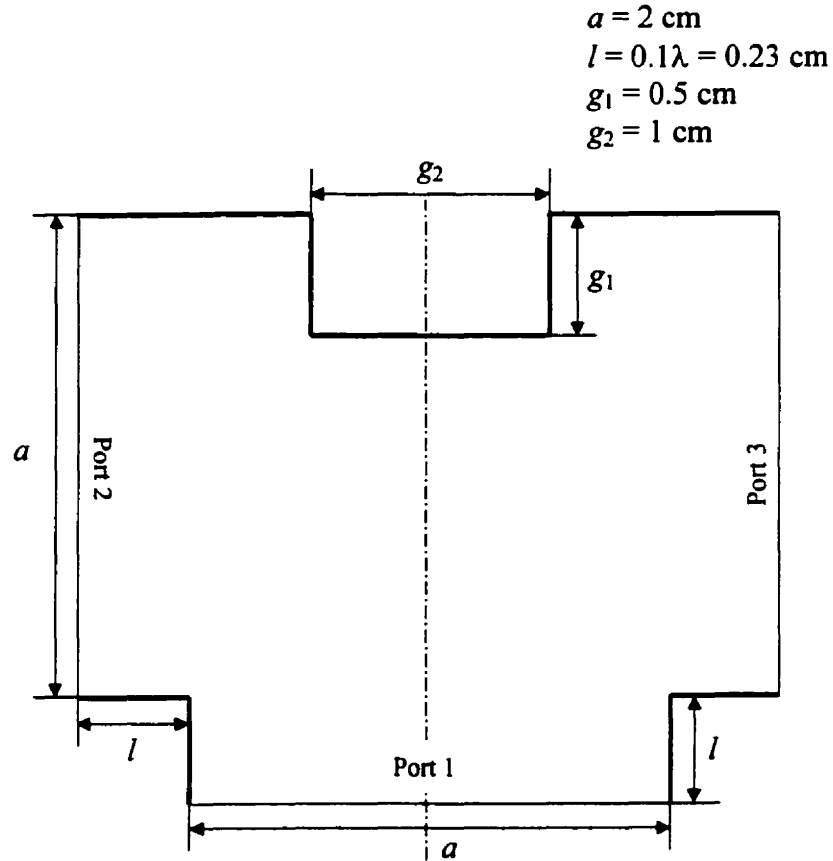


Figure 6.3: Initial geometry and dimensions of the waveguide T-junction.

The single frequency cost function is the coefficient of reflection, i.e.

$$c_i = |S_{11}(f_i)|^2 \quad (6.7)$$

The frequency range over which the reflection is minimized is:

$$8 \text{ GHz} < f < 12 \text{ GHz} \quad (6.8)$$

with a center frequency of $f_a = 10 \text{ GHz}$. The number of discrete points sampled by a frequency sweep of a CFE is $N_f = 5$. The wavelength of the device at its center frequency is calculated to be $\lambda_a = 4.53557 \text{ cm}$.

The bounds for the two geometric parameters are chosen to be:

$$\begin{aligned} 0.1 \text{ cm} < g_1 < 1.9 \text{ cm} \\ 0.1 \text{ cm} < g_2 < 2.70712 \text{ cm} \end{aligned} \quad (6.9)$$

The choice of the limits in (6.9) is, of course, subjective. For any problem, it is important to choose constraints that limit the geometric parameters to a range where any geometry of the device within that range is physically possible. A poorly constrained optimization problem (or an unconstrained one) may allow edges or boundaries to overlap – making it physically impossible to build and a useless design. A well-constrained optimization problem will converge more quickly than a poorly constrained or unconstrained problem. Although the constraints for the inductive post are quite simple, the choice of upper and lower limits must be made. An additional concern in the optimization of a device is meshing. While a particular geometry may be feasible (in a physical sense), meshing difficulties can easily arise when modeling the device. Edges too close together or too small in themselves may force the generation of many small elements – which may not be necessary. To prevent this, a minimum “cushion” of space was left between any two boundaries – thus affecting the choice of bounds for the parameters. For the T-junction meshes, a cushion of 0.1 cm worked well. The lower limits for both parameters are taken

to be 0.1 cm. The upper limit for g_1 (1.9 cm) was chosen because at $g_1 = 1.9$ cm, the bottom edge of the post is 0.1 cm shorter than the port width of 2 cm. The upper limit of g_2 was chosen because when g_2 is at its upper limit, the left and right edges of the post are 0.1 cm away from ports 2 and 3, respectively. Four inequality constraints, of the form $h_i(\underline{g}) \leq 0$, based on those bounds are:

$$\begin{aligned} h_1(\underline{g}) &= g_1 - 1.9 \\ h_2(\underline{g}) &= 0.1 - g_1 \\ h_3(\underline{g}) &= g_2 - 2.70712 \\ h_4(\underline{g}) &= 0.1 - g_1 \end{aligned} \tag{6.10}$$

The starting point of the optimization is arbitrarily taken to be $\underline{g} = [0.5 \ 1]^T$ cm. The mesh of the initial geometry of the T-junction is given in Figure 6.4.

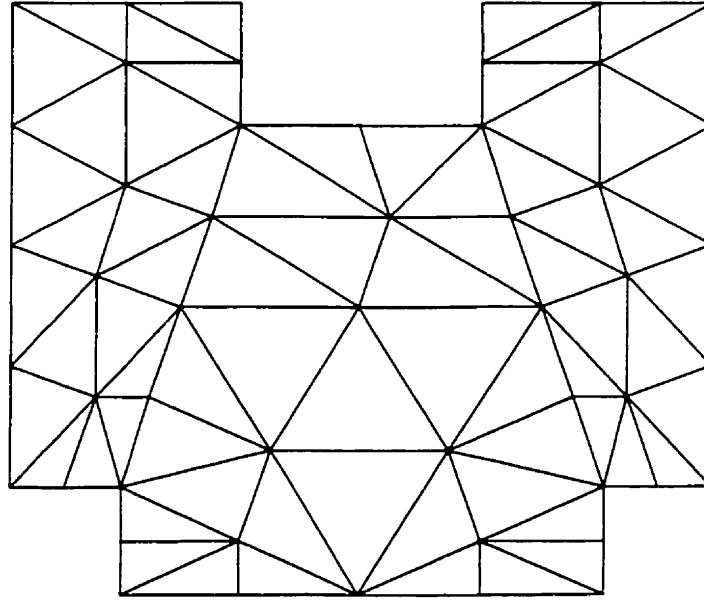


Figure 6.4: Mesh of the initial geometry of the waveguide T-junction.

The mesh generated for the initial geometry contains 17 elements. However, this mesh does not reach the criterion for the minimum number of internal nodes necessary for error estimation. The mesh in Figure 6.4 is the result of one uniform mesh refinement, yielding 68 triangular elements and a good discretization of internal nodes. In the Results section (section 6.3), the 17 element mesh was used for most of the results except for testing the error estimator, where the 68 element mesh was used. The initial mesh of the optimization was the one of 68 elements, shown, because every CFE underwent the scrutiny of meeting the minimum generated 5 internal nodes.

Figure 6.3 shows a plane of symmetry for the waveguide T-junction. While computational cost can be reduced by exploiting this plane of symmetry, practically, the number of elements in the mesh is quite small. Solving the problem in 2-D is quite speedy (even with higher order basis functions) so the decision was made to model the whole T-junction instead of half.

6.2.3 Miter Bend with Dielectric Column

A right-angled bend is a common microwave component that changes the direction of microwave energy by 90° . The interest in such a device is again to minimize the loss of power from one port to another. Once again, the loss is attributed to reflection of the wave. In order to create a smoother corner, the corner is cut at 45° to create a mitered right-angle bend [116]. The 45° chamfer has a mirror like effect and reflects the incoming wave around the corner. In addition, performance can be substantially improved by adding a square column of dielectric. The H -plane device is shown in Figure 6.5.

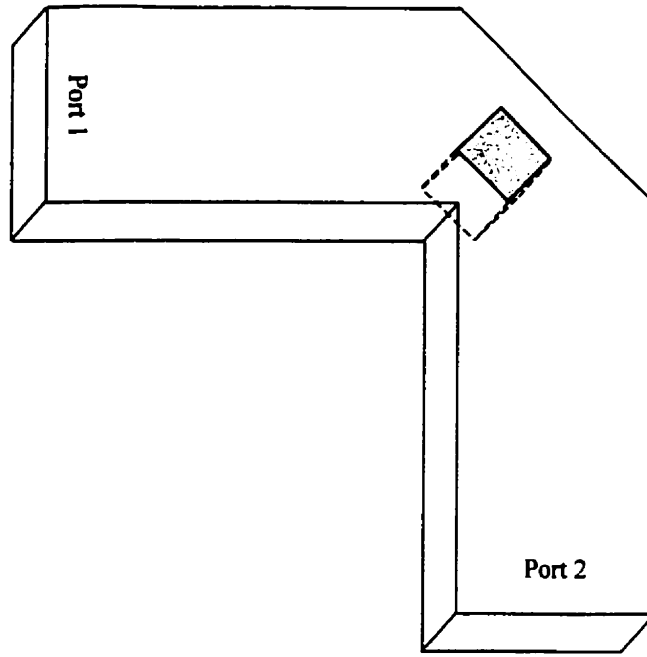


Figure 6.5: Miter bend with dielectric column.

In order to minimize the reflection, the same cost function is used as in the T-junction problem (defined in (6.6) and (6.7)). However, in this case, the cost function varies with four geometric parameters: length of the chamfer (g_1), dimensions of the dielectric block (g_2 and g_3) and distance from the corner to the center of the column (g_4). The parameters and the dimensions of the initial geometry are given in Figure 6.6.

The frequency range is also chosen to be $8 \text{ GHz} < f < 12 \text{ GHz}$ with a center frequency of 10 GHz and wavelength of $\lambda_a = 4.53557 \text{ cm}$. The number of points in a CFE frequency sweep is $N_f = 5$. The relative permittivity of the dielectric block is $\epsilon_r = 2.1$.

Creating constraints for the miter bend with dielectric column (or block) is quite complex. Varying the size and position of the dielectric block can result in many collisions between edges and boundaries of the model. The fact that the dielectric block is not necessarily square increases the complexity of choosing proper constraints. The

constraints that were “derived” ensure that a “cushion” of δ is kept between any two edges or corners (parameterized by \underline{g}).

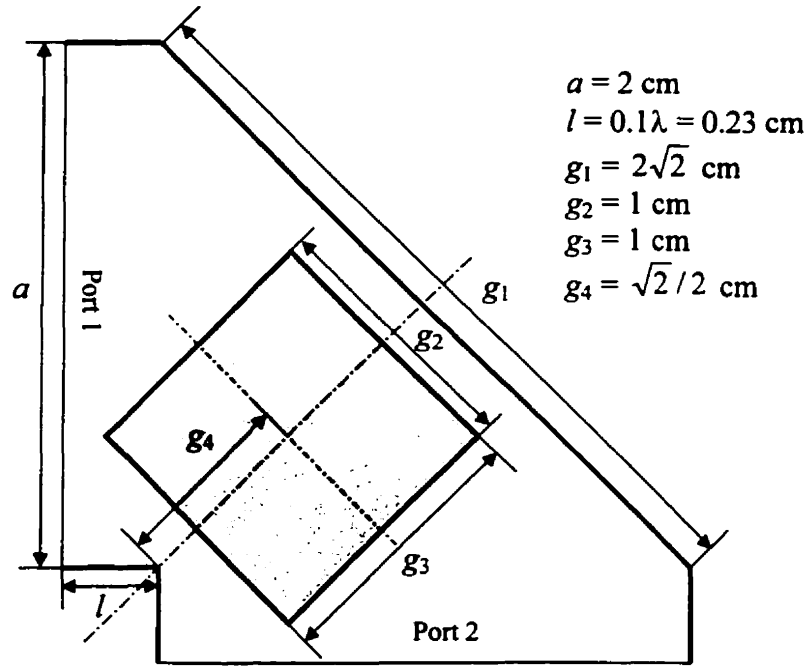


Figure 6.6: Initial geometry and dimensions of the miter bend.

The nine inequality constraints for the problem are:

$$\begin{aligned}
 h_1(\underline{g}) &= g_1 - \sqrt{2}(1 + l) + \delta \\
 h_2(\underline{g}) &= \delta - g_1 \\
 h_3(\underline{g}) &= \delta - g_2 \\
 h_4(\underline{g}) &= \delta - g_3 \\
 h_5(\underline{g}) &= 2\delta - g_4 \\
 h_6(\underline{g}) &= g_1 + g_3 + 2g_4 + 2(\delta - \sqrt{2}) \\
 h_7(\underline{g}) &= g_3 - 2g_4 + 2\delta \\
 h_8(\underline{g}) &= g_2 + g_3 - 2g_4 + 2\sqrt{2}(-l + \delta) \\
 h_9(\underline{g}) &= g_2 + g_3 + 2g_4 - 2\sqrt{2}(1 - \delta)
 \end{aligned} \tag{6.11}$$

where $\delta = 0.1$ cm and $l = 0.1\lambda = 0.453557$ cm.

The starting point of the optimization is arbitrarily taken to be $\underline{g} = [2\sqrt{2} \ 1 \ 1 \ \sqrt{2}/2]^T$ cm. This initial geometry centers the square dielectric block (of dimension 1 cm \times 1 cm) in the middle of the device. In both the miter bend and T-junction problems, the entire frequency range is “all-pass” and there is no danger of choosing a bad starting point that would lead to total reflection (which can happen in a bandpass filter). However, the miter bend problem has more than one local optimum, and starting the optimization at a point where a smaller block is centered, an interesting different local optimum is obtained (see Appendix B). This alternate starting point is $\underline{g}^* = [2\sqrt{2} \ \sqrt{2}/2 \ \sqrt{2}/2 \ \sqrt{2}/2]^T$ cm. Many of the results in section 6.3 use a mesh based on this alternate starting point (arbitrarily). However, while starting an optimization from this point yields an interesting optimum (shown in Appendix B), the results for the accuracy linked optimization-adaption system in section 6.3 are based on the former starting point.

The mesh generated for the initial geometry contains 16 elements. However, the initial mesh did not have a single internal node. A uniform refinement of all elements is necessary for estimating errors in ∇c_a . The refined, 64 element mesh of the initial geometry is shown in Figure 6.7.

Once again, there is a plane of symmetry (the diagonal cutting through the center of the block, in Figure 6.6) that it was not thought necessary to exploit due to the relatively small number of elements.

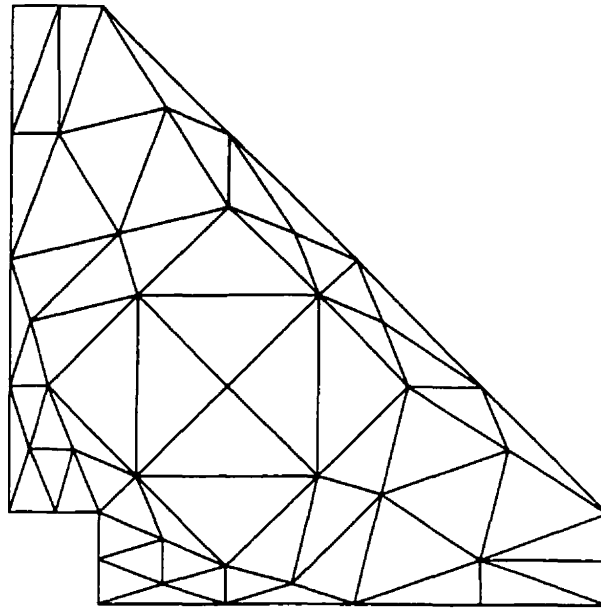


Figure 6.7: Mesh of the initial geometry of the miter bend with dielectric block.

6.2.4 Two – Cavity, Iris Coupled, Waveguide Filter

Consider the problem of designing a waveguide device to have the characteristics of a two pole, Chebyshev bandpass filter. The design and optimization of iris coupled, waveguide cavities using CAD tools is common in achieving bandpass filter characteristics [117–119,32]. Filters of this type have resonating cavities (where each guide cavity is roughly $\lambda/2$ in length), coupled by thin (or thick) irises with coupling apertures between any two cavities. While some designs make use of varying the thickness of each iris as design parameters [117,119], the 2-cavity filter described below has irises of constant, small, thickness (see Figure 6.8).

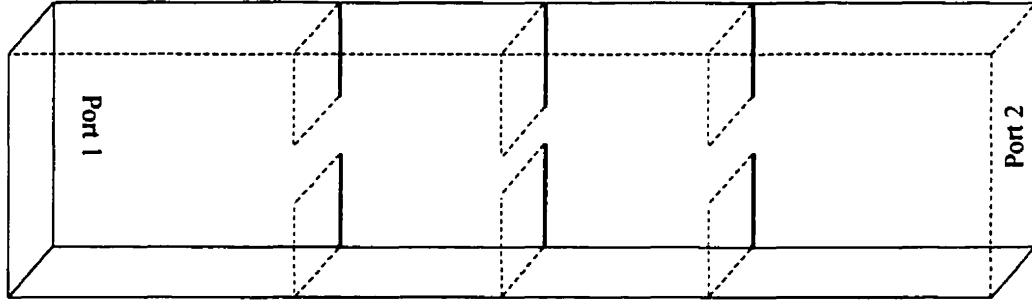


Figure 6.8: 2-Cavity, iris coupled, waveguide filter.

The characteristics of a bandpass filter are to maximize transmission over a band of frequencies and stop transmission outside that band. While filter design specifications are commonly given by pass band and stop band attenuation, a slightly different approach is used here: minimize reflection in the pass band and minimize transmission in the stop band. The cost function for the problem is:

$$C(\underline{g}) = \sum_{i=1}^{N_f^p} c_i^p + \sum_{i=1}^{N_f^s} c_i^s \quad (6.12)$$

where there are N_f^p discrete frequency points sampled in the pass band and N_f^s points in the stop band. The single frequency cost functions (in the summation terms of (6.12)) are given by:

$$\begin{aligned} c_i^p &= |S_{11}(f_i^p)|^2 \\ c_i^s &= |S_{12}(f_i^s)|^2 \end{aligned} \quad (6.13)$$

The design problem has three varying geometric parameters. The aperture width of the left-most and right-most irises (symmetrical) is g_1 . The aperture width of the center iris is

g_2 . The lengths of the two, symmetrical cavities is g_3 . The geometric parameters and dimensions are illustrated for a quarter of the problem in Figure 6.9.

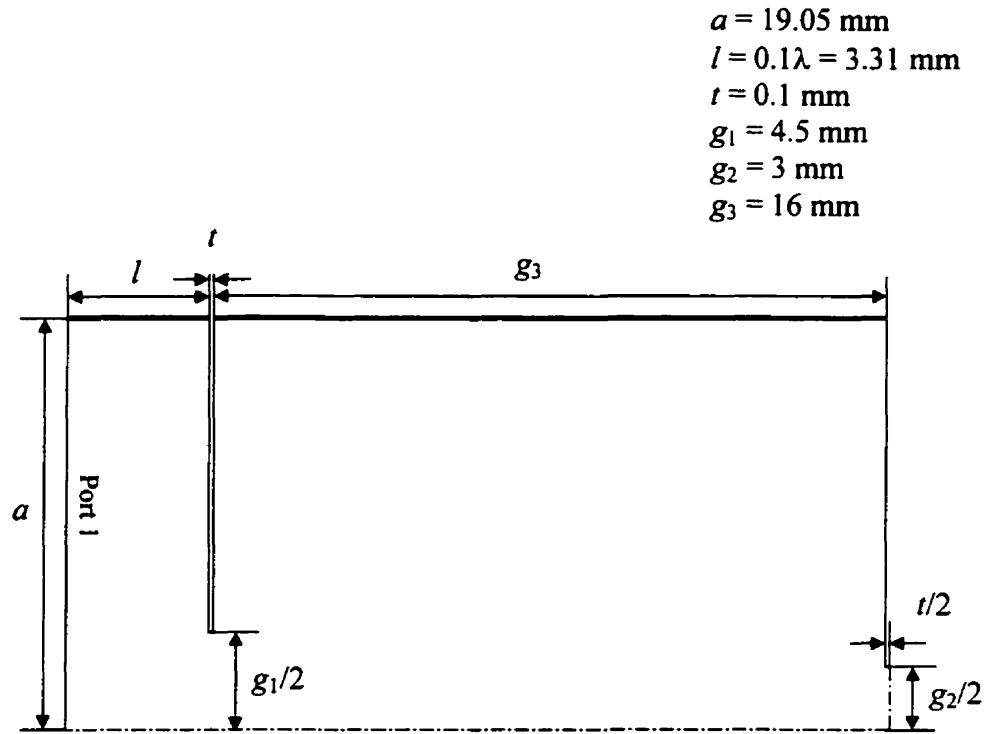


Figure 6.9: Initial geometry and dimensions of the 2-cavity filter.

The frequency range is taken to be between 11.8 GHz and 12.2 GHz with a pass band (of 100 MHz bandwidth) centered at $f_a = 12 \text{ GHz}$. The pass and stop bands are therefore

$$\text{Pass band : } 11.95 \text{ GHz} < f < 12.05 \text{ GHz}$$

$$\begin{aligned} \text{Stop bands : } & 11.80 \text{ GHz} < f < 11.95 \text{ GHz} \\ & 12.05 \text{ GHz} < f < 12.20 \text{ GHz} \end{aligned}$$

(6.14)

The total number of frequency points sampled for a CFE is $N_f = 21$, with $N_f^p = 15$ points in the pass band and $N_f^s = 6$ points in the stop band. A greater number of points are taken in the pass band because in the optimal design of the filter, both cavity resonances are within this range of frequency. The guide wavelength at 12 GHz is $\lambda_g = 33.12948$ mm.

Choosing constraints for the geometric parameters is quite simple because as long as the parameters are positive quantities, there cannot be any overlapping of edges. The bounds for each parameter are loosely chosen as:

$$\begin{aligned} 2 \text{ mm} < g_1 < 10 \text{ mm} \\ 2 \text{ mm} < g_2 < 10 \text{ mm} \\ 10 \text{ mm} < g_3 < 20 \text{ mm} \end{aligned} \tag{6.15}$$

and six inequality constraints can be written as:

$$\begin{aligned} h_1(\underline{g}) &= g_1 - 10 \\ h_2(\underline{g}) &= 2 - g_1 \\ h_3(\underline{g}) &= g_2 - 10 \\ h_4(\underline{g}) &= 2 - g_2 \\ h_5(\underline{g}) &= g_3 - 20 \\ h_6(\underline{g}) &= 10 - g_3 \end{aligned} \tag{6.16}$$

The thickness of the irises was taken to be 0.1 mm.

The choice of initial geometry is more difficult in the bandpass filter than in both the T-junction and miter bend test cases. To optimize the filter effectively, both resonant frequencies must be within the pass band frequency range (also meaning they are close together because of a relatively narrow bandwidth). A poor initial choice for \underline{g} might have resonant frequencies that are far apart and far from the pass band. In addition, a resonant cavity frequency might be outside the entire range of frequencies sampled (outside the stop band, as well). Using a frequency sweep for such a poor design will not

work because there will be no detection of the presence of the cavity resonance in the sampled frequency range. However, microwave filter design is a good example of where an initial geometry is found by an initial design method [9] and “tuned” by the optimizer to enhance performance. The method of filter design from [9] is used to design a filter to have the frequency response within a desired initial range (roughly). The initial geometry used is $\underline{g} = [4.5 \ 3 \ 16]^T$ mm.

There are two planes of symmetry in the filter model. Since the thickness of each iris is very small, the mesh generated for any geometry will contain many elements. While the filter can be modeled with a mesh of fewer elements (by modeling an iris as infinitely thin), it was found that for adequately accurate results, a fine discretization around the iris apertures is necessary. Therefore, both planes of symmetry can be used and in the results of the optimization, only a 1/4 of the problem is modeled. The mesh for the initial geometry is given in Figures 6.10 and 6.11. Due to a large number of internal nodes, there is no need for further refinement.

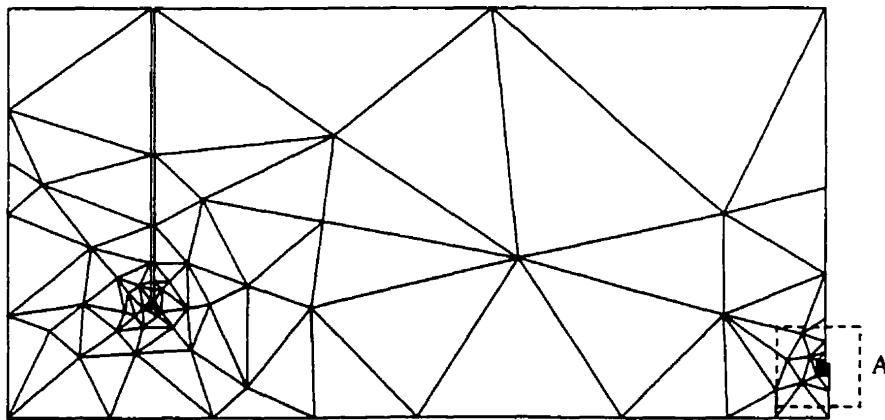


Figure 6.10: Mesh of the initial geometry of the 2-cavity filter.

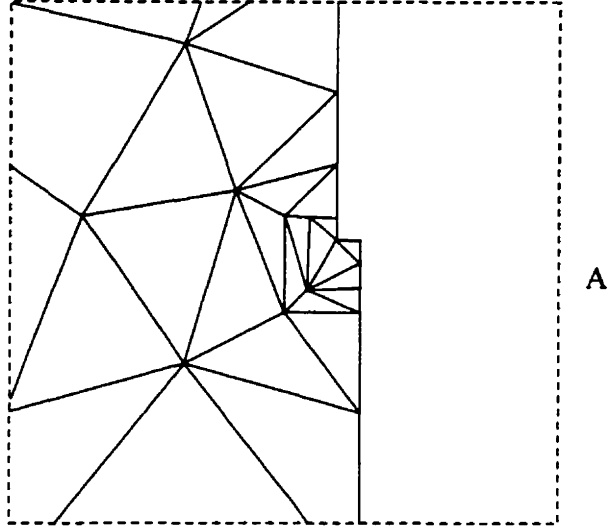


Figure 6.11: Zoom of region A from the mesh in Figure 6.10.

6.3 Numerical Results

The combination of optimization and FE p -adaption requires each component within the system to function properly. In addition to presenting the results of the optimization-adaption system, this section first validates the different aspects of the adaptive FE process on the pre-defined test cases. Each aspect of the adaptive process is tested on the different problems for several cost functions and geometric parameters. All adaption is performed at a single frequency, within a test case's defined range.

6.3.1 Error Indication and Port Elements for P -Adaptive Cost Function Calculation

The T-junction, miter bend and 2-cavity filter are analyzed three times, with different lengths of waveguide connected to the ports. With 1.5λ of guide length attached to each junction, all evanescent modes decay to very small amplitudes by the time they reach the ports. In this case, the ports have virtually nothing to absorb and they can alternatively be modeled by a constrained-field boundary condition. However, when the

length of guide is 0.1λ , the fields at the ports are no longer pure TE_{10} and the port element must be used to absorb evanescent waves.

In order to show the convergence of a p -adaptive scheme for each problem, the Targeted Error Indicator (TEI) from section 3.4.2 is used to guide the adaption, based on a defined cost function. To reduce the error, 25% of the elements (those with the highest errors in the cost function, determined by the TEI) are increased in order by one, at each adaptive step. The order of each element is initially 2. The error at an adaptive step is defined as the difference between current value of the cost function, c , and a reference value, i.e. $|c - c_{conv}^{1.5\lambda}|$. The reference value ($c_{conv}^{1.5\lambda}$) is obtained by converging the value of c in the 1.5λ case by repeatedly refining the mesh (dividing each element into four new elements) with all elements at 10^{th} order.

Figures 6.12 through 6.14 show the convergence of p -adaptive schemes for the three

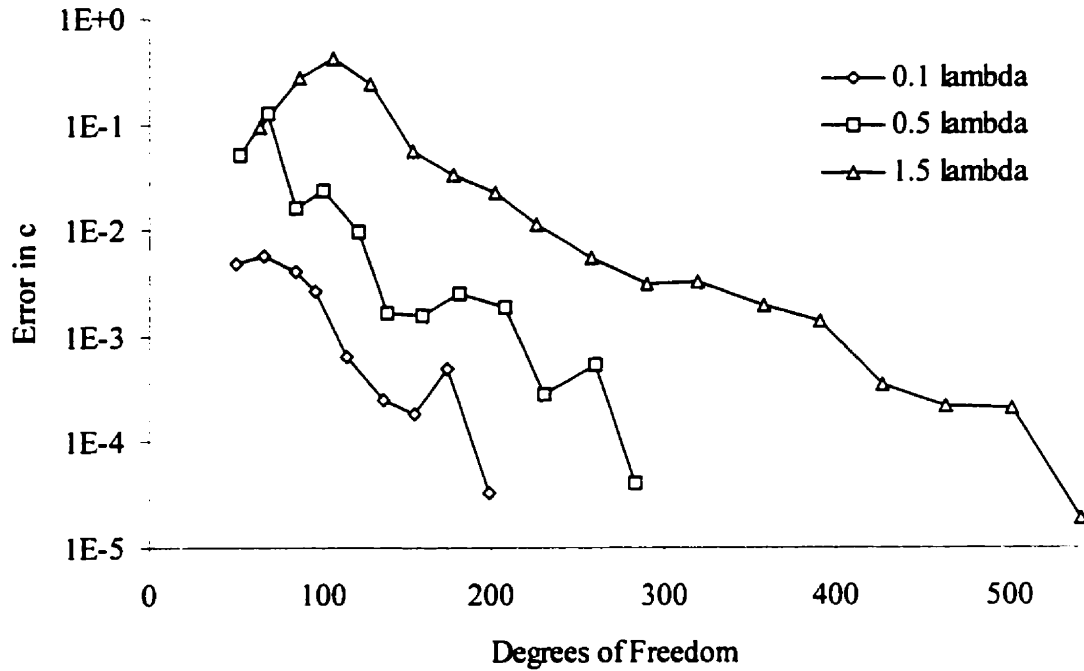


Figure 6.12: Convergence of p -adaption for the T-junction, with ports placed three different distances away. The cost function is $c = |S_{12}|^2 + |S_{13}|^2$; frequency is $f = 10 \text{ GHz}$ and geometry is $\underline{g} = [0.5 \ 1]^T \text{ cm}$.

problems with guide lengths of 1.5λ , 0.5λ and 0.1λ .

The cost function used for the T-junction is the transmission coefficient from port 1 to ports 2 and 3: $c = |S_{12}|^2 + |S_{13}|^2$. The meshes for both 0.5λ and 0.1λ cases have 17 elements while the mesh based on 1.5λ guide length has 21 elements. The solution was adapted at a frequency of 10 GHz and the T-junction model used has the geometry $\underline{g} = [0.5 \ 1]^T$ cm. The adaption was terminated when the error in the cost function was below 10^{-4} (the error tolerance).

The miter bend cost function is chosen to be the coefficient of reflection at port 1, or $c = |S_{11}|^2$. The mesh with a guide length of 1.5λ has 28 elements; the mesh for the 0.5λ case has 30 elements and the 0.1λ case has 16 elements. The test frequency is $f = 11.25$ GHz with the geometry $\underline{g} = [2\sqrt{2} \ \sqrt{2}/2 \ \sqrt{2}/2 \ \sqrt{2}/2]^T$ cm. The error tolerance is 10^{-4} .

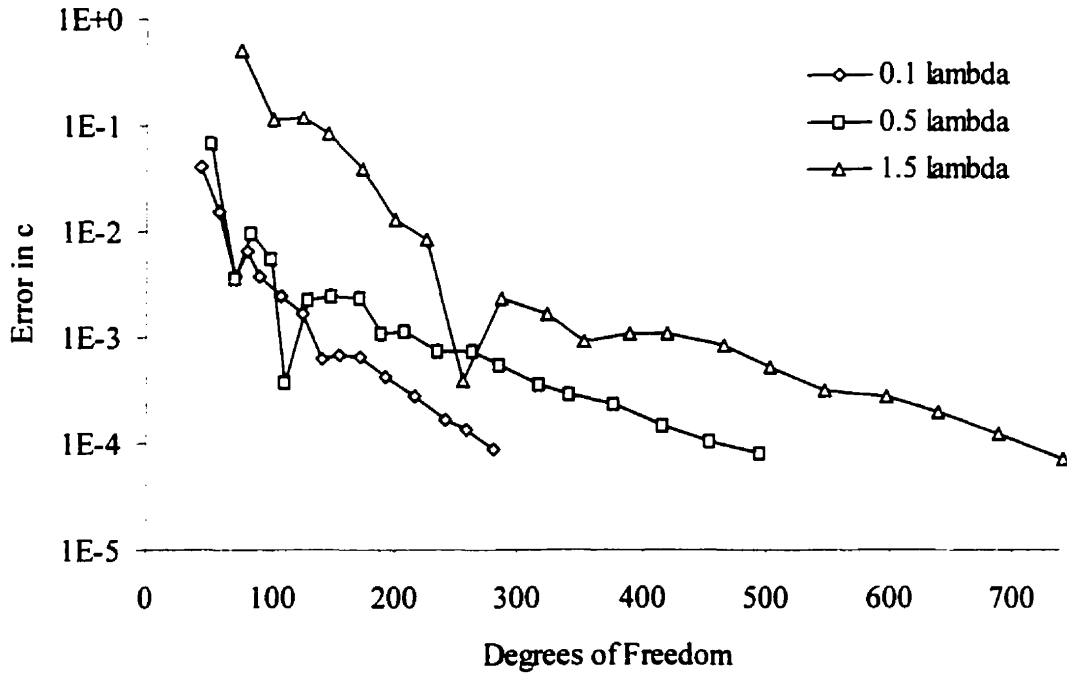


Figure 6.13: Convergence of p -adaption for the miter bend, with ports placed three different distances away. The cost function is $c = |S_{11}|^2$; frequency is $f = 11.25$ GHz and geometry is $\underline{g} = [2\sqrt{2} \ \sqrt{2}/2 \ \sqrt{2}/2 \ \sqrt{2}/2]^T$ cm.

The full model of the 2-cavity filter (no symmetry used) generates meshes of 490, 540 and 494 elements for guide lengths of 1.5λ , 0.5λ and 0.1λ respectively. The cost function for this problem was the transmission coefficient, $c = |S_{12}|^2$, and the frequency of the adaption was 12 GHz. The geometry for this test case is $\underline{g} = [4.5 \ 3 \ 16]^T$ mm. The error tolerance for the filter was chosen to be 10^{-3} .

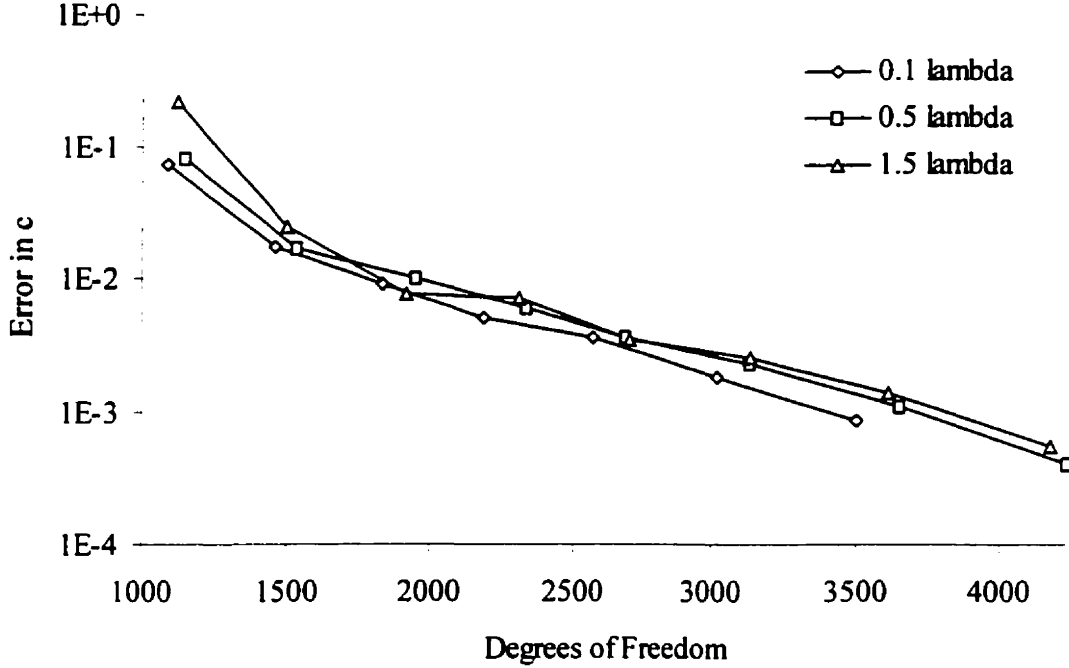


Figure 6.14: Convergence of p -adaption for the 2-cavity filter, with ports placed three different distances away. The cost function is $c = |S_{12}|^2$; frequency is $f = 12$ GHz and geometry is $\underline{g} = [4.5 \ 3 \ 16]^T$ mm.

From the viewpoint of computational efficiency, the above results suggest that 0.1λ is the best distance to place the ports. However, there are other problems for which 0.1λ is not the best choice [29]. When the guide length is large (such as 1.5λ), more elements are generated in the mesh, but since the fields near the ports vary slowly, they require low polynomial orders. Alternatively, placing ports too close to the junction can cause the generation of a large number of small elements [29] – which may not be needed because

high order elements are available. The ideal placement of the ports may vary from problem to problem.

Overall, while it is difficult to conclude what the optimal distance is from port to device, the important point is that by using the port element, accurate, converged results can be obtained regardless of where the port is placed. In addition, all three graphs show the TEI to be an effective indicator in guiding the adaption to accurate and quick convergence of different cost functions, for a variety of problems.

6.3.2 Derivatives of Cost Functions using Hierarchal, *P*-Type Elements

The cost function derivatives are computed for hierarchal, *p*-type elements using (2.69) and the theory from Chapters 2 and 3. The results are validated in two ways.

- a) The true error in a derivative is computed when design sensitivities can be calculated analytically.
- b) For problems with no analytic solution, the discrepancy between the derivatives calculated with (2.69) and a finite difference formula is used for validation.

A typical central finite difference formula is used to calculate the derivative of a cost function with respect to parameter g :

$$\frac{\partial c}{\partial g_k} \equiv \frac{c(g_k + \Delta g_k) - c(g_k - \Delta g_k)}{2\Delta g_k} \quad (6.17)$$

where $\Delta g_k = 10^{-3}$ cm for the T-junction and miter bend test cases and $\Delta g_k = 10^{-3}$ mm for the 2-cavity filter problem. The values $c(g_k + \Delta g_k)$ and $c(g_k - \Delta g_k)$ are obtained from FE analyses. The “dragging” method was used instead of re-meshing (as described in section 5.4.2) and thus variation by a small amount Δg_k does not change the structure of the mesh (i.e. no diagonal flipping). Cost function discontinuity, therefore, is not a concern.

For test cases with no analytic solution, there are two types of comparisons made between derivatives computed within the FE solution, with (2.69), and those computed with (6.17). The first type plots the percent discrepancy between the two methods for a p -adaptive FE solution. The second type plots the percent discrepancy for a range of frequencies, but a fixed number (and distribution) of DOFs. For every adaptive step (and frequency point, for the second set of results), computation of $c(g_k + \Delta g_k)$ and $c(g_k - \Delta g_k)$ is required for finite difference calculation of the derivative. The number of DOFs in computing these additional cost functions is the same as computing $c(g_k)$. Therefore, the derivative found with (6.17) has the same accuracy as $\partial c / \partial g_k$ computed by (2.69). For this reason, the discrepancy between derivatives is not expected to change with the number of DOFs or with the test frequency.

The error in a cost function derivative can be calculated exactly for a length of uniform, air-filled, rectangular waveguide. The derivative of the cost function, $c = \angle S_{12}$, with respect to length is calculated analytically using (6.5). Figure 6.15 plots the percent error in dc/dl (for uniform, 1st, 2nd and 4th order elements) computed with (2.69), for a range of wavenumbers. The mesh has 24 elements and the derivative was evaluated for 51 discrete points in the wavenumber range: $1.1\pi < k_0 < 1.9\pi$.

As expected, the percent error in the derivative reduces with an increase in element order and varies little with change of wavenumber.

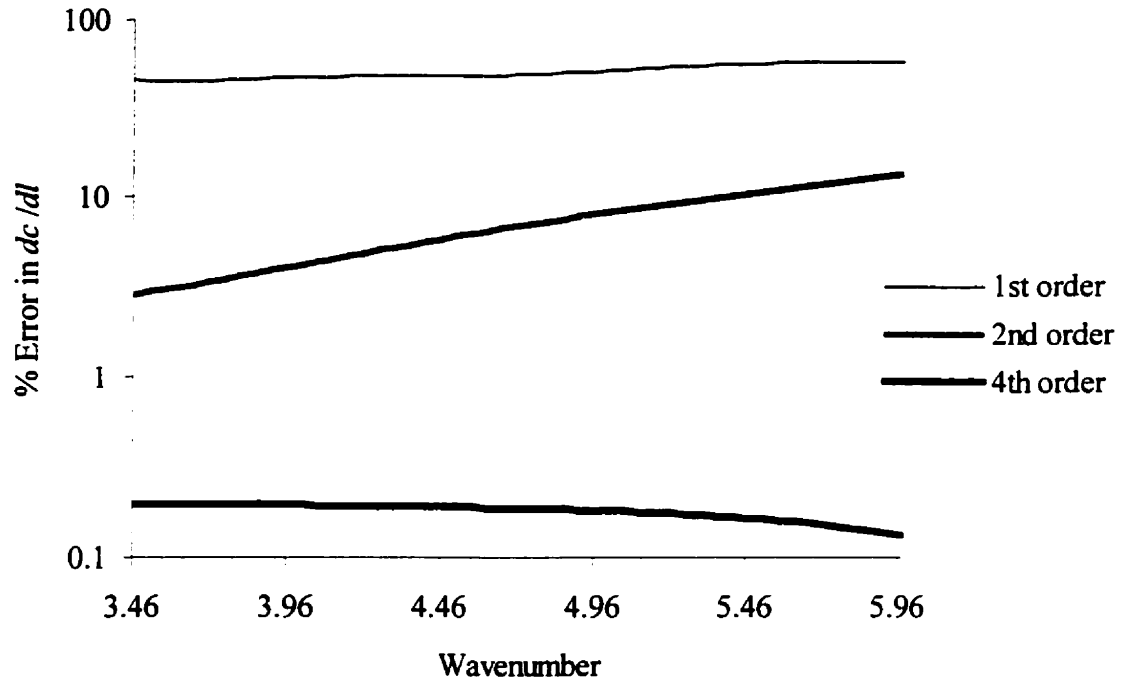


Figure 6.15: Percent error in dc/dl , calculated using (2.69), with an increase in element order for a length of uniform rectangular waveguide. The cost function is $c = \angle S_{12}$; guide length is $l = 2$ m and there are 51 discrete, equally spaced, data points sampled for the range of wavenumbers.

The percent discrepancy in the calculation of $\partial c / \partial g_1$, where $c = \angle S_{11}$, is plotted in Figure 6.16 for the T-junction problem. The error in the cost function is reduced by increasing the order of 25% of the elements by one, at each adaptive step (25% adaption). The adaptive process is terminated when all elements reach 10th order. The frequency of adaption is 10 GHz and the T-junction geometry is $\underline{g} = [0.5 \ 1]^T$ cm – modeled with a 17 element mesh. A plot of percent discrepancies in the derivative throughout the adaption is given in Figure 6.16. Figure 6.17 plots the percent discrepancy over a frequency range of $8 \text{ GHz} < f < 12 \text{ GHz}$ for 51 different points. The element orders used for the result in Figure 6.17 are those at the 10th step of the adaption (225 DOFs) from Figure 6.16.

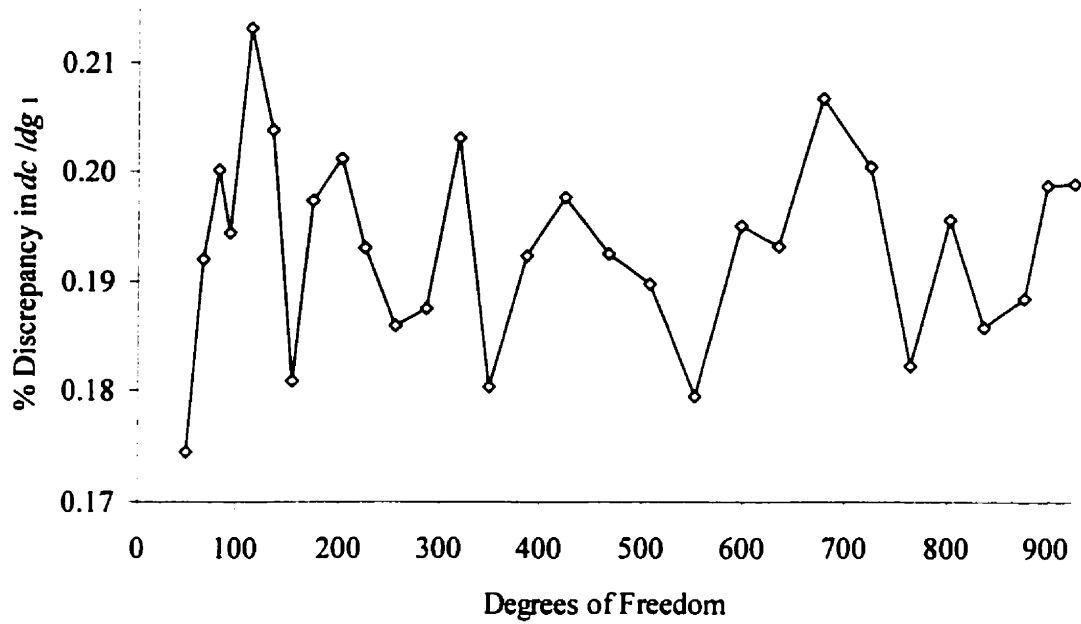


Figure 6.16: Percent discrepancy in $\partial c/\partial g_1$ using (2.69) and (6.17) for the T-junction problem – solved at 25% adaption. The cost function is $c = \angle S_{11}$; frequency is $f = 10$ GHz and geometry is $\underline{g} = [0.5 \ 1]^T$ cm.

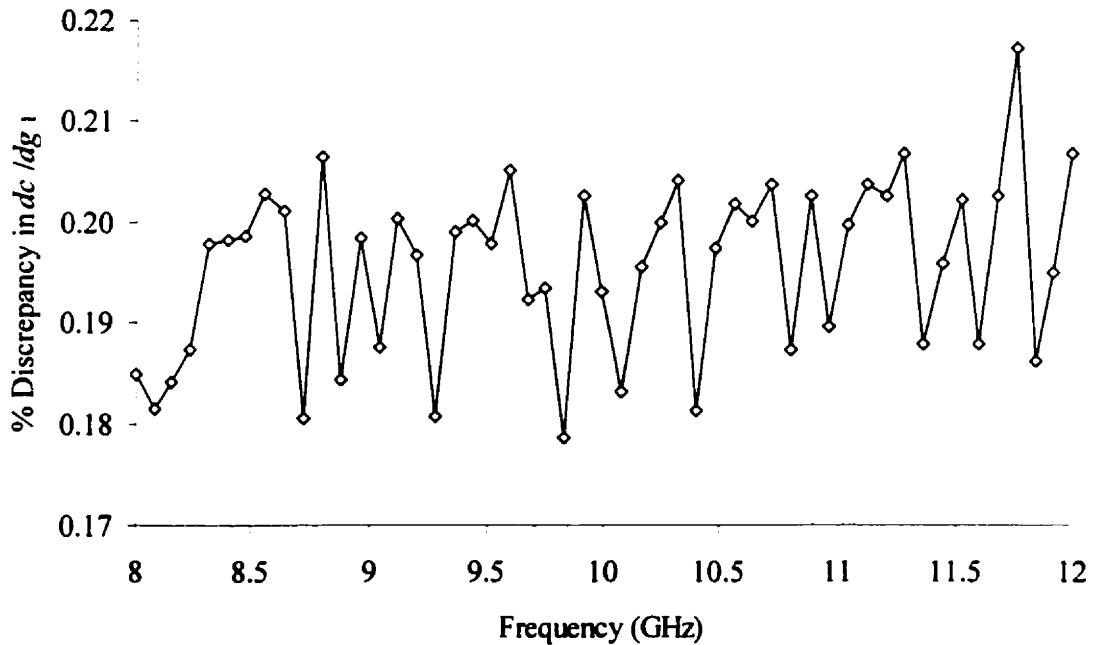


Figure 6.17: Percent discrepancy in $\partial c/\partial g_1$ using (2.69) and (6.17) for the T-junction problem – solved at 51 different frequencies using the same distribution of DOFs (225 DOFs) as the 10th adaptive step of Figure 6.16.

Both Figures 6.16 and 6.17 show good agreement (less than 0.22%) between the derivative calculation from (2.69) and the finite difference calculation of (6.17). As expected, there is little change in the discrepancy of the design sensitivity for the T-junction when varying the number of DOFs or frequency.

Figures 6.18 and 6.19 plot the percent discrepancies in $\partial c / \partial g_3$, where $c = |S_{12}|$, for the miter bend problem. The 25% adaptive process reduces the error in $c = |S_{12}|$ until all elements are 10^{th} order. The frequency of adaption is 11.25 GHz and the miter bend geometry is given by $\underline{g} = [2\sqrt{2} \quad \sqrt{2}/2 \quad \sqrt{2}/2 \quad \sqrt{2}/2]^T$ cm – generating a mesh of 16 elements. Each derivative calculation in Figure 6.19 is based on the element configuration from the 10^{th} step of the adaption, which has 172 DOFs in its FE solution. The frequency range is $8 \text{ GHz} < f < 12 \text{ GHz}$ and 51 frequency points are sampled.

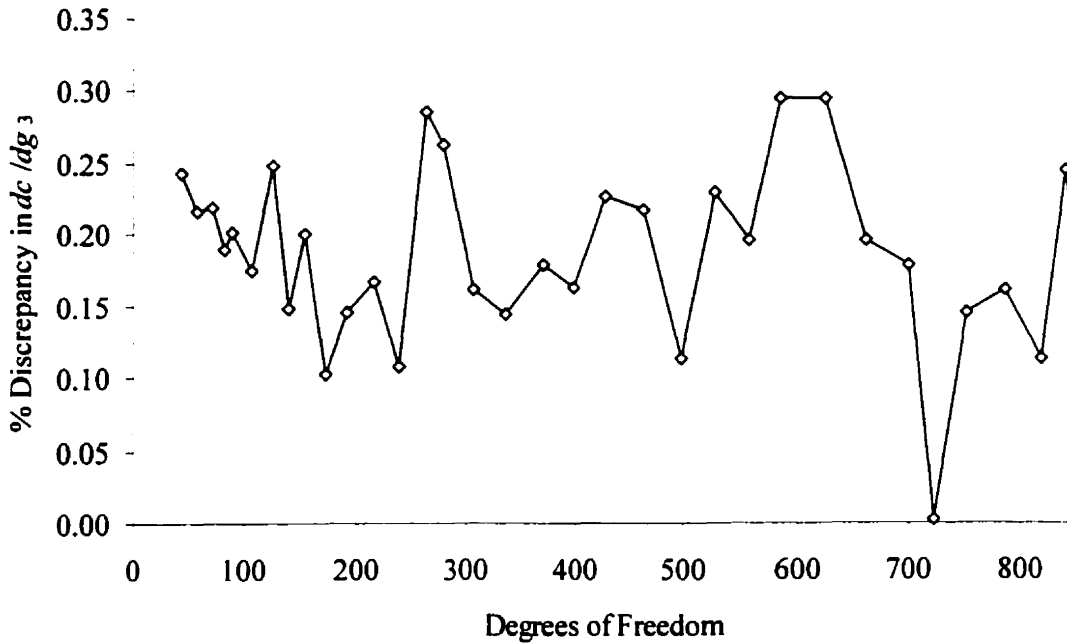


Figure 6.18: Percent discrepancy in $\partial c / \partial g_3$ using (2.69) and (6.17) for the miter bend problem – solved at 25% adaption. The cost function is $c = |S_{12}|$; frequency is $f = 11.25 \text{ GHz}$ and geometry is $\underline{g} = [2\sqrt{2} \quad \sqrt{2}/2 \quad \sqrt{2}/2 \quad \sqrt{2}/2]^T$ cm.

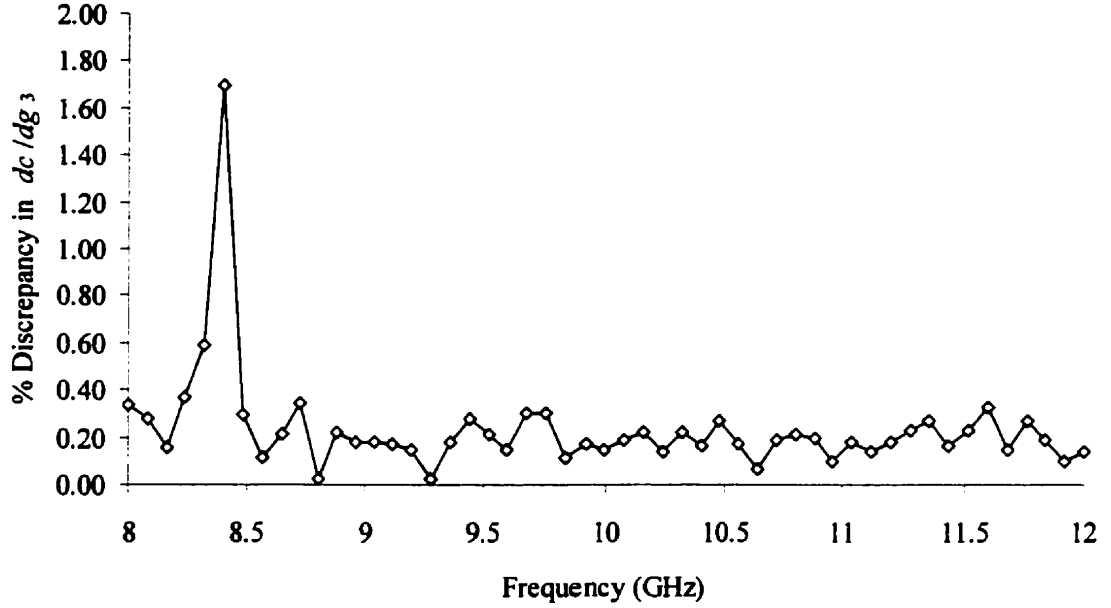


Figure 6.19: Percent discrepancy in $\partial c/\partial g_3$ using (2.69) and (6.17) for the miter bend problem – solved at 51 different frequencies using the same distribution of DOFs (172 DOFs) as the 10th adaptive step of Figure 6.18.

Figure 6.18 shows good agreement of derivative calculations for the adaptive FE solution at 11.25 GHz. Figure 6.19 shows a similar level of discrepancy as that of Figure 6.18, except at around $f = 8.4$ GHz. The higher discrepancy (about 1.7 %) at 8.4 GHz is a result of a resonant frequency at that point, e.g. RL (return loss) $\rightarrow -\infty$ or $|S_{11}| \cong 0$. Derivatives of S -parameters around this point are very sensitive to small variations in g_3 – thus making it difficult to calculate with a finite difference equation using the same step size, Δg_3 , as at the other frequencies. It is likely that the derivative calculation of (2.69) is more accurate at such a point.

For the 2-cavity filter, graphs for percent discrepancy of $\partial c/\partial g_3$, $c = |S_{12}|$, are plotted in Figures 6.20 and 6.21. The 25% adaption in Figure 6.20 reduces the error of c until each element is 10th order. The frequency of adaption is 12 GHz and geometry of the problem is $\underline{g} = [4.5 \ 3 \ 16]^T$ mm. The full geometry (no symmetry) is meshed with 494 elements. The results in Figure 6.21 are based on the element order distribution of the

10th step of the adaption (5277 DOFs). The 51 frequencies sampled are from the frequency range $11.8 \text{ GHz} < f < 12.2 \text{ GHz}$.

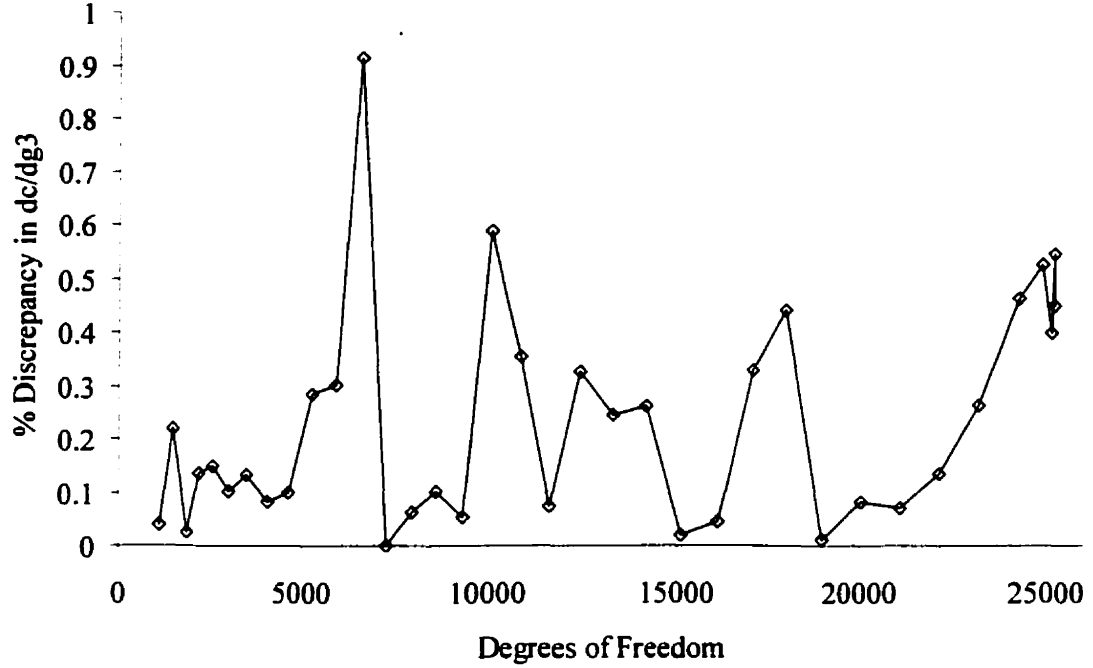


Figure 6.20: Percent discrepancy in $\partial c / \partial g_3$ using (2.69) and (6.17) for the 2-cavity filter problem – solved at 25% adaption. The cost function is $c = |S_{12}|$; frequency is $f = 12 \text{ GHz}$ and geometry is $g = [4.5 \ 3 \ 16]^T \text{ cm}$.

The discrepancies in errors are a little higher for the filter than the other test cases – probably due to the presence of two resonant frequencies (one per cavity) in the frequency range. Nonetheless, the Figures 6.20 and 6.21 show good agreement between derivatives (less than 1% discrepancy) for variations of DOF and frequency.

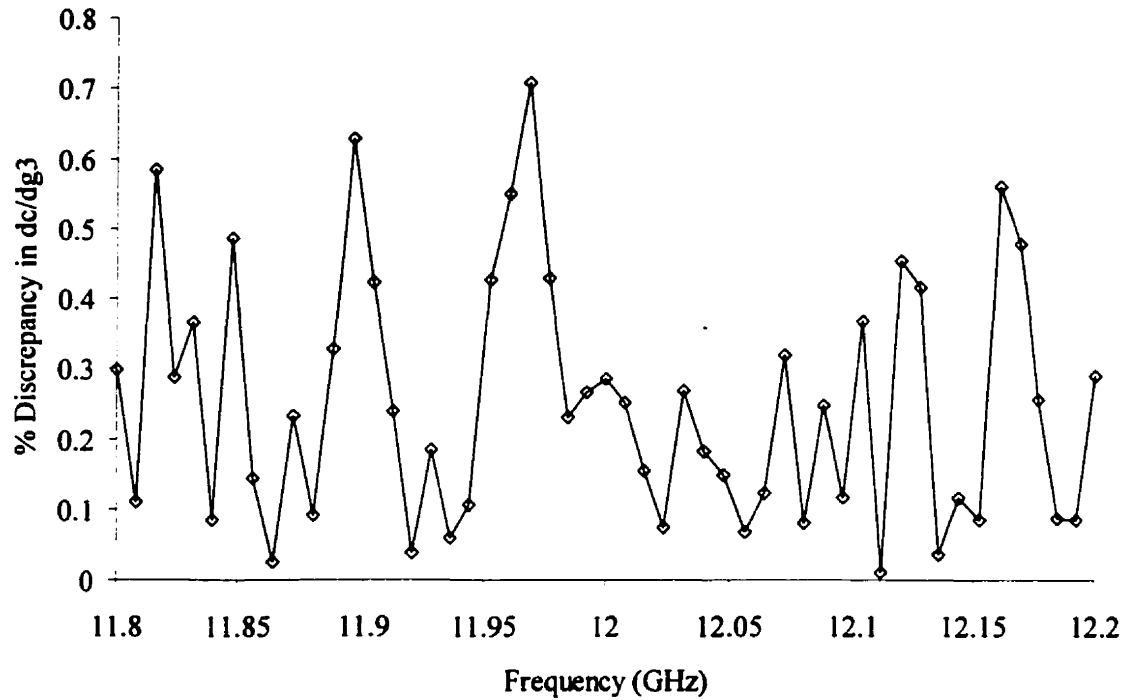


Figure 6.21: Percent discrepancy in $\partial c/\partial g_3$ using (2.69) and (6.17) for the 2-cavity filter problem – solved at 51 different frequencies using the same distribution of DOFs (5277 DOFs) as the 10th adaptive step of Figure 6.20.

Usually, calculating the derivatives of various cost functions with (2.69) yields similar results to the finite difference calculations. Computing derivatives using the adjoint variable method not only saves a great deal of computational effort (by avoiding extra CFEs required by the finite difference method), it appears to be immune to the effects of resonant frequencies. The choice of Δg_k is not always clear when using a finite difference approach and a convergence study is necessary to ensure an accurate calculation. Reducing the size of Δg_k too much may lead to numerical round-off error (because of single precision accuracy used in the FE solution) in calculation of the derivative. The choice of $\Delta g_k = 10^{-3}$ (cm or mm) was based on a convergence study (for each problem) at the center frequency. The study showed the derivative slowly started to diverge after Δg_k was reduced further.

Using (2.69) requires very little extra computational effort. The only controllable limits to the accuracy of the derivative in this case are the same that limit any FE solution

– the discretization of the problem, whether it be the order of the hierarchal elements used in the FE model or the number of elements in the mesh.

6.3.3 Estimates of Error in Cost Function Derivatives

In order to show the effectiveness of the error estimator from Chapter 5, estimates of the percent error in cost function derivatives for the T-junction, miter bend and 2-cavity filter are compared with their corresponding “true” percent errors. The “true” error in the derivative of a cost function is calculated from the converged value, obtained by uniformly refining a mesh consisting of 10^{th} order elements. The percent error estimates are compared for different steps of adaptive processes. The cases of 100% and 25% adaption are examined. Note that in the case of 100% adaption, there is no need for any error indication because all the elements are turned up in order at each adaptive step. In addition, 100% adaption can start with a uniform solution of 1st order, where 25% adaption requires the initial element orders to be 2. The error indicator in the 25% adaption targets the cost function (from section 3.4.2).

The estimates of the error in $\partial c / \partial g_2$, $c = |S_{23}|$, for the T-junction are shown for the cases of 100% and 25% adaption in Figures 6.22 and 6.23, respectively. The problem was solved at the center frequency of the range, $f = 10$ GHz. The geometric parameters used are $\underline{g} = [0.5 \ 1]^T$ cm. Since the initial mesh (of 17 elements, used in sections 6.3.1 and 6.3.2) does not contain enough internal nodes, the mesh of 68 elements is used.

In all the plots of error estimates in this section, any values that lie above the diagonal line are overestimates while underestimates of the error lie below the line. The estimated percent error in Figure 6.22 shows a slight, but consistent overestimate of the true percent error in the derivative. Figure 6.23 shows that the error estimator is no less effective in true p -adaptive schemes, when using elements of different orders in the mesh.

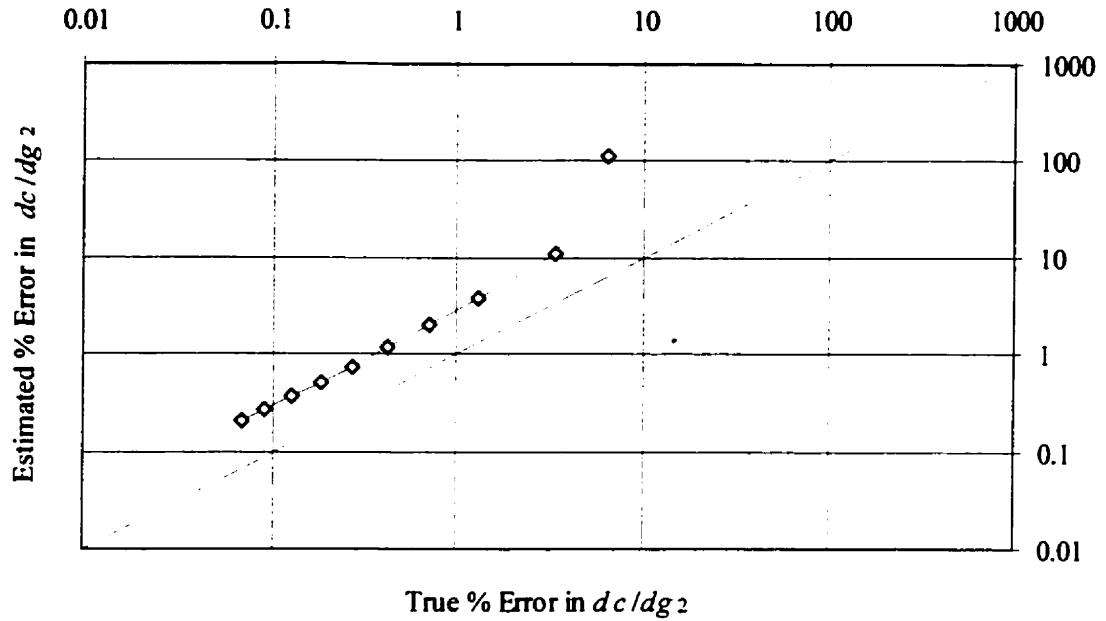


Figure 6.22: Estimated vs. true percent error in $\partial c/\partial g_2$ ($c = |S_{23}|$) for 100% adaption in the T-junction problem. The initial element order is 1; frequency is $f = 10$ GHz and geometry is $\underline{g} = [0.5 \ 1]^T$ cm.

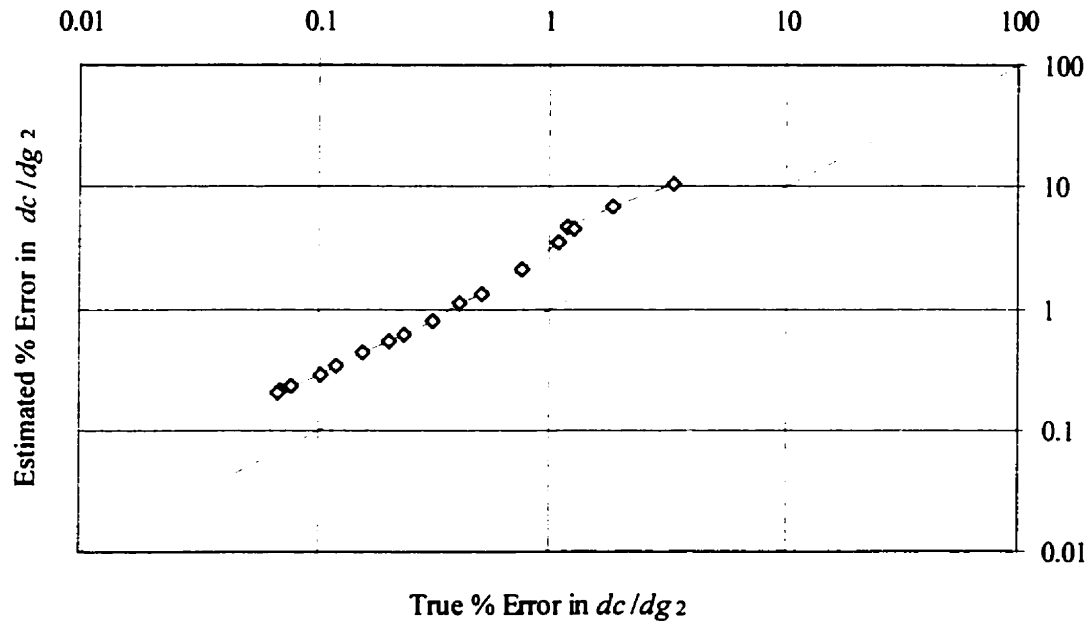


Figure 6.23: Estimated vs. true percent error in $\partial c/\partial g_2$ ($c = |S_{23}|$) for 25% adaption in the T-junction problem. The initial element order is 2; frequency is $f = 10$ GHz and geometry is $\underline{g} = [0.5 \ 1]^T$ cm.

The miter bend problem was solved adaptively at a test frequency of 11.25 GHz for the geometry $\underline{g} = [2\sqrt{2} \quad \sqrt{2}/2 \quad \sqrt{2}/2 \quad \sqrt{2}/2]^T$ cm. The estimates of the error in the derivative $\partial c/\partial g_1$, $c = \angle S_{12}$ are plotted in Figures 6.24 and 6.25. The initial mesh of 16 elements does not contain a single internal node, so the uniformly refined mesh of 64 elements was used.

The results in Figures 6.24 and 6.25 show very good agreement between estimated and true error, especially in the later adaptive steps. Once again, there is no loss of estimate accuracy in using p -adaptive (25% case) elements to reduce the error.

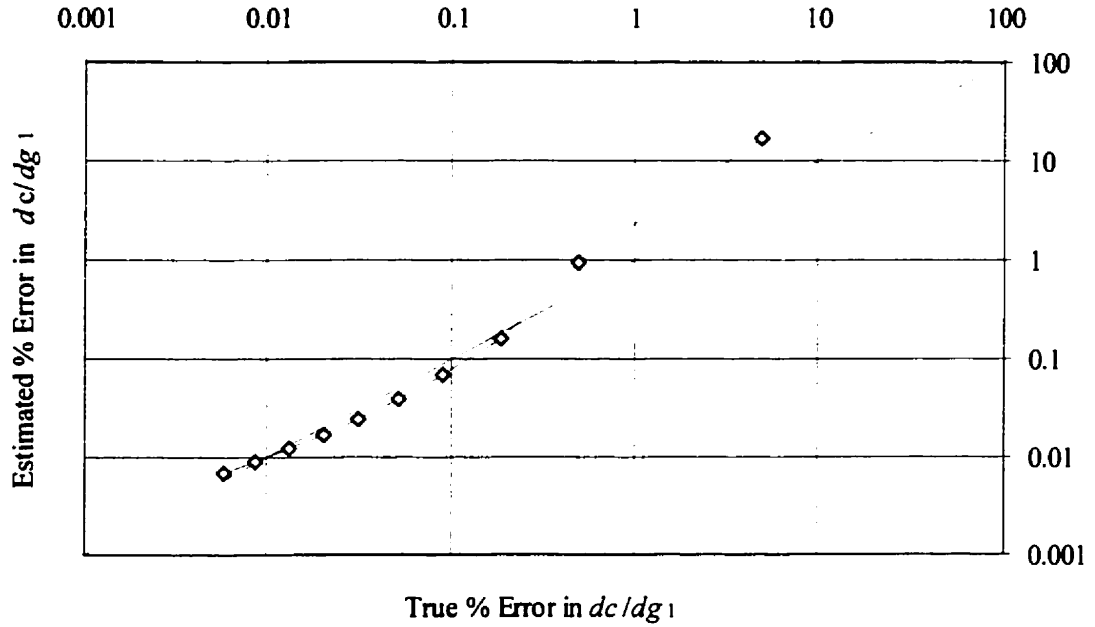


Figure 6.24: Estimated vs. true error in $\partial c/\partial g_1$ ($c = \angle S_{12}$) for 100% adaption in the miter bend problem. The initial element order is 1; frequency is $f = 11.25$ GHz and geometry is $\underline{g} = [2\sqrt{2} \quad \sqrt{2}/2 \quad \sqrt{2}/2 \quad \sqrt{2}/2]^T$ cm.

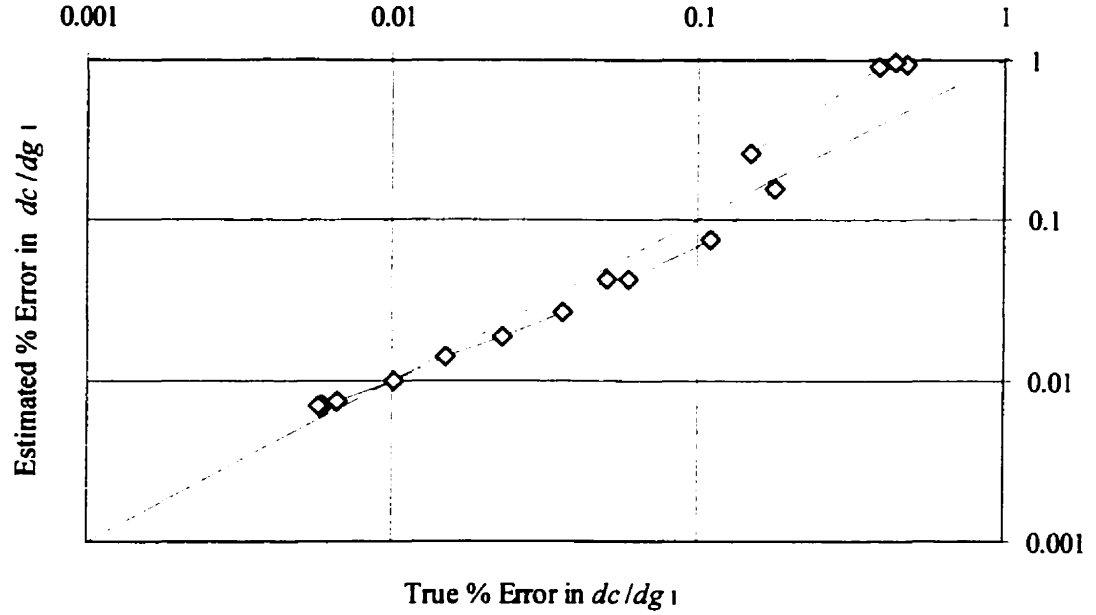


Figure 6.25: Estimated vs. true error in $\partial c/\partial g_1$ ($c = \angle S_{12}$) for 25% adaption in the miter bend problem. The initial element order is 2; frequency is $f = 11.25$ GHz and geometry is $\underline{g} = [2\sqrt{2} \quad \sqrt{2}/2 \quad \sqrt{2}/2 \quad \sqrt{2}/2]^T$ cm.

The estimates of the error in $\partial c/\partial g_1$, $c = |S_{11}|^2$ for the 2-cavity filter problem are given in Figures 6.26 and 6.27. The problem is solved with the geometric parameters $\underline{g} = [4.5 \quad 3 \quad 16]^T$ mm. The mesh (of the full model) has 494 elements and the frequency of analysis is taken to be the center frequency, $f = 12$ GHz.

While there is an underestimate (under an order of magnitude) in both Figures 6.26 and 6.27, it is important to note that with an increase of DOFs, there is a reduction of discrepancy between estimated and true errors. By the last steps of the 100% case and step 11 of the 25% adaption, the estimated percent error has practically converged to the true percent error.

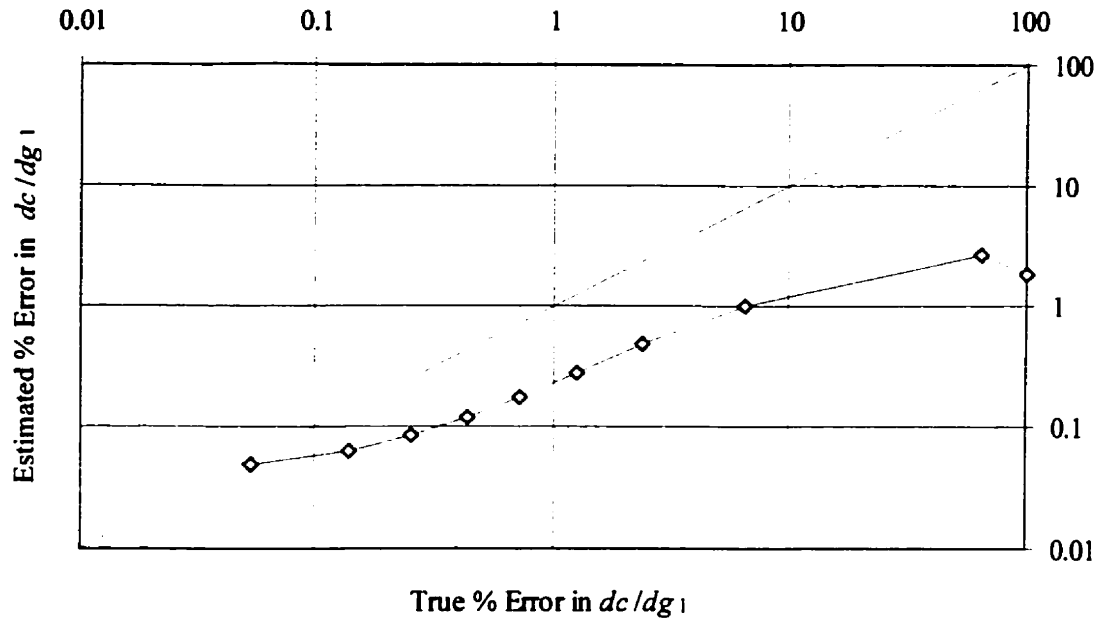


Figure 6.26: Estimated vs. true error in $\partial c/\partial g_1$ ($c = |S_{11}|^2$) for 100% adaption in the 2-cavity filter problem. The initial element order is 1; frequency is $f = 12$ GHz and geometry is $\underline{g} = [4.5 \ 3 \ 16]^T$ mm.

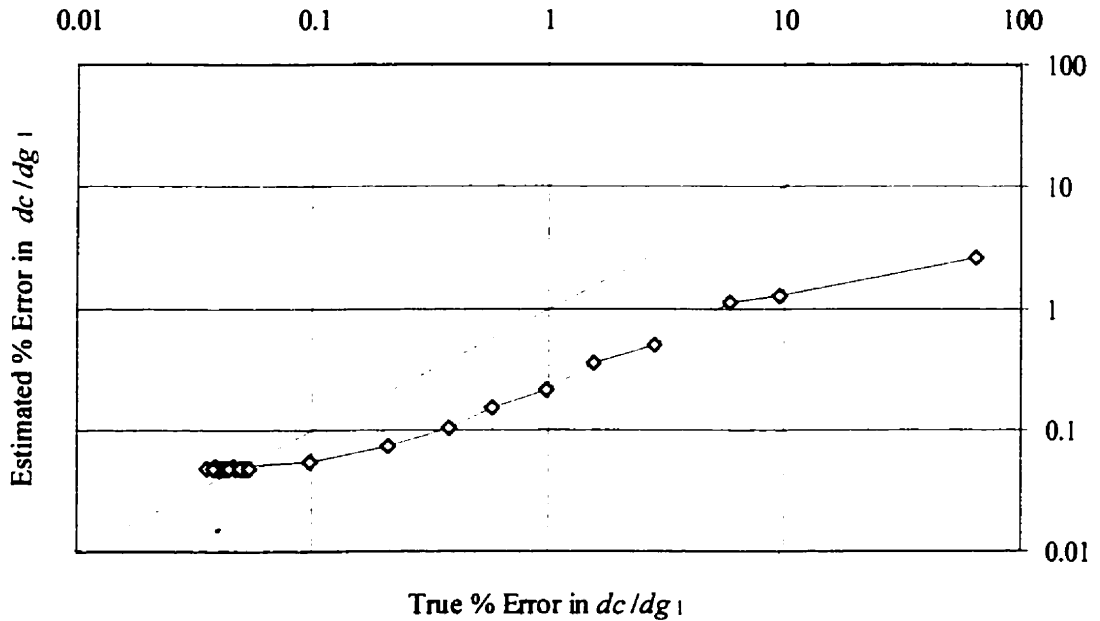


Figure 6.27: Estimated vs. true error in $\partial c/\partial g_1$ ($c = |S_{11}|^2$) for 25% adaption in 2-cavity filter problem. The initial element order is 2; frequency is $f = 12$ GHz and geometry is $\underline{g} = [4.5 \ 3 \ 16]^T$ mm.

In each test problem, the 25% adaptive solution yielded equally accurate estimates of gradient error as in the 100% adaptive scheme. An interesting side note is the effectiveness of using a truly adaptive scheme, such as 25%. While termination is taken to be the saturation of every element at 10th order, many of the steps from the 25% adaptive scheme were unnecessary because a similar level of error as the 10th order solution was reached with a small number of steps. The best example of this is in the 2-cavity filter, where Figure 6.26 shows that to reduce the true error in the derivative to a level of about 0.05 %, a uniform 10th order solution is needed. Alternatively, the scheme in Figure 6.27 requires only 11 adaptive steps (25% of the elements turned up in order at each step) to reach roughly the same error.

In the development of the error estimator of section 5.3, a difficulty in accurately estimating errors of gradients near optima was overcome by defining the value S_m in (5.11). Early versions of the error estimator used an average value (for the internal nodes) of the maximum of the partial derivatives of the *cost function itself* – not the maximum of the partial derivatives of both real and imaginary parts of the S -parameters (as is used in calculating S_m). Depending on the definition of the cost function, there was sometimes a cancellation effect between the real and imaginary S -parameters that occurred in the calculation of $\partial c / \partial x$, $\partial c / \partial y$ or $\partial c / \partial z$. Using such an approach to estimate errors risks massive underestimation when the derivative is very small.

The miter bend test case was re-analyzed at a different frequency and geometry, where the derivative of the cost function is very small. The estimates of error in the derivative $\partial c / \partial g_1$, $c = |S_{12}|^2$, are given for adaptive solutions in Figures 6.28 and 6.29. The frequency of the analysis is 10 GHz and the geometry of the device is $\underline{g} = [1.1208 \ 0.4989 \ 0.06713 \ 0.5685]^T$ cm. The mesh has 50 elements and the converged value of the derivative is $\partial c / \partial g_1 = 4.3 \times 10^{-3}$ (the derivative value in the previous miter bend test case was $\partial \angle S_{12} / \partial g_1 \cong 136$).

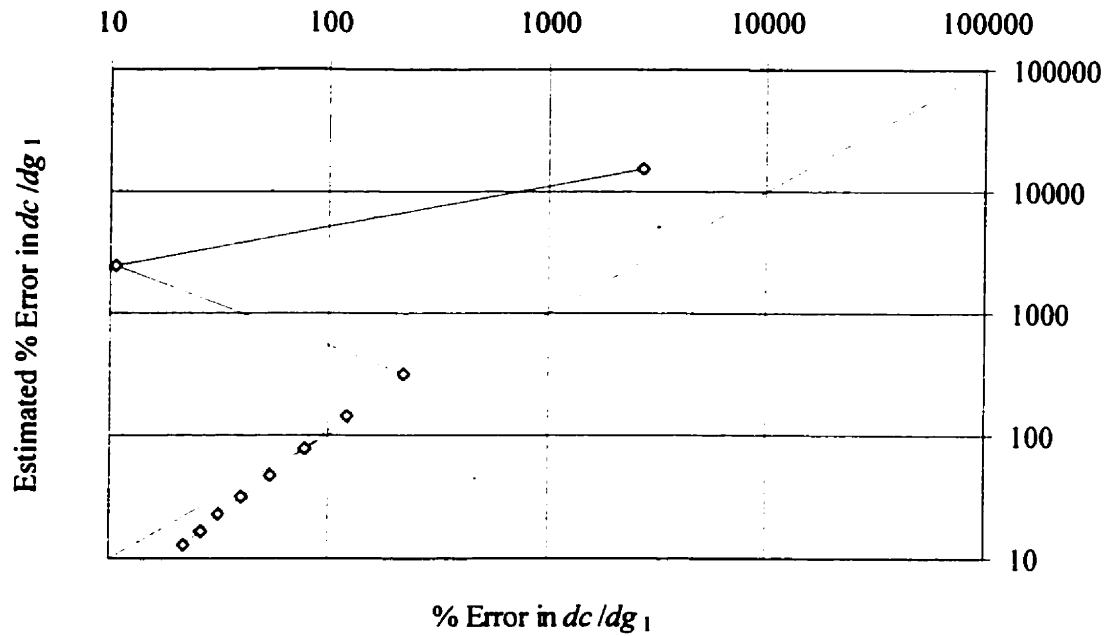


Figure 6.28: Estimated vs. true error in $\partial c / \partial g_1$ ($c = |S_{12}|^2$) for 100% adaption in the miter bend problem. The initial element order is 1; frequency is $f = 10$ GHz and geometry is $\underline{g} = [1.1208 \ 0.4989 \ 0.06713 \ 0.5685]^T$ cm.

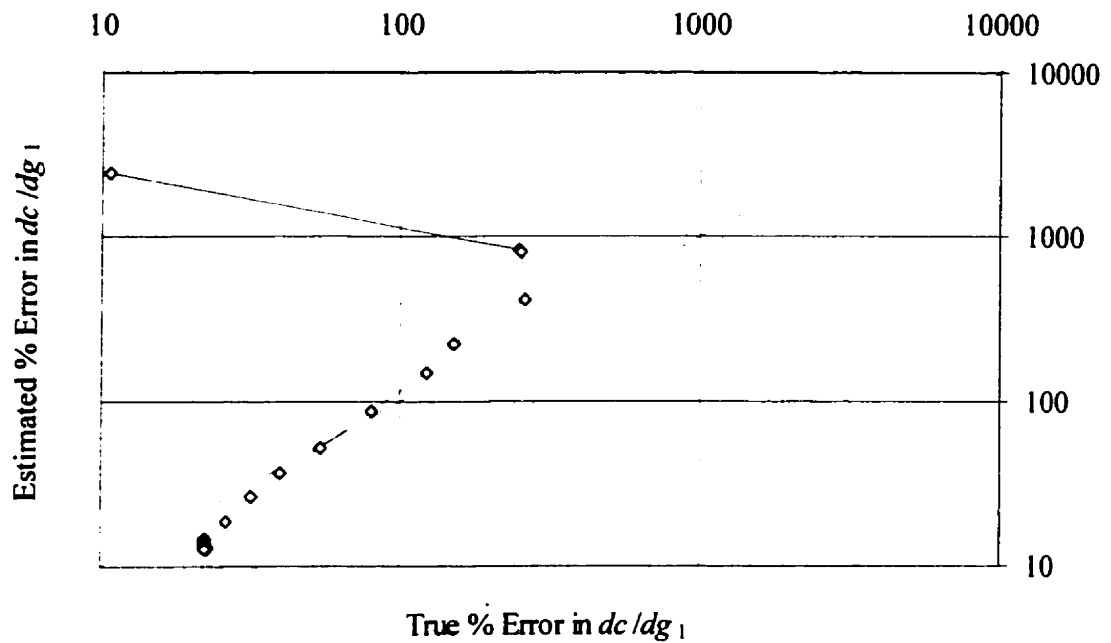


Figure 6.29: Estimated vs. true error in $\partial c / \partial g_1$ ($c = |S_{12}|^2$) for 25% adaption in the miter bend problem. The initial element order is 2; frequency is $f = 10$ GHz and geometry is $\underline{g} = [1.1208 \ 0.4989 \ 0.06713 \ 0.5685]^T$ cm.

Expectedly, the percent errors are quite high in Figures 6.28 and 6.29 – due to the relatively low value of the derivative. The results in these plots are a reminder that as the values of any derivatives in cost function reduce, the difficulty in reducing the percent error increases greatly. However, the point is that regardless of the size of the gradient of the cost function, the error estimates continue to be accurate. One of the error estimates in both figures seems to be a drastic overestimate (error estimate of about 2400% while the true error is 10%). This is attributed to “spurious” convergence in the problem. Figure 6.13 also shows how convergence of S -parameters (and in this case, their derivatives) is not necessarily monotonic in the early stages of an adaption. While the solution may be very small in error for a particular distribution of DOFs, it is purely incidental and not necessarily converged (and unfortunately unpredictable). In this case, the error estimate (2400%) is more accurate than the actual error (10%) in dictating at what stage of the convergence the adaption has reached.

6.3.4 Accuracy Controlled Optimization

The optimization-adaption system from Figure 5.3 was used to optimize the T-junction, miter bend and 2-cavity filter. Three optimizations with different accuracy control (or accuracy links) are compared for each problem: $\alpha = 0.1$, $\alpha = 0.01$ and a third, benchmark case. The benchmark optimization used requires each CFE to be of the same accuracy as the CFE at termination. The termination criterion of (4.31) guarantees that the final gradient of the Lagrangian (when the optimization terminates) is $1/R^{\text{th}}$ its initial value. Combining (5.5) and (4.31) gives the accuracy link for any CFE for the benchmark, or “fixed accuracy” case:

$$N_f \left\| \text{estimate of error in } \nabla c_a(\underline{g}^k) \right\|_2 < \frac{\alpha}{R} \left\| \nabla L(\underline{g}^{\text{initial}}) \right\|_2 \quad (6.18)$$

where $R = 1000$ and $\alpha = 0.1$. There was no need to try a lower level of α in any test case because even with $\alpha = 0.1$, the FE p -adaption saturated each element at 10^{th} order before it could reach the actual accuracy controlled by (6.18).

For each test case, $R = 1000$ was chosen as the termination criterion of (4.31). In addition, 25% adaption was performed at the center frequency to reduce the error in the gradient, $\nabla c_a(\underline{g}^k)$.

Optimizing the T-junction problem for the three different accuracy links yielded similar design parameters. The T-junction geometries for initial and optimal designs are illustrated in Figure 6.30.

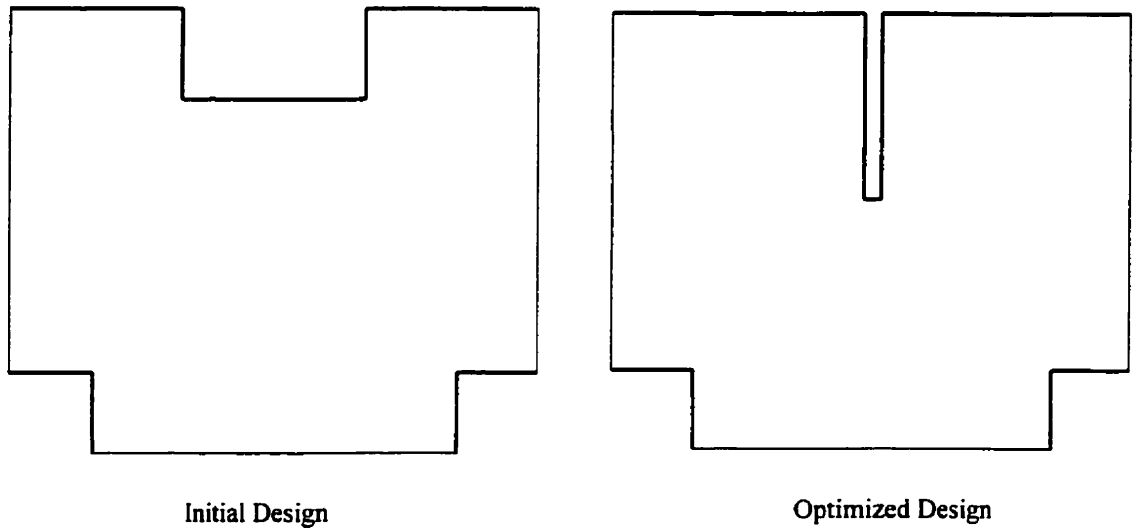


Figure 6.30: Initial and optimized designs of the inductive post for the waveguide T-junction test case. The initial geometry is $\underline{g}^{\text{initial}} = [0.5 \ 1]^T$ cm and final design is $\underline{g}^{\text{optimized}} = [1.0444 \ 0.1000]^T$ cm.

The optimization (for the fixed accuracy case) reduced the cost function from $C(\underline{g}^{\text{initial}}) = 8.95193 \times 10^{-1}$ to an optimal value of $C(\underline{g}^{\text{optimized}}) = 7.9992 \times 10^{-2}$. Both values of cost function are based on uniform 10^{th} order solutions, i.e. the maximum accuracy that can be achieved at any step by p -adaption. The optimized geometry has a mesh of 47 elements. The accuracy of the optimum can be verified by refining the mesh. Uniformly refining the mesh twice gives cost function values of $C(\underline{g}^{\text{optimized}}) = 8.0231 \times 10^{-2}$ (for a

188 element mesh) and $C(\underline{g}^{\text{optimal}}) = 8.0263 \times 10^{-2}$ (for a 752 element mesh) – thus validating the accuracy of a 10^{th} order solution for the optimization of the T-junction.

While the effectiveness of the optimization can be seen through the reduction of the cost function, the cost function is based on a discrete sum of reflection coefficients in the frequency domain. A more “continuous” display of the effectiveness of the optimization is a plot of the return loss throughout the frequency range for a large number of points. Figure 6.31 shows a plot of the return loss for 200 evenly spaced frequencies in the given range. Each evaluation of return loss is based on a 10^{th} order solution.

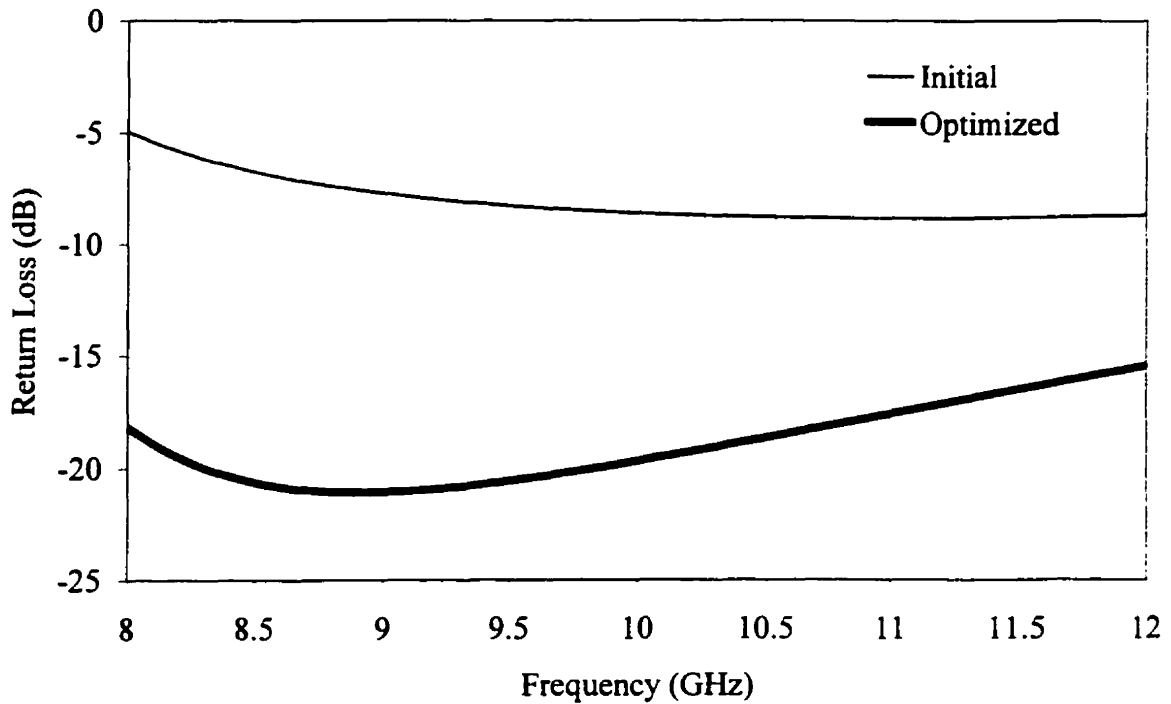


Figure 6.31: Initial and optimized frequency responses of the return loss for T - junction problem. Each curve is based on a discrete sampling of 200 points in the frequency range.

The three different accuracy links lead to different levels of accuracy for CFE throughout each optimization. Figure 6.32 plots the variation of number of DOFs used for CFEs at major optimization steps.

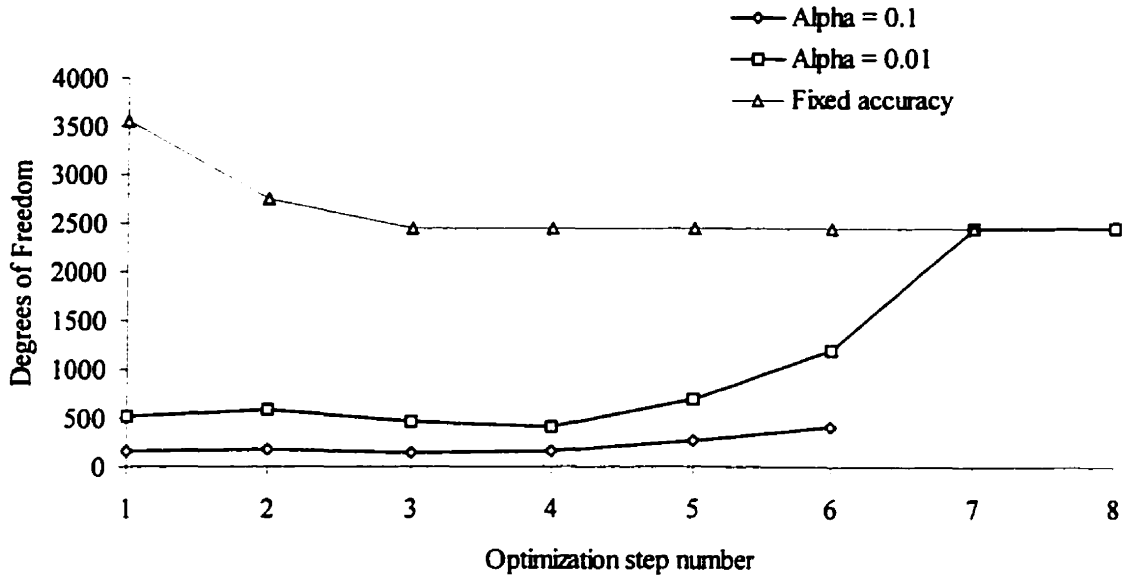


Figure 6.32: Number of DOFs used by CFEs at major optimization steps for the three different accuracy linked optimizations of the T-junction.

As expected, requiring the accuracy of a CFE to be the same as at termination for every optimization step leads to a large number of DOFs at each step. The number of DOFs for the most accurate accuracy link corresponds to meshes consisting of 10^{th} order elements. Normally, a straight line would be expected because the required accuracy is a fixed number. However, varying geometry changes the number of elements in a generated mesh (sometimes by a significant amount) and makes it difficult to keep a perfectly constant level of DOFs throughout the optimization. Both the cases $\alpha = 0.1$ and $\alpha = 0.01$ started with few DOFs and increased throughout the optimization – which is desired in an accuracy controlled optimization. Requiring $\alpha = 0.01$ accuracy at every step resulted in a final CFE accuracy corresponding to a uniform 10^{th} order solution while the $\alpha = 0.1$ case terminated at a lower accuracy level (fewer DOFs).

The three curves in Figure 6.32 are exactly what is desired in the accuracy controlled optimization from Figure 5.3. However, the goal of such a system is to reduce the computational costs of designing a microwave device. Table 6.1 gives the final cumulative computational cost (CCC) for each optimization and a speed-up factor in each

case. The speed-up factor is taken simply to be the CCC of the fixed accuracy optimization divided by the CCC for that case.

Table 6.1: Cumulative computational costs and speed-up factors for different accuracy-links in the T-junction.

| Accuracy Link | Cumulative Computational Cost | Speed-up Factor |
|-----------------|----------------------------------|-----------------|
| Fixed accuracy | 2.72E+09 | 1 |
| $\alpha = 0.01$ | 2.32E+07 | 4.6 |
| $\alpha = 0.1$ | 5.92E+08 | 117 |

The results in Table 6.1 show the remarkable speed-up that can be attained by controlling accuracy using α . In the case of $\alpha = 0.1$, the design process finds the same optimum as the most accurate case, but at 117 times the speed. However, Table 6.1 gives information only about final CCCs. Figure 6.33 plots the values of the cost functions and the CCC throughout the optimization. In order to provide a fair comparison of cost function values, each cost function must be of the same accuracy level. Since 10th order elements achieve the highest accuracy (for a fixed mesh size), each point in Figure 6.33 is the 10th order cost function value at the geometry of that particular optimization step. In other words, it is a plot of the “true” value of the cost function versus the CCCs obtained using the different accuracy links.

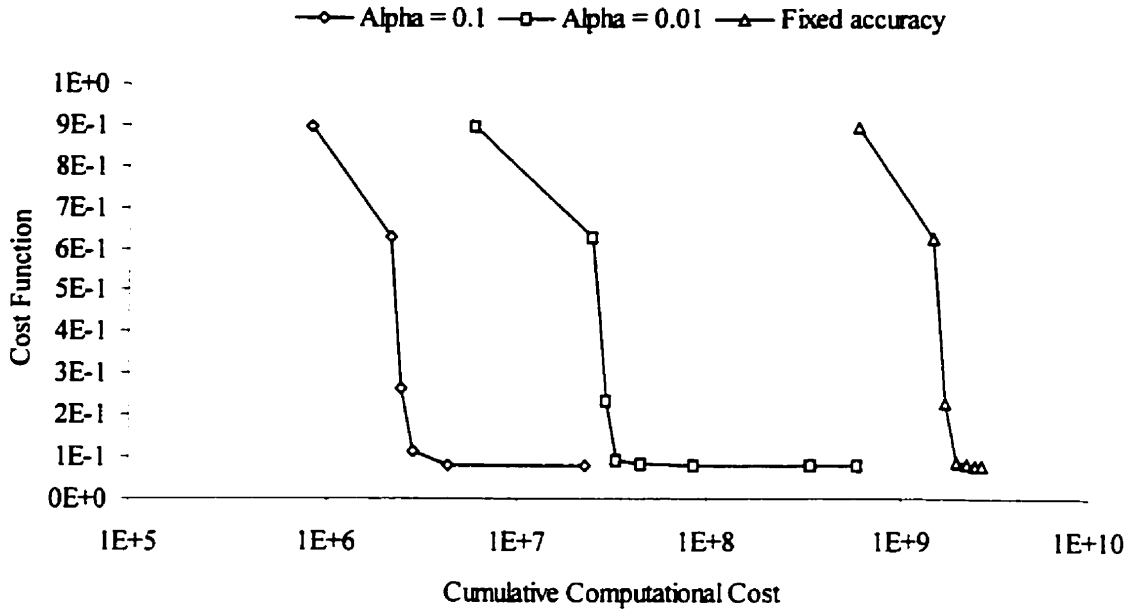


Figure 6.33: “True” values of the cost function and the corresponding cumulative computational costs for different accuracy links in the optimization of the T-junction.

Figure 6.33 shows that a great deal of the computational costs in the optimizations lie in the final few steps, which is expected because the CFEs require higher accuracy levels as an optimum is approached. In addition, Figure 6.33 shows that throughout an optimization (not only at the end), controlling accuracy can drastically reduce computation time without sacrificing the validity of the cost function. In other words, despite the fact that CFEs are of the lower accuracy at the initial stages of an optimization (for $\alpha = 0.1$ and $\alpha = 0.01$), the cost function does genuinely reduce by the same amount as the fixed accuracy optimization— at a much lower cumulative computational cost.

The miter bend with dielectric block was also optimized three times, once for each accuracy link. The three optimizations resulted in the same local optimum. There is negligible distinction between minima found in all three cases. The initial and optimized (based on the fixed accuracy case) geometries for chamfer and dielectric block of the miter bend are shown in Figure 6.34. Other interesting local optima are shown in Appendix B, along with the frequency response of their return losses.

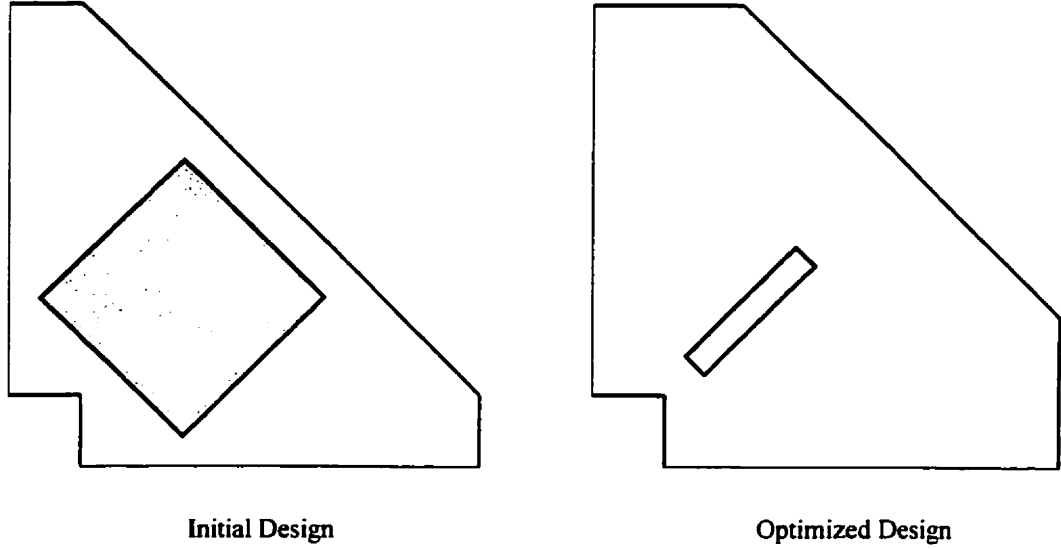


Figure 6.34: Initial and optimized designs of the chamfer and dielectric block for the miter bend test case. The initial geometry is $\underline{g}^{\text{initial}} = [2\sqrt{2} \ 1 \ 1 \ \sqrt{2}/2]^T$ cm and final design is $\underline{g}^{\text{optimized}} = [2.2489 \ 1.0167 \ 0.1346 \ 1.1280]^T$ cm.

The optimization of highest accuracy (fixed accuracy case) reduced the cost function from $C(\underline{g}^{\text{initial}}) = 7.5645 \times 10^{-1}$ to $C(\underline{g}^{\text{optimized}}) = 5.8413 \times 10^{-6}$. The mesh at the optimum (for the highest accuracy case) has 56 elements. Refining the mesh once and re-solving with uniform 10^{th} order elements yields a cost function of $C(\underline{g}^{\text{optimized}}) = 1.8597 \times 10^{-5}$ (for 224 elements). Another uniform refinement creates a mesh of 896 elements and a cost function value of $C(\underline{g}^{\text{optimized}}) = 6.0797 \times 10^{-5}$. The percent change in cost function when uniformly refining the mesh is larger than in the case of the T-junction because of the small value of C . The important point is that 10^{th} order uniform elements are accurate enough that the cost function for 56 elements is genuinely an accurate minimum. Figure 6.35 shows the frequency response of the return loss in the miter bend for the initial and optimized geometries.

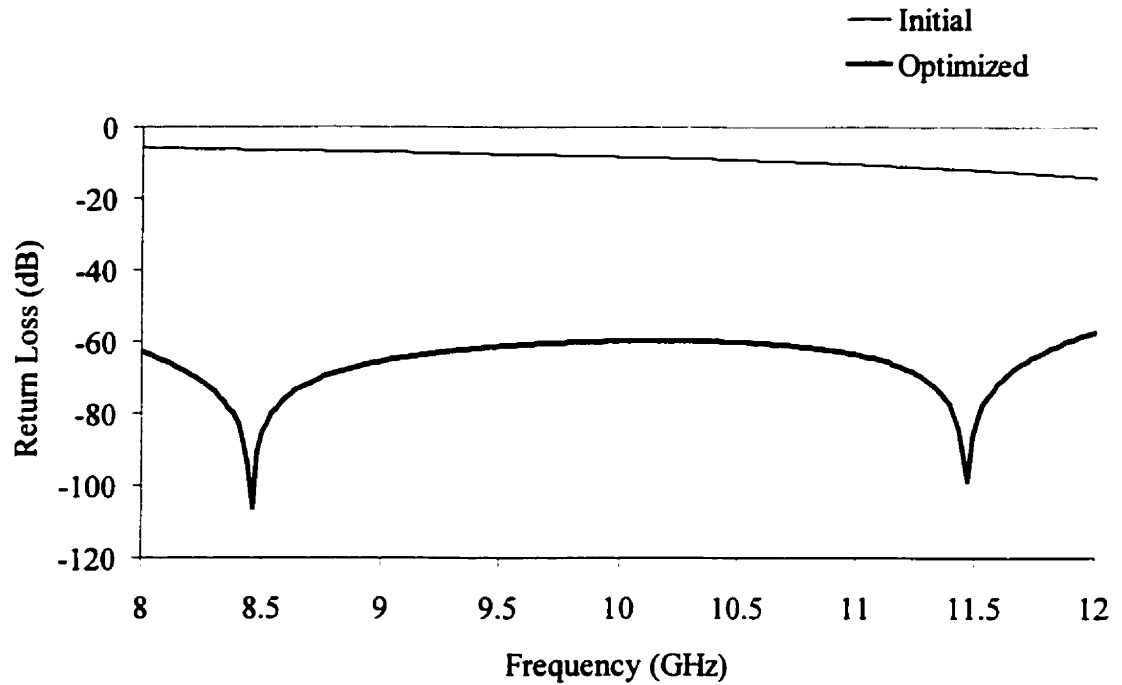


Figure 6.35: Initial and optimized frequency responses of the return loss for the miter bend. Each curve is based on a discrete sampling of 200 points in the frequency range.

The existence of two resonance frequencies in the frequency range (at around 8.5 GHz and 11.5 GHz in the optimized case) explains the dramatic decrease in the cost function throughout the optimization. Without the dielectric block, there are no longer resonant frequencies in the range and the cost function of the same geometry has a much higher value: $C(\underline{g}^{\text{optimized}}) = 1.4464 \times 10^{-1}$! The frequency response of the return loss in the miter bend (with the same optimized chamfer length) without the dielectric block is given in Appendix B.

Figure 6.36 illustrates the variation of number of DOFs used for CFEs at major optimization steps of the miter bend problem.

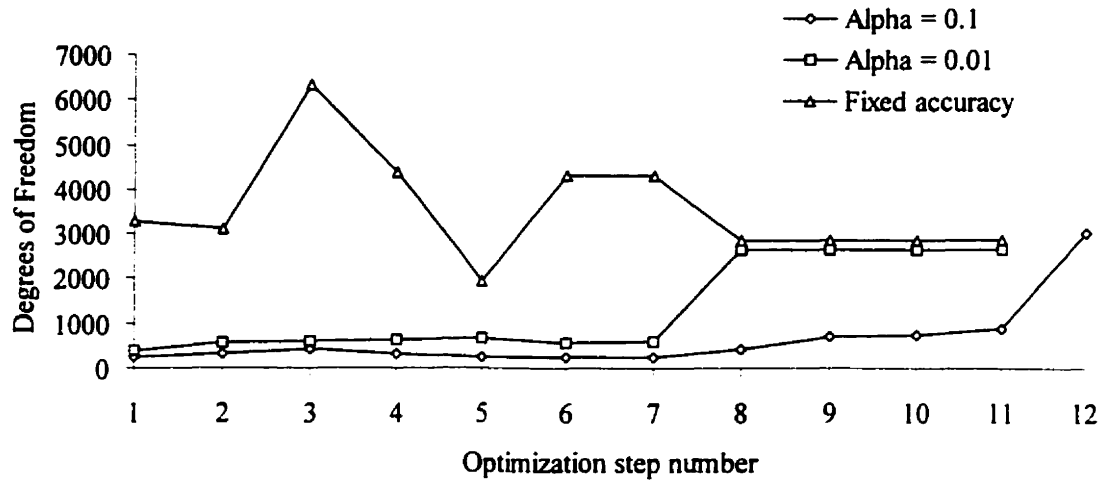


Figure 6.36: Number of DOFs used by CFEs at major optimization steps for the three different accuracy-linked optimizations of the miter bend problem.

The three different accuracy links results in final CFEs of the same accuracy (corresponding to a uniform, 10^{th} order solution). The optimization of fixed accuracy once again oscillates dramatically due to relatively large changes in geometry and thus mesh size.

Table 6.2 gives the final cumulative computational cost for each optimization and a speed-up factor in each case.

Table 6.2: Cumulative computational costs and speed-up factors for different accuracy-links in the miter bend problem.

| Accuracy Link | Cumulative Computational Cost | Speed-up Factor |
|-----------------|----------------------------------|-----------------|
| Fixed accuracy | 6.95E+09 | 1 |
| $\alpha = 0.01$ | 2.54E+09 | 2.7 |
| $\alpha = 0.1$ | 3.19E+08 | 22 |

The results in Table 6.2 show the speed-up that was attained by varying α . Figure 6.37 plots the values of the cost functions and the CCC throughout the optimization.

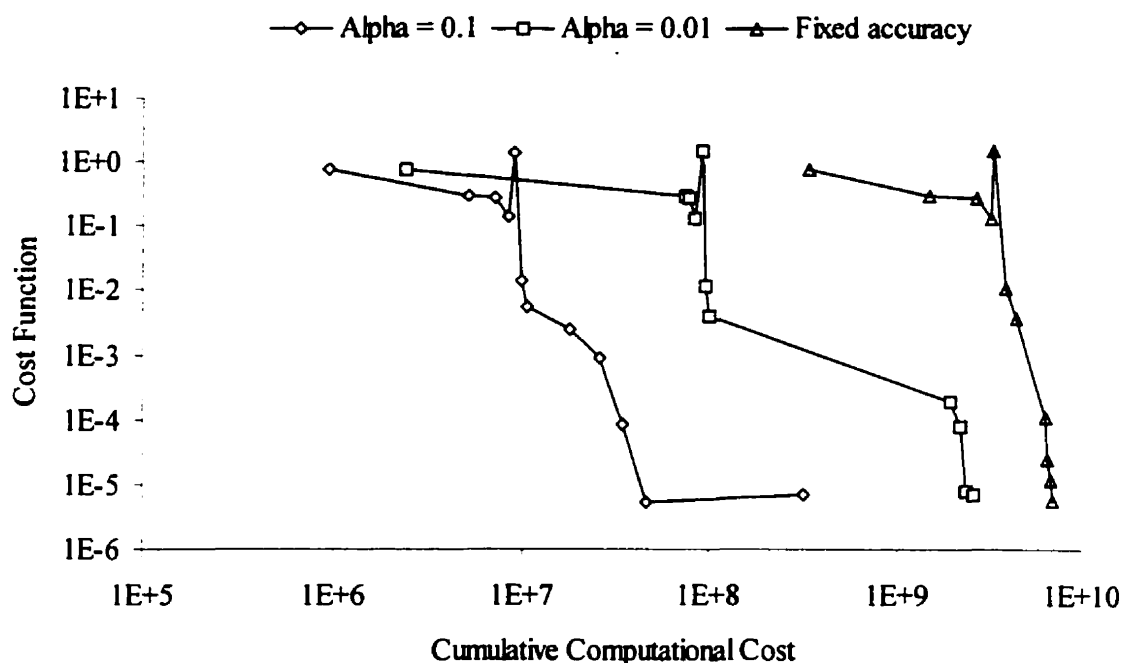


Figure 6.37: “True” values of the cost function and the corresponding cumulative computational costs for different accuracy links in the optimization of the miter bend.

Figure 6.37 shows that great savings can be realized at any step of the optimization when the size of α is increased. The last value of C for the $\alpha = 0.1$ optimization is actually higher than that of the previous step. This is a result of terminating the optimization within a line search. While the actual cost function may increase slightly, the gradient of the Lagrangian has satisfied the termination criterion.

The initial design and final frequency responses for the return loss of the 2-cavity filter are shown in Figure 6.38. The cost function (in the fixed accuracy case) is reduced from $C(\underline{g}^{\text{initial}}) = 9.6227$ to $C(\underline{g}^{\text{optimal}}) = 1.8833$. Due to the sensitivity of the device, the geometric parameters do not change by a large amount in the optimization of the device (maximum parameter change is roughly 20%), yet produce a big difference in the performance of the device.

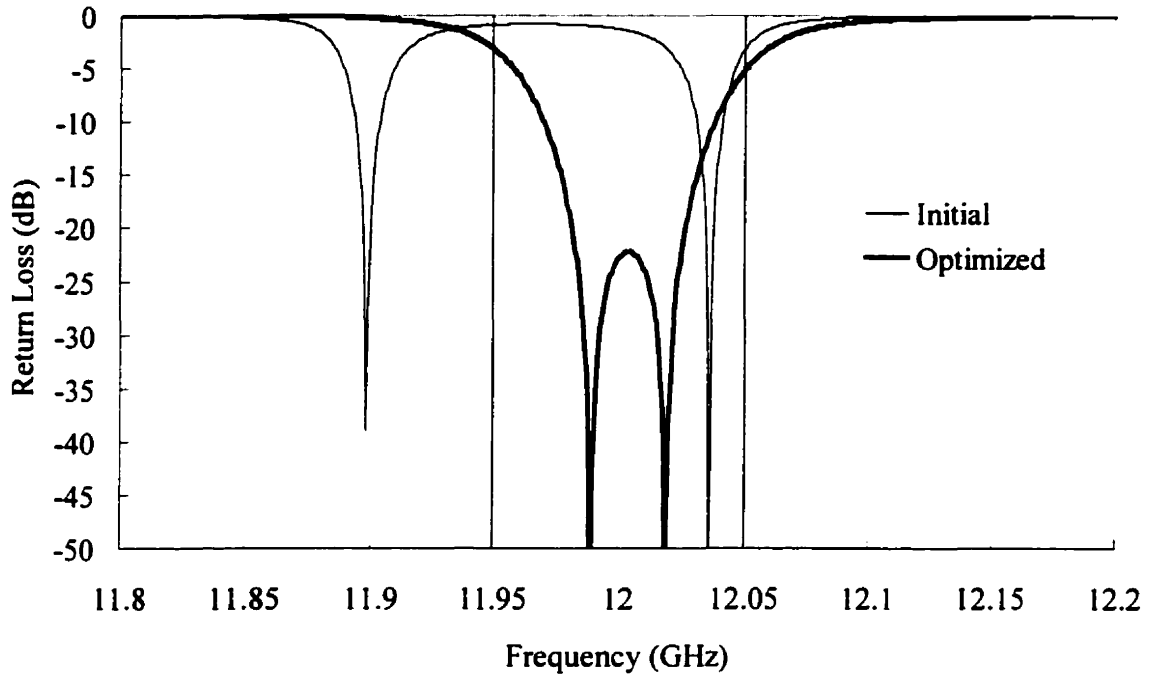


Figure 6.38: Initial and optimized frequency responses of the return loss for the 2 - cavity filter. Each curve is based on a discrete sampling of over 500 points in the frequency range. The initial geometry is $\underline{g}^{\text{initial}} = [4.5 \ 3 \ 16]^T$ mm and the optimized design is $\underline{g}^{\text{optimized}} = [5.4208 \ 2.3938 \ 15.7488]^T$ mm.

All optimizations find the same minimum, where the optimized response corresponds to maximum transmission in the pass band of $11.95 \text{ GHz} < f < 12.05 \text{ GHz}$, as shown in the above figure. In order to save computation time, symmetry was used in the FE analysis and $\frac{1}{4}$ of the problem was solved for each CFE. The final mesh had 149 elements. The meshes of this level are very well discretized and refining the final mesh to 596 elements changes the cost function (by very little) at the optimum to $C(\underline{g}^{\text{optimized}}) = 1.883297$ (from $C(\underline{g}^{\text{optimizedI}}) = 1.883300$)

Figure 6.39 plots the variation of DOFs for each optimization step for the three accuracy-links.

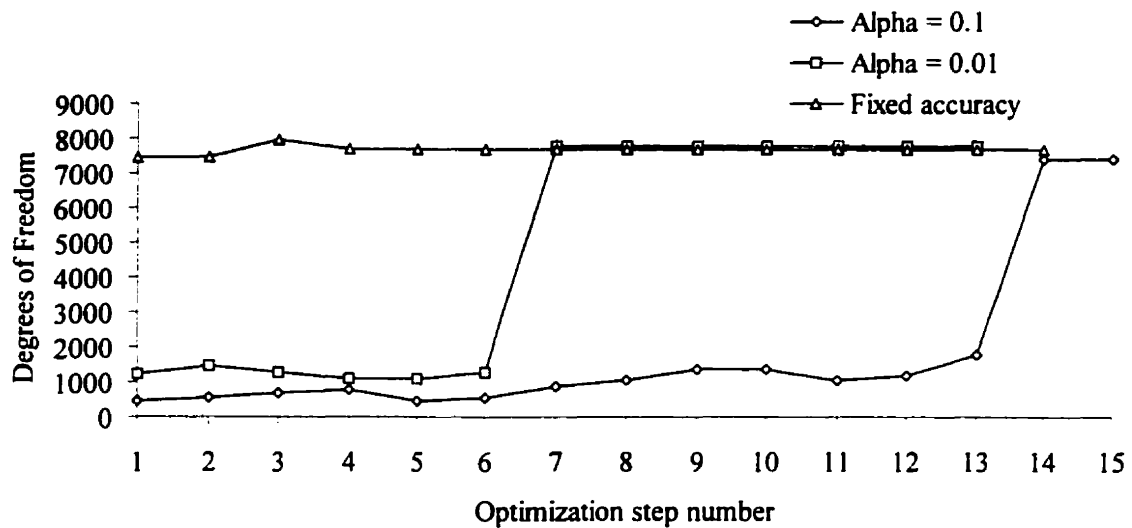


Figure 6.39: Number of DOFs used by CFEs at major optimization steps for the three different accuracy-linked optimizations of the 2-cavity filter problem.

The optimization with the highest control of accuracy in the 2-cavity filter has almost a constant level of DOFs throughout the optimization (corresponding to uniform 10^{th} order elements). The flatness of this curve is expected because small changes in geometry correspond to similar meshes throughout. For the accuracy controls of $\alpha = 0.01$ and $\alpha = 0.1$, there is a big jump to a higher level of accuracy, but as expected, in the case of $\alpha = 0.1$, the large change in accuracy takes place only near the end of the optimization.

Table 6.3 gives the final cumulative computational cost for each optimization and a speed-up factor in each case.

Table 6.3: Cumulative computational costs and speed-up factors for different accuracy-links in the 2-cavity filter problem.

| Accuracy Link | Cumulative Computational Cost | Speed-up Factor |
|-----------------|----------------------------------|-----------------|
| Fixed accuracy | 8.39E+10 | 1 |
| $\alpha = 0.01$ | 3.01E+10 | 2.8 |
| $\alpha = 0.1$ | 7.33E+09 | 11 |

Once again there is great computational cost saved when optimizing with CFEs of lower initial accuracy. The variation of the cost functions and their CCCs are shown in Figure 6.40.

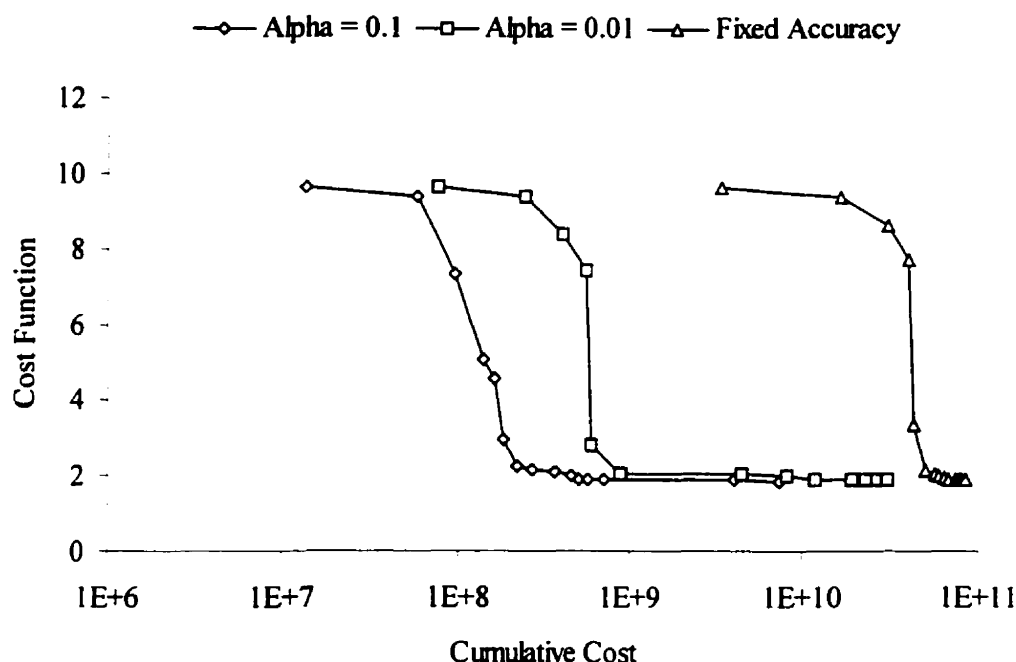


Figure 6.40: “True” values of the cost function and the corresponding cumulative computational costs for different accuracy links in the optimization of the 2-cavity filter problem.

Figure 6.40 shows that once again, while much of the computational effort lies in the final steps of the optimization, there is consistent savings at any point in the optimization. In addition, the figure shows that even when the accuracy demand is low, the cost function is reduced in value at the same rate as the more accurate case.

The results for the optimization of the three problems demonstrate that costs can be saved by controlling the accuracy of the CFEs. While $\alpha = 0.1$ is the best choice in accuracy links for the problems optimized, it is important to remember that there is a saturation (in most cases) of all the elements to 10^{th} order. Had there been looser restrictions of increasing the number of DOFs further (by h -adaptive refinement, for example), the final costs and accuracy at the optimum would be different. The choice of

α (as mentioned in section 5.4.1) implies a control over the final accuracy of the problem and also the path taken to get there. For the most part, the choice of α does not change the accuracy of the final answer because of the 10th order saturation, but does affect the path – where $\alpha = 0.1$ clearly is the best choice for the optimizations performed on the test cases.

Chapter 7

Conclusions

The objective of the work in this thesis has been to illustrate the dramatic cost savings possible when controlling cost function accuracy throughout the optimization of a microwave device. The accuracy-controlled optimization system requires two major components: a gradient-based, constrained optimizer and a CAD tool to perform evaluations of the cost function and its gradient. While an “off-the-shelf” optimizer is used, a p -adaptive FE method was developed to allow accuracy control in a CFE. Developing an effective p -adaptive FE method that satisfies the needs of an accuracy controlled optimization required research in several different areas.

While the gradient of a cost function can be found by a finite difference approach, embedded calculation of the gradient in the FE code saves a lot of computation time. The theory behind the adjoint variable method was extended to allow calculation of design sensitivities with hierarchal elements – thus allowing inexpensive calculation (computationally) of the gradient at any adaptive step.

The port element for p -adaptive, S -parameter calculation was developed in order to allow some flexibility in the modeling of the microwave devices. The port element uses a port boundary condition to allow placement of the port at an arbitrary distance from a microwave junction. As the p -adaption proceeds, the number of waveguide modes taken into account on the port increases and so does the accuracy of the port boundary condition.

Error indication and estimation are two key concepts in FE p -adaption. An error indicator was developed to target elements in the mesh that had the largest impact on a global cost function. Accurate error estimation, while important in itself, is vital to combining FE adaption with optimization. To suit the needs of the optimization, the error estimator developed assesses errors in the gradient of a cost function. Calculating the error in a cost function (indicator) and its gradient (estimator) is very computationally inexpensive.

The combination of optimization and adaption was automated by defining an accuracy link (from optimizer to error estimator) to control the error tolerance required for termination of the adaptive process – thus controlling the accuracy at any step of the optimization. A number of unforeseeable, important issues arose in the practical implementation of the system, and were addressed satisfactorily.

The results validated all aspects of the system for a variety of test cases. The optimization results were compared to an optimization of fixed accuracy that required the accuracy of the gradient of any CFE to be the same as that at termination. Varying the levels of accuracy throughout the optimization led to speed-up factors (for $\alpha = 0.1$) of computational costs by at least an order of magnitude in each case. For the T-junction, the $\alpha = 0.1$ case achieved a speed-up of 117. In addition, it was shown that even when low accuracy is used at the early stages of an optimization, the cost function does genuinely reduce by the same amount as the fixed accuracy optimization, but at a much smaller, cumulative computational cost.

7.1 Original Contributions

In the course of developing an accuracy-controlled optimization, the following original contributions to knowledge have been made:

- a) A method of controlling accuracy by linking optimization to finite element p -adaption.
- b) An error estimator for design sensitivities for finite element analysis of microwave devices.
- c) A new element for the treatment of ports in p -adaptive finite element analysis of microwave devices.

- d) An error indicator that targets S -parameters for p -adaptive finite element analysis of microwave devices.
- e) Efficient calculation of S -parameter derivatives for scalar, hierarchal finite elements.

7.2 Suggestions for Further Work

Given the large number of components needed to assemble the optimization-adaption system, there are several research directions that can be followed from this work.

Most of the theory from this work can be extended to suit 3-D problems. The FE formulations for p -adaptive schemes can be generalized to 3-D by implementing tetrahedral finite elements with hierarchal vector basis functions [120]. In terms of time taken to complete microwave designs, using accuracy-controlled optimization to reduce computational costs in problems requiring 3-D FE analyses would be even more beneficial than in 2-D.

There are two issues in mesh generation (in 2-D or 3-D) that require improvement. Error estimation in the gradient would benefit from a mesh that is well discretized in terms of internal nodes. A mesh generator of this type would generate internal nodes in the vicinity of each node located on a parameterized boundary or interface. In addition, an interesting research question deals with modeling infinitely thin objects (like the iris of the 2-cavity filter). Suppose the thickness of such a post is a parameter of an optimization (e.g. the width of the inductive post in the T-junction), and the optimal geometry is an infinitely thin width. At the optimum, the post can be modeled as infinitely thin in a mesh. However, in the earlier stages of the optimization, the post might have a considerable width that must be modeled as such. An interesting question is when the width is being reduced throughout the optimization, at what point should the mesh generator begin to model the post as infinitely thin. Automating such a process would add much more flexibility to the mesh generation process.

All the theory proposed for the FE formulation is easily applicable to h -adaptive finite elements. Furthermore, a combination of h and p refinement schemes would probably be the most logical choice in creating a more powerful adaptive process. Using both h and p type refinement for the adaptive process would resolve the problem of not reaching a desired accuracy (because 10^{th} order elements for a particular mesh were just not accurate enough).

An improved control of accuracy for the optimization is to create two different α 's: one to control the final accuracy of the answer (α^{final}) and the other to control the path to get there (α^{path}). Furthermore, a study could be performed to find (if there is one) the best choice for α^{path} .

It is feasible to couple the entire system to a stochastic optimization method – thus increasing the chance of finding a global optimum. While there have been attempts to optimize in noisy search spaces [121], an interesting research path would be to visualize errors in CFEs as “numerical noise” and use this concept to develop an accuracy control in a combination of both gradient-based and stochastic optimization.

On a more general note, the concept of controlling accuracy throughout an optimization is not necessarily limited to one type of optimization method or CAD tool. The system developed in this thesis proves that great time-savings and reduction in computational effort can be achieved by controlling accuracy of CFEs throughout an optimization. For a problem to be a candidate for accuracy controlled optimization, it must have two main features: an expensive CFE and a way of trading cost for accuracy. Most problems whose cost functions are computed numerically (e.g. finite elements, finite difference, etc.) possess these two features. This leaves open the possibility for such a system to be applied to a range of fields outside microwaves and electromagnetics – such as fluid dynamics, structural analysis, etc. Furthermore, it would be fascinating to see if this idea can be applied to areas where the cost function is determined by sampling. For example, the result of a poll of N voters is defined to be a cost function. Controlling accuracy of such a cost function could correspond to the number of voters, N . The greater the size of the poll, the more accurate an assessment is achieved by the cost function.

References

- [1] N. Marcuvitz, *Waveguide Handbook*, McGraw-Hill Book Co., New York City, NY, 1951.
- [2] D. K. Cheng, *Field and Wave Electromagnetics*, Addison-Wesley Publishing Company, Reading, MA, 1983.
- [3] C. T. A. Johnk, *Engineering Electromagnetic Fields and Waves*, Addison-Wesley Publishing Company, Reading, MA, 1988.
- [4] F. Alessandri, M. Mongiardo and R. Sorrentino, "New efficient full wave optimization of microwave circuits by the adjoint network method," *IEEE Microwave and Guided Wave Letters*, Vol. 3, No. 11, pp. 414–416, November, 1993.
- [5] T. Itoh, "An overview on numerical techniques for modeling miniaturized passive components," *Annales des Telecommunications*, September–October, 1986.
- [6] R. Sorrentino, editor, *Numerical Methods for Passive Microwave and Millimeter Wave Structures*, IEEE Press, Piscataway, NJ, 1989.
- [7] T. Itoh, G. Pelosi and P. P. Silvester, *Finite Element Software for Microwave Engineering*, 1996.
- [8] J. M. Jin and J. L. Volakis, "A hybrid finite element method for scattering and radiation by microstrip patch antennas and arrays residing in a cavity," *IEEE Transactions on Antennas and Propagation*, Vol. 39, No. 11, pp. 1598–1604, November, 1991.
- [9] G. Matthaei, L. Young and E.M.T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, McGraw-Hill Book Company, MA, 1985.
- [10] S. S. Rao, *Engineering Optimization, Theory and Practice 3rd edition*, John Wiley and Sons, 1996.
- [11] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading MA, 1989.
- [12] P. van Laarhoven and E. Aarts, *Simulated Annealing: Theory and Applications*, D. Reigdel, Boston, MA, 1987.

- [13] David Isaak, Infolytica Corporation.
- [14] FullWave is a registered trademark of Infolytica Corp., www.infolytica.com, Montreal, QC, Canada.
- [15] Emag is a registered trademark of ANSYS, Inc., www.ansys.com, Canonsburg, PA 15317, USA.
- [16] HFSS is a registered trademark of Ansoft Corp., www.ansoft.com, Pittsburgh, PA 15219-1119, USA.
- [17] Z. J. Cendes, "EM simulators: Cae tools," *IEEE Spectrum*, Vol. 27, No. 11, pp. 73–77, November, 1990.
- [18] O. C. Zienkiewicz, *The Finite Element Method*, McGraw-Hill, London, 1977.
- [19] P. P. Silvester, "Finite element solution of homogeneous waveguide problems," *Alta Frequenza*, Vol. 38, pp. 313–317, 1969.
- [20] P. P. Silvester, "A general high-order finite element waveguide analysis program," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 17, pp. 204–210, April, 1969.
- [21] S. Ahmed, P. Daly, "Waveguide solutions by the finite element method," *Radio and Electronic Engineer*, Vol 38, pp. 217–223, 1969.
- [22] P. P. Silvester and R. L. Ferrari, *Finite Elements for Electrical Engineers*, Third Edition, Cambridge University Press, Cambridge, 1996.
- [23] J. M. Jin, *The Finite Element Method in Electromagnetics*, John Wiley & Sons, Inc., New York, 1993.
- [24] M. Calamia, G. Pelosi and P. P. Silvester, *Second International Workshop on Finite Element Methods for Electromagnetic Wave Problems: Siena, Italy, May 1994*, James and James, London, UK, 1994.
- [25] J. P. Webb, "Application of the finite element method to electromagnetic and electrical topics," *Reports on Progress in Physics*, Vol. 58, pp. 1673–1712, December, 1995
- [26] D. M. Pozar, *Microwave Engineering*, Addison-Wesley Publishing Company, New York, 1990.
- [27] R. E. Collin, *Foundations for Microwave Engineering*, McGraw-Hill, New York, 1966.

- [28] H. Akel and J. P. Webb, "Design sensitivities for scattering-matrix calculation with tetrahedral edge elements," *Conference on the Computation of Electromagnetic Fields (COMPUMAG)*, Sapporo, Japan, October 25-28, 1999; to appear in *IEEE Transactions on Magnetics*, July, 2000.
- [29] M. M. Gavrilovic and J. P. Webb "A port element for p -adaptive S-parameter calculation," *IEEE Transactions on Magnetics*, Vol. 35, No. 3, pp. 1530–1533, May, 1999.
- [30] J. P. Webb, "Finite element methods for junctions of microwave and optical waveguides," *IEEE Transactions on Magnetics*, Vol. 26, No. 5, pp. 1754–1758, September, 1990.
- [31] F. Alessandri, M. Mongiardo and R. Sorrentino, "Computer-aided design of beam forming networks for modern satellite antennas," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 40, no. 6, pp. 1117–1127, June, 1992.
- [32] F. Alessandri, M. Dionigi and R. Sorrentino, "A fullwave CAD tool for waveguide components using a high speed direct optimizer," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 43, no. 9, pp. 2046–2052, September, 1995.
- [33] J. P. Webb and S. Parihar, "Finite element analysis of H -plane rectangular waveguide problems," *Proc. Inst. Elec. Eng.*, Vol. 133, Pt. H, pp.103–109, January 1986.
- [34] Hong-bae Lee and T. Itoh, "A systematic optimum design of waveguide-to-microstrip transition," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 5, pp. 803–809, May, 1997.
- [35] Il-han Park, In-gu Kwak, Hyang-beom Lee, Ki-sik Lee and Song-yop Hahn, "Optimal design of transient eddy current systems driven by voltage source," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 33, No. 2, pp. 1624–1629, March, 1997.
- [36] J. F. Lee and Z. J. Cendes, "An adaptive spectral response modeling procedure for multiport microwave circuits," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 35, No. 12, pp. 1240–1247, December, 1987.

- [37] S. Kumashiro, R. A. Rohrer, A. J and Strojwas, "Asymptotic waveform evaluation for transient analysis of 3-D interconnect structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 12, No. 7, pp. 988–996, July, 1993.
- [38] S. V. Polstyanko, V. Sergey, R. Dyczij-Edlinger and J. F. Lee, "Fast frequency sweep technique for the efficient analysis of dielectric waveguides," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 7, pp. 1118–1126, July, 1997.
- [39] R. Sanaie, E. Chiprout, M. S. Nakhla and Q. Zhang, "A fast method for frequency and time domain simulation of high-speed VLSA interconnects," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 42, No. 12, pp. 2562–2571, December, 1994.
- [40] J. Gong and J. L. Volakis, "AWE implementation for electromagnetic FEM analysis," *Electronics Letters*, Vol. 21, November, 1996.
- [41] P. P. Silvester, "High-order polynomial triangular elements for potential problems," *International Journal for Numerical Methods in Engineering*, Vol. 17, pp. 849–861, 1969.
- [42] O. C. Zienkiewicz, J. P. De S. R. Gago and D. W. Kelly, "The hierarchical concept in finite element analysis," *Computers and Structures*, Vol. 16, No.1-4, pp. 53–65, 1983.
- [43] M. P. Rossow and I. N. Katz, "Hierarchal finite elements and precomputed arrays," *International Journal for Numerical Methods in Engineering*, Vol. 12, pp. 977–999, 1978.
- [44] O. C. Zienkiewicz, B. M. Irons, F. C. Scott and J. S. Campbell, "Three dimensional stress analysis," *Proceedings of the International Union of Theoretical and Applied Mechanics*, Liege, Belgium, 1971.
- [45] D. Rodger, "Experience with hierarchical finite elements in 2D electromagnetics," *IEEE Transactions on Magnetics*, Vol. 23, No. 5, pp. 3560–3562, September, 1987.
- [46] A. Darcherif, A. Raizer, G. Meunier, J. F. Imhoff and J. C. Sabonadiere, "New techniques in FEM field calculation applied to power cable characteristics

- computation," *IEEE Transactions on Magnetics*, Vol. 26, No. 5, pp. 2388–2390, September, 1990.
- [47] J. P. Webb and S. McFee, "The use of hierarchal triangles in finite-element analysis of microwave and optical devices," *IEEE Transactions on Magnetics*, Vol. 27, No. 5, pp. 4040–4043, September, 1991.
 - [48] J. P. Webb and R. Abouchacra, "Hierarchal triangular elements using orthogonal polynomials," *International Journal for Numerical Methods in Engineering*, Vol. 38, pp. 245–257, 1995.
 - [49] P. P. Silvester, "Construction of triangular finite element universal matrices," *International Journal for Numerical Methods in Engineering*, Vol. 12, pp. 237–244, 1978.
 - [50] N. S. Bardell, "The application of symbolic computing to the hierarchal finite element method," *International Journal for Numerical Methods in Engineering*, Vol. 28, pp. 1181–1204, 1989.
 - [51] P. Garcia and J. P. Webb, "Optimization of planar devices by the finite element method," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 38, No. 1, pp. 48–53, January, 1990.
 - [52] E. D. Rainville, *Special Functions*, Macmillan, New York, 1960, Chapter 16.
 - [53] M. Koshiba and M. Suzuki, "Finite-element analysis of *H*-plane waveguide junction with arbitrarily shaped ferrite post," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 34, pp. 103–109, January, 1986.
 - [54] Z. J. Cendes and J. F. Lee, "The transfinite element method for modeling MMIC devices," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 36, No. 12, pp. 1639–1649, December, 1988.
 - [55] K. L. Wu, G. Y. Delisle, D. G. Fang and M. Lecours, "Waveguide discontinuity analysis with a coupled finite-boundary element method," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 37, No. 6, pp. 993–998, June, 1989.
 - [56] J. F. Lee, "Analysis of passive microwave devices by using three-dimensional tangential vector finite elements," *International Journal of Numerical Modeling: Electronic Networks, Devices and Fields*, Vol. 3, No. 4, pp. 235–246, December, 1990.

- [57] S. J. Polak, C. den Heijer and Th. Beelen, "The future of mesh adaptive solving," *IEEE Transactions on Magnetics*, Vol. 18, No. 2, March, 1982.
- [58] S. McFee and J. P. Webb, "Adaptive finite element analysis of microwave and optical devices using hierarchal triangles," *IEEE Transactions on Magnetics*, Vol. 28, No. 2, pp. 1708–1711, March, 1992.
- [59] D. K. Sun, Z. J. Cendes and J. F. Lee, "Adaptive mesh refinement, h -version, for solving multiport microwave devices in three dimensions," *Conference on the Computation of Electromagnetic Fields (COMPUMAG)*, Sapporo, Japan, October 25-28, 1999.
- [60] F. Ling, C. F. Wang and J. M. Jin, "Application of adaptive integral method to scattering and radiation analysis of arbitrarily shaped planar structures," *IEEE Antennas and Propagation Society, AP-S International Symposium (Digest)*, Vol. 3, pp. 1778–1781, June, 1998.
- [61] J. Wang and J. P. Webb, "Hierarchal vector boundary elements and p -adaption for 3-D electromagnetic scattering," *IEEE Transactions on Antennas and Propagation*, Vol. 45, No. 12, pp. 1869–1879, December, 1997.
- [62] C. W. Steele, *Numerical Computation of Electric and Magnetic Fields*, Van Nostrand, New York, 1987, Chapter 8.
- [63] O. C. Zienkiewicz and R. L. Taylor, *The Finite-Element Method Vol. 1*, Fourth Edition, McGraw-Hill, London, 1988, Chapter 14.
- [64] O. C. Zienkiewicz, D. W. Kelly, J. Gago and I. Babuska, *The Mathematics of Finite Elements and Applications*, Ed., J.R. Whiteman, Academic Press, 1982.
- [65] M. M. Gavrilovic and J. P. Webb "An error indicator for the calculation of global quantities by the p -adaptive finite element method," *IEEE Transactions on Magnetics*, Vol. 33, No. 5, pp. 4128–4130, September, 1997.
- [66] I. Babuska, B. A. Szabo and I. N. Katz, "The p -version of the finite-element method," *SIAM Journal of Numerical Analysis*, No. 18, pp. 515–545, 1981.
- [67] W. Daigang and J. Kexun, " P -version. adaptive computation of FEM," *IEEE Transactions on Magnetics*, Vol. 30, No. 5, pp. 3515–3518, September, 1994.
- [68] I. Babuska and M. Sun, "The p and hp versions of the finite element method, basic principles and properties," *SIAM Review*, Vol. 36, No. 4, pp. 578–632, 1994.

- [69] A. Patra and J. T. Oden, "Computational techniques for adaptive *hp* finite element methods," *Finite Elements in Analysis and Design*, Vol. 25, pp. 27–39, 1997.
- [70] J. T. Oden, T. Strouboulis and P. Devloo, "Adaptive finite element methods for the analysis of inviscid compressible flow. Part 1: Fast refinement / unrefinement and moving mesh methods for unstructured meshes," *Computer Methods in Applied Mechanics and Engineering*, Vol. 59, pp. 327–362, 1986.
- [71] D. W. Kelly, J. P. De S. R. Gago and O. C. Zienkiewicz, "A *posteriori* error analysis and adaptive processes in the finite element method: part I: error analysis," *International Journal for Numerical Methods in Engineering*, Vol. 19, pp. 1593–1619, 1983.
- [72] A. Pinchuk and P. Silvester, "Error estimation for automatic adaptive finite element mesh generation," *IEEE Transactions on Magnetics*, Vol. 21, No. 6, pp. 2551–2554, November, 1985.
- [73] O. C. Zienkiewicz and J. Z. Zhu, "A simple error estimator and adaptive procedure for practical engineering analysis," *International Journal for Numerical Methods in Engineering*, Vol. 24, pp. 337–357, 1987.
- [74] S. Hahn, C. Calmels, G. Meunier and J. L. Coulomb, "Posteriori error estimate for adaptive finite element mesh generation," *IEEE Transactions on Magnetics*, Vol. 24, No. 1, pp. 315–317, January, 1988.
- [75] G. Drago, P. Molfino, M. Nervi and M. Repetto, "A 'local field error problem' approach for estimation in finite element analysis," *IEEE Transactions on Magnetics*, Vol. 28, No. 2, pp. 1743–1746, March, 1992.
- [76] M. G. Vanti, A. Raizer and J. P. A. Bastos, "Magnetostatic 2D comparison of local error estimates in FEM," *IEEE Transactions on Magnetics*, Vol. 29, No. 2, pp. 1902–1905, March, 1993.
- [77] J-F. Remacle, P. Dular, A. Genon and W. Legros, "Posteriori error estimation and adaptive meshing using error in constitutive relation," *IEEE Transactions on Magnetics*, Vol. 32, No. 3, pp. 1369–1372, May, 1996.
- [78] M. M. Gavrilovic and J. P. Webb "Targeted error indicators for use in finite-element *p*-adaption," *IEEE Transactions on Magnetics*, Vol. 34, No. 5, pp. 3280–3283, September, 1998.

- [79] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 17, pp. 533–552, 1969.
- [80] G. C. Temes and D. A. Calahan, "Computer-aided network optimization and the state-of-the-art," *Proceedings IEEE*, Vol. 55, pp. 1832–1863, 1967.
- [81] J. W. Bandler, "Circuit optimization: the state of the art," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 36, No. 2, pp. 424–443, February, 1988.
- [82] N. E. Sifi, G. Angenieux and Ph. Ferrari, "A combined algorithm for optimizing microwave component models in time domain," *IEEE Transactions on Magnetics*, Vol. 31, No. 3, pp. 1980–1983, May, 1995.
- [83] W. D. Becker, P. H. Harms and R. Mittra, "Time-domain electromagnetic analysis of interconnects in computer chip package," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 40, No. 12, pp. 2155–2163, December, 1993.
- [84] J. W. Bandler, R. M. Biernacki, S. H. Chen, L. W. Hendrick and D. Omeragic, "Electromagnetic optimization of 3-D structures," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 5, pp. 770–779, May, 1997.
- [85] F. Arndt, R. Beyer, J. M. Reiter, T. Sieverding and T. Wolf, "Automated design of waveguide components using hybrid mode-matching/numerical EM building blocks in optimization-oriented CAD frameworks—state-of-the-art and recent advances," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 5, pp. 747–759, May, 1997.
- [86] N. Jain and P. Onno, "Methods of using commercial electromagnetic simulators for microwave and millimeter-wave circuit design and optimization," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 5, pp. 724–745, May, 1997.
- [87] Ansoft Corporation, "Parametrics and optimization using Ansoft HFSS," *Microwave Journal*, Vol. 42, No. 11, Horizon House Publications, Inc., pp. 156–166, November, 1999.
- [88] J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobelny and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 42, pp. 2536–2544, December, 1994.

- [89] J. W. Bandler, R. M. Biernacki, S. H. Chen, R. H. Hemmers and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 43, No. 12, pp. 2874–2882, December, 1995.
- [90] D. A. Lowther, "Knowledge-based and numerical optimization techniques for the design of electromagnetic devices," *International Journal for Numerical Modeling*, Vol. 9, pp. 35–44, January–April, 1996.
- [91] A. H. Zaabab, Qi-Jun Zhang and M. Nakhla, "A neural network modeling approach to circuit optimization and statistical design," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 43, No. 6, pp. 1349–1358, June, 1995.
- [92] P. Alotto et. al, "Stochastic algorithms in electromagnetic optimization," *IEEE Transactions on Magnetics*, Vol. 34, No. 5, pp. 3674–3684, September, 1998.
- [93] F. G. Uhler, O. A. Mohammed and C. S. Koh, "Utilizing genetic algorithms for the optimal design of electromagnetic devices," *IEEE Transactions on Magnetics*, Vol. 30, pp. 4296–4298, 1994.
- [94] C. A. Magele, K. Preis, W. Renhart, R. Dyczij-Edlinger and R. Richter, "Higher order evolution strategies for global optimization of electromagnetic devices," *IEEE Transactions on Magnetics*, Vol. 29, No. 2, pp. 1775–1778, March, 1993.
- [95] P. Alotto and C. Magele, "Optimization in electromagnetics," *International Compumag society newsletter*, Vol. 6, No. 3, pp. 7–9, November, 1999.
- [96] D. N. Dyck, D. A. Lowther and E. M. Freeman, "A method of computing sensitivity of electromagnetic quantities to changes in materials and sources," *IEEE Transactions on Magnetics*, Vol. 30, pp. 341–344, September, 1994.
- [97] D. N. Dyck, "Automating the topological design of magnetic devices," Ph.D. thesis, McGill University, Canada, September, 1995.
- [98] S. W. Director and R. A. Rohrer, "Generalized adjoint network and network sensitivities," *IEEE Transactions on Circuit Theory*, Vol. 16, pp. 318–323, 1969.
- [99] J. L. Coulomb et. al, "Sensitivity analysis using high order derivatives," *Applied Computational Electromagnetics Society Journal*, Vol. 12, No. 2, pp. 20–25, March, 1997.

- [100] In-gu Kwak, Young-woo Ahn, Song-yop Hahn and Il-han Park, "Shape optimization of electromagnetic devices using high order derivatives," *IEEE Transactions on Magnetics*, Vol. 35, No. 3, pp. 1726–1729, May, 1999.
- [101] J. W. Bandler, R. M. Biernacki and S. H. Chen, "Parameterization of arbitrary geometrical structures for automated electromagnetic optimization," *IEEE MTT-S International Microwave Symposium Digest*, San Francisco, CA, June, 1996, pp. 1059–1062.
- [102] A. Grace, *Optimization Toolbox for use with MATLAB*, The Math Works Inc., Mass., pp. 2-22–2-31, 1992.
- [103] P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, Academic Press, London, 1981.
- [104] M. C. Biggs, "Constrained minimization using recursive quadratic programming," *Towards Global Optimization* (L. C. W. Dixon and G. P. Szergo, editors), North-Holland, pp. 341–349, 1975.
- [105] S. P. Han, "A globally convergent method for nonlinear programming," *Journal of Optimization Theory and Applications*, Vol. 22, p. 297, 1977.
- [106] M. J. D. Powell, "A fast algorithm for nonlinearly constrained optimization calculations," *Numerical Analysis* (G. A. Watson, editor), Lecture Notes in Mathematics, Springer Verlag, Vol. 630, 1978.
- [107] M. J. D. Powell, "The convergence of variable metric methods for nonlinearly constrained optimization calculations," *Nonlinear Programming 3* (O. L. Mangasarian, R. R. Meyer and S. M. Robinson, editors), Academic Press, New York, 1978.
- [108] P. E. Gill, W. Murray and M. H. Wright, *Numerical Linear Algebra and Optimization*, Vol. 1, Addison-Wesley, 1991.
- [109] P. E. Gill, W. Murray, M. A. Saunders and M. H. Wright, "Procedures for optimization problems with a mixture of bounds and general linear constraints," *ACM Transactions Math. Software*, Vol. 10, pp. 282–298, 1984.
- [110] G. Dantzig, *Linear Programming and Extensions*, Princeton University Press, Princeton, 1963.

- [111] Fortran PowerStation Version 4.0 is a registered trademark of Microsoft Corp., USA.
- [112] MATLAB Version 4.2c is a registered trademark of Mathworks Inc., Natick, MA 01760, USA.
- [113] Windows NT Version 4.0 is a registered trademark of Microsoft Corp., USA.
- [114] B. M. Irons, "A frontal solution program for finite element analysis," *International Journal for Numerical Methods in Engineering*, Vol. 2, pp. 5–32, 1970.
- [115] Hong-bae Lee, Hyun-Kyo Jung, Song-yop Hahn, C. Cheon and Ki-Sik Lee, "Shape optimization of H -plane waveguide tee junction using edge finite element method," *IEEE Transactions on Magnetics*, Vol. 31, No. 3, pp. 1928–1931, May, 1995.
- [116] J. Baden Fuller, *Microwaves, an introduction to microwave theory and techniques*, Pergamon Press, second edition, 1979, Section 10.4.
- [117] J. T. Alos and M. Guglielmi, "Simple and effective EM-based optimization procedure for microwave filters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 6, pp. 856–858, June, 1997.
- [118] K. Shamsaifar, "Designing iris-coupled waveguide filters using the mode-matching technique," *Microwave Journal*, pp. 156–154, January, 1992.
- [119] J-Fuh Liang, Hsin-Chin Chang and K. A. Zaki, "Design and tolerance analysis of thick iris waveguide bandpass filters," *IEEE Transactions on Magnetics*, Vol. 29, No. 2, pp. 1605–1608, March, 1993.
- [120] J. P. Webb, "Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements," *IEEE Transactions on Antennas and Propagation*, Vol. 47, No. 8, pp. 1244–1253, August, 1999.
- [121] Z. Badics, J. Pavo, H. Komatsu, S. Kojima, Y. Matsumoto and K. Aoki, "Fast flaw reconstruction from 3D eddy current data," *IEEE Transactions on Magnetics*, Vol. 34, No. 5, pp. 2823–2828, September, 1998.

Appendix A

Derivative of Matrix $[K]$ with respect to r_k^l

For a vertex k , vector r^k is permitted to move in direction a_l while all other nodes attached to k via adjacent elements have fixed positions. Figure A.1 shows the vertex within region Ω_k , discretized into triangles Ω_k^e .

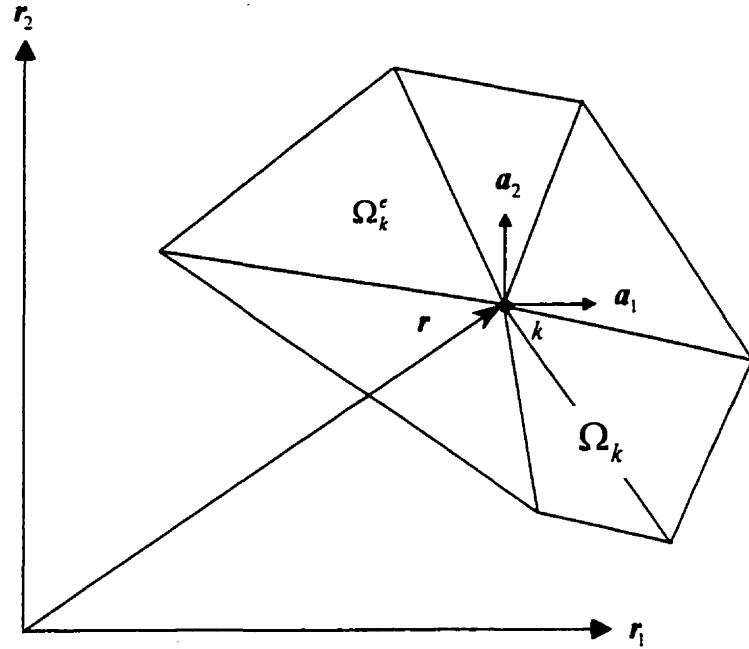


Figure A.1: Vertex k among surrounding elements in region Ω_k .

Unit vectors a_1 and a_2 are parallel to axes r_1 and r_2 (x and z in the H -plane defined in Chapter 2). As node k moves, it is assumed that all points in Ω_k move also, in a continuous way. The general position of a point r in Ω_k becomes

$$r + \gamma v \tag{A.1}$$

where γ is a parameter controlling the movement and \mathbf{v} is a continuous displacement field, having value \mathbf{a}_l at node k and 0 at all other nodes:

$$\mathbf{v}^e = \mathbf{a}_l \zeta_k \quad \text{for } l = 1, 2; \quad k = 1, 2, 3 \quad (\text{A.2})$$

where \mathbf{v}^e is the local denotation for \mathbf{v} . Since $\zeta_1, \zeta_2, \zeta_3$ are simplex co-ordinates defined locally for each element, \mathbf{v}^e conveniently becomes equal to \mathbf{a}_l at $\mathbf{r} = \mathbf{r}^k$ and zero outside region Ω_k .

Consider element e connected to the k^{th} vertex. The derivative of local matrix $[K]$ becomes

$$\begin{aligned} \frac{\partial K_{ij}}{\partial r_i^k} &= \frac{\partial B(N_i, N_j)}{\partial r_i^k} = \frac{\partial}{\partial r_i^k} \frac{1}{jk_o \eta_o} \int_{\Omega_k^e} \left(\frac{1}{\mu_r} \nabla N_i \cdot \nabla N_j - k_o^2 \epsilon_r N_i N_j \right) d\Omega \\ &\quad + \frac{\partial}{\partial r_i^k} \sum_{q=1}^N \sum_{l=0}^{\infty} \gamma_l^{(q)}(N_i) \gamma_l^{(q)}(N_j) \end{aligned} \quad (\text{A.3})$$

Provided that the ports do *not* vary with parameter g (or rather do not contain any vertices affected by g), the last term becomes zero and (A.3) can be shown [51] to be

$$\begin{aligned} \frac{\partial K_{ij}}{\partial r_i^k} &= \frac{1}{jk_o \eta_o} \int_{\Omega_k^e} \nabla \cdot \mathbf{v}^e \left(\frac{1}{\mu_r} \nabla N_i \cdot \nabla N_j - k_o^2 \epsilon_r N_i N_j \right) d\Omega \\ &\quad - \frac{1}{jk_o \eta_o \mu_r} \int_{\Omega_k^e} \sum_{m=1}^2 \sum_{n=1}^2 \frac{\partial v_m^e}{\partial r_n} \left[\frac{\partial N_i}{\partial r_m} \frac{\partial N_j}{\partial r_n} + \frac{\partial N_i}{\partial r_n} \frac{\partial N_j}{\partial r_m} \right] d\Omega \end{aligned} \quad (\text{A.4})$$

where the divergence of \mathbf{v}^e is given by

$$\nabla \cdot \mathbf{v}^e = \nabla \zeta_k \cdot \mathbf{a}_l \quad (\text{A.5})$$

and

$$\frac{\partial v_m^e}{\partial r_n} = \nabla v_m^e \cdot \mathbf{a}_n = \delta_{ml} \nabla \zeta_k \cdot \mathbf{a}_n \quad (\text{A.6})$$

Both $\frac{\partial v_m^e}{\partial r_n}$ and $\nabla \cdot \mathbf{v}^e$ are constant over element e because \mathbf{v}^e is defined to be linear.

Substituting (A.5) and (A.6) into (A.4) yields

$$\begin{aligned} \frac{\partial K_y}{\partial r_l^k} = & \frac{\nabla \zeta_k \cdot \mathbf{a}_l}{jk_o n_o} \int_{\Omega_k^e} \left(\frac{1}{\mu_r} \nabla N_i \cdot \nabla N_j - k_o^2 \epsilon_r N_i N_j \right) d\Omega \\ & - \frac{1}{jk_o n_o \mu_r} \int_{\Omega_k^e} \sum_{m=1}^2 \sum_{n=1}^2 \nabla \zeta_k \cdot \mathbf{a}_n \delta_{ml} \left[\frac{\partial N_i}{\partial r_m} \frac{\partial N_j}{\partial r_n} + \frac{\partial N_i}{\partial r_n} \frac{\partial N_j}{\partial r_m} \right] d\Omega \end{aligned} \quad (\text{A.7})$$

(A.7) can be rearranged to contain 3 explicit terms:

$$\begin{aligned} \frac{\partial K_y}{\partial r_l^k} = & \frac{\nabla \zeta_k \cdot \mathbf{a}_l}{jk_o n_o \mu_r} \int_{\Omega_k^e} (\nabla N_i \cdot \nabla N_j) d\Omega \\ & - \frac{1}{jk_o n_o \mu_r} \sum_{m=1}^2 \sum_{n=1}^2 \nabla \zeta_k \cdot \mathbf{a}_n \delta_{ml} \int_{\Omega_k^e} \left[\frac{\partial N_i}{\partial r_m} \frac{\partial N_j}{\partial r_n} + \frac{\partial N_i}{\partial r_n} \frac{\partial N_j}{\partial r_m} \right] d\Omega \\ & - \frac{\nabla \zeta_k \cdot \mathbf{a}_l k_o^2 \epsilon_r}{jk_o n_o} \int_{\Omega_k^e} (N_i N_j) d\Omega \end{aligned} \quad (\text{A.8})$$

If the basis functions N_i and N_j are expressed in terms of simplex co-ordinates, using (3.6) following relations apply:

$$\nabla N_i \cdot \nabla N_j = \sum_{p=1}^2 \sum_{q=1}^2 \nabla \zeta_q \cdot \nabla \zeta_p \frac{\partial N_i}{\partial \zeta_p} \frac{\partial N_j}{\partial \zeta_q} \quad (\text{A.9})$$

$$\frac{\partial N_i}{\partial r_m} \frac{\partial N_j}{\partial r_n} = \sum_{p=1}^2 \sum_{q=1}^2 (\nabla \zeta_p \cdot \mathbf{a}_m) (\nabla \zeta_q \cdot \mathbf{a}_n) \frac{\partial N_i}{\partial \zeta_p} \frac{\partial N_j}{\partial \zeta_q} \quad (\text{A.10})$$

Let

$$S_{ij}^{(1)} = \int_{\Delta} \frac{\partial N_i}{\partial \zeta_1} \frac{\partial N_j}{\partial \zeta_1} d\Delta \quad (\text{A.11})$$

$$S_{ij}^{(2)} = \int_{\Delta} \left(\frac{\partial N_i}{\partial \zeta_1} \frac{\partial N_j}{\partial \zeta_2} + \frac{\partial N_i}{\partial \zeta_2} \frac{\partial N_j}{\partial \zeta_1} \right) d\Delta \quad (\text{A.12})$$

$$S_{ij}^{(3)} = \int_{\Delta} \frac{\partial N_i}{\partial \zeta_2} \frac{\partial N_j}{\partial \zeta_2} d\Delta \quad (\text{A.13})$$

where Δ ranges from $\zeta_2 = 0 \dots 1 - \zeta_1$ and $\zeta_1 = 0 \dots 1$ and $d\Delta = d\zeta_1 d\zeta_2$. The integral of (A.9) can now be written as a function of these “universal” matrices.

$$\int_{\Omega_k^e} \nabla N_i \cdot \nabla N_j d\Omega = 2A (\nabla \zeta_1^2 S_{ij}^{(1)} + \nabla \zeta_1 \cdot \nabla \zeta_2 S_{ij}^{(2)} + \nabla \zeta_2^2 S_{ij}^{(3)}) \quad (\text{A.14})$$

where A is the area of the triangle and $2A$ is the Jacobian of the mapping from Ω to Δ .

In using (A.10), the integrand of the second term of (A.8) can be expanded so

$$\frac{\partial N_i}{\partial r_m} \frac{\partial N_j}{\partial r_n} + \frac{\partial N_i}{\partial r_n} \frac{\partial N_j}{\partial r_m} = \sum_{p=1}^2 \sum_{q=1}^2 [(\nabla \zeta_p \cdot \mathbf{a}_m) (\nabla \zeta_q \cdot \mathbf{a}_n) + (\nabla \zeta_p \cdot \mathbf{a}_n) (\nabla \zeta_q \cdot \mathbf{a}_m)] \frac{\partial N_i}{\partial \zeta_p} \frac{\partial N_j}{\partial \zeta_q} \quad (\text{A.15})$$

and its integral can similarly be re-written as a function of the $[S^{(i)}]$ matrices, so with the use of (A.15):

$$\begin{aligned} \int_{\Omega_k^e} \left[\frac{\partial N_i}{\partial r_m} \frac{\partial N_j}{\partial r_n} + \frac{\partial N_i}{\partial r_n} \frac{\partial N_j}{\partial r_m} \right] d\Omega &= 4A (\nabla \zeta_1 \cdot \mathbf{a}_m) (\nabla \zeta_1 \cdot \mathbf{a}_n) S_{ij}^{(1)} \\ &+ 2A [(\nabla \zeta_1 \cdot \mathbf{a}_m) (\nabla \zeta_2 \cdot \mathbf{a}_n) + (\nabla \zeta_1 \cdot \mathbf{a}_n) (\nabla \zeta_2 \cdot \mathbf{a}_m)] S_{ij}^{(2)} \\ &+ 4A (\nabla \zeta_2 \cdot \mathbf{a}_m) (\nabla \zeta_2 \cdot \mathbf{a}_n) S_{ij}^{(3)} \end{aligned} \quad (\text{A.16})$$

The integral of the last term of (A.8) can be recognized as the metric, of the form $2A[T]$, where

$$T_y = \int_{\Delta} N_i N_j d\Delta \quad (\text{A.17})$$

By combining (A.14), (A.16) and (A.17), (A.8) can be expressed in the convenient form of universal matrices:

$$\frac{\partial K_y}{\partial r_k^I} = \frac{2A}{jk_o n_o} \left(\frac{1}{\mu_r} \sum_{p=1}^3 d_{ikp} S_y^{(p)} - k_o^2 \epsilon_r \nabla \zeta_k \cdot \mathbf{a}_I T_y \right) \quad (\text{A.18})$$

where

$$\begin{aligned} d_{ik1} &= \nabla \zeta_1^2 \nabla \zeta_k \cdot \mathbf{a}_I - 2(\nabla \zeta_1 \cdot \mathbf{a}_I)(\nabla \zeta_1 \cdot \mathbf{a}_k) \\ d_{ik2} &= (\nabla \zeta_1 \cdot \nabla \zeta_2) \nabla \zeta_k \cdot \mathbf{a}_I - (\nabla \zeta_1 \cdot \mathbf{a}_I)(\nabla \zeta_2 \cdot \mathbf{a}_k) - (\nabla \zeta_2 \cdot \mathbf{a}_I)(\nabla \zeta_1 \cdot \mathbf{a}_k) \\ d_{ik3} &= \nabla \zeta_2^2 \nabla \zeta_k \cdot \mathbf{a}_I - 2(\nabla \zeta_2 \cdot \mathbf{a}_I)(\nabla \zeta_2 \cdot \mathbf{a}_k) \end{aligned} \quad (\text{A.19})$$

Appendix B

Local Optima of the Miter Bend Problem

Three local optima were found in the optimization of the miter bend with dielectric block: optima 1, 2 and 3.

The geometry of the miter bend for optimum 1 (same as that in section 6.3.4) is illustrated in Figure B.1, with and without the dielectric block. For the same mesh and accuracy level (all 10th order elements), removing the block increases the cost function from $C(\underline{g}^{\text{optimum 1}}) = 5.8413 \times 10^{-6}$ to $C(\underline{g}^{\text{optimum 1}}) = 1.4464 \times 10^{-1}$. Frequency responses of the return loss for optimum 1 with and without the dielectric block are given in Figure B.2 (for 200 discretely sampled points).

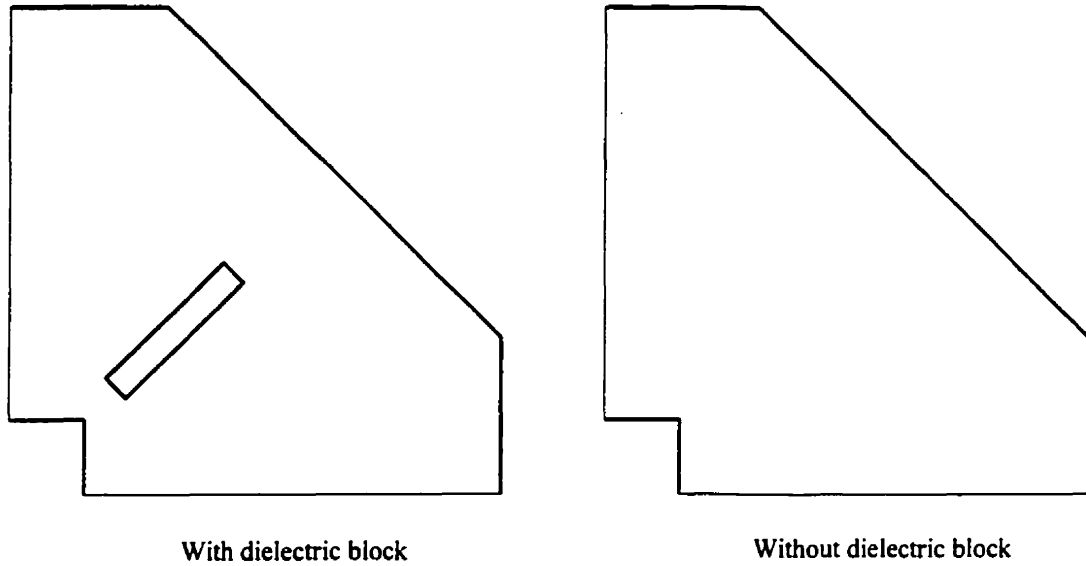


Figure B.1: Miter bend with and without dielectric block for optimum 1. The geometry of the device with the block is $\underline{g}^{\text{optimum 1}} = [2.2489 \ 1.0167 \ 0.1346 \ 1.1280]^T$ cm and without is $g_1^{\text{optimum 1}} = 2.2489$ cm .

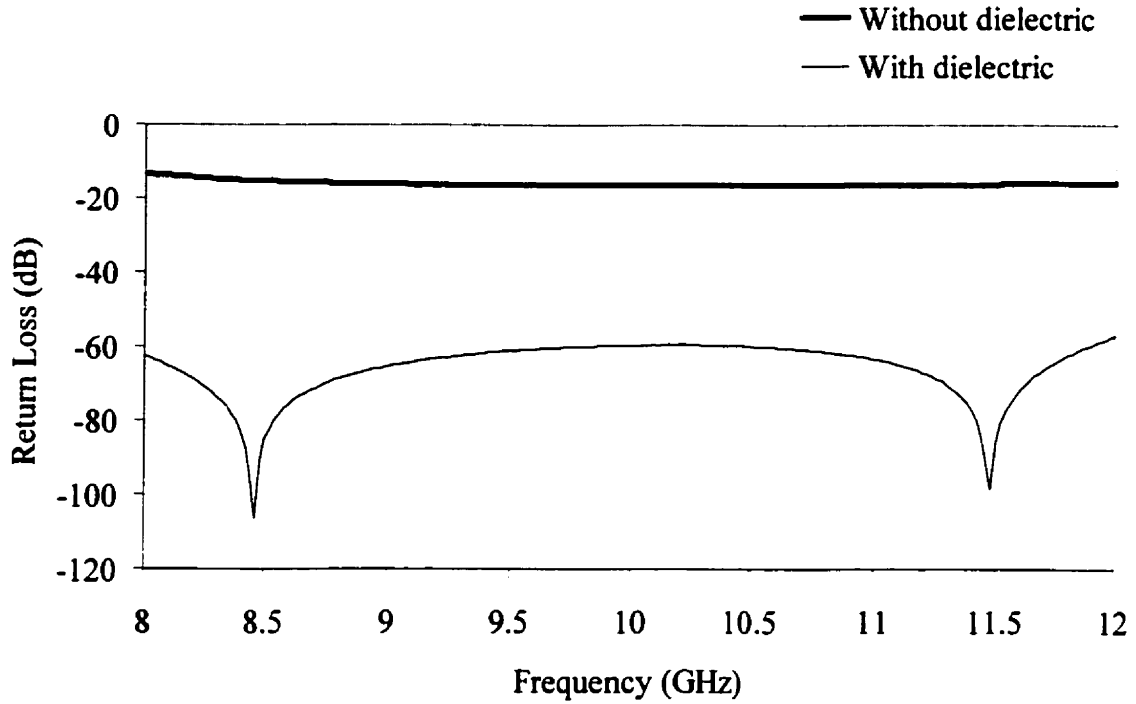


Figure B.2: Frequency responses of the return loss for the miter bend with and without dielectric block. $\underline{g}^{\text{optimum}^1} = [2.2489 \ 1.0167 \ 0.1346 \ 1.1280]^T \text{ cm}$.

Figures B.3 through B.6 illustrate optima 2, 3 and the corresponding return losses (computed using all 10th order elements) plotted for 200 points throughout the frequency range. Although the geometry is different, optimum 2 yields a similar frequency response (and thus, a similar cost function value) to optimum 1 – with 2 computed resonances (see Figure B.4). Optimum 3, a result of starting the optimization at $\underline{g}^{\circ} = [2\sqrt{2} \ \sqrt{2}/2 \ \sqrt{2}/2 \ \sqrt{2}/2]^T \text{ cm}$, is interesting because the dielectric block is reduced to its minimum size, as close to the corner as possible (Figure B.5). Although it is a genuine local optimum, the cost function is considerably higher than for the two other optima and its frequency response has only one computed resonance (Figure B.6).

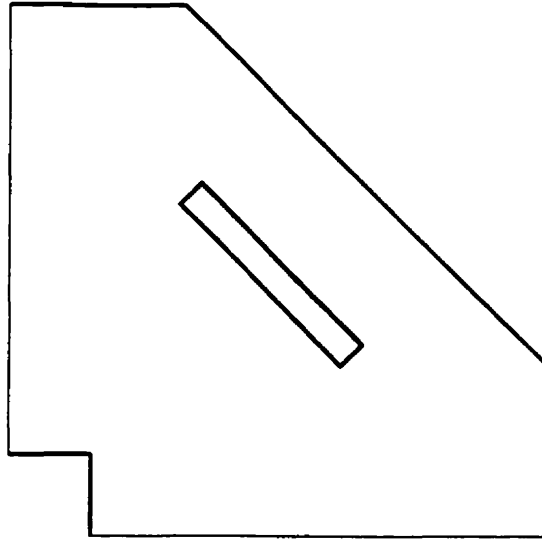


Figure B.3: Miter bend with dielectric block at optimum 2. The geometry of the device $\underline{g}^{\text{optimum } 2} = [2.2680 \ 0.1382 \ 0.7979 \ 0.6176]^T$ cm and the cost function at this optimum is $C(\underline{g}^{\text{optimum } 2}) = 7.3910 \times 10^{-6}$.

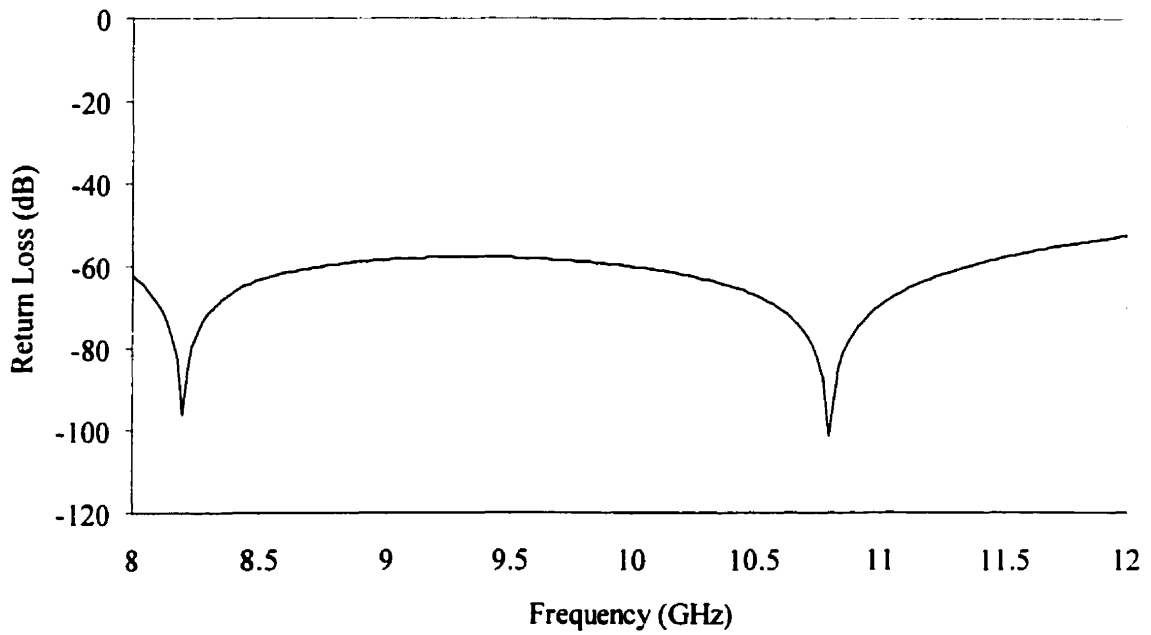


Figure B.4: Frequency response of the return loss for the miter bend with dielectric block at optimum 2, shown in Figure B.3.

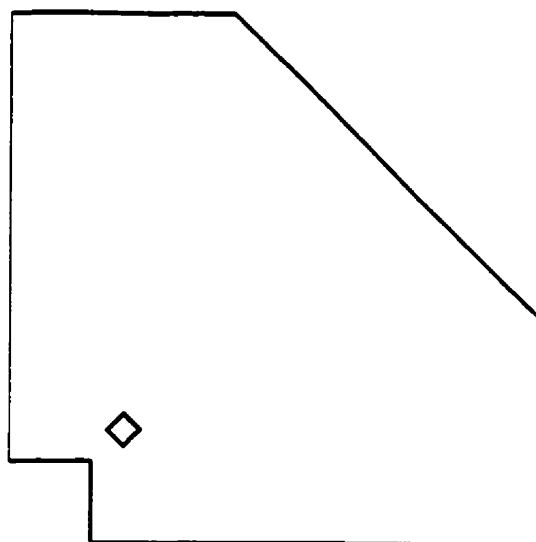


Figure B.5: Miter bend with dielectric block at optimum 3. The geometry of the device $\underline{g}^{\text{optimum } 3} = [1.9441 \ 0.1 \ 0.1 \ 0.2]^T$ cm and the cost function at this optimum is $C(\underline{g}^{\text{optimum } 3}) = 3.8874 \times 10^{-2}$.

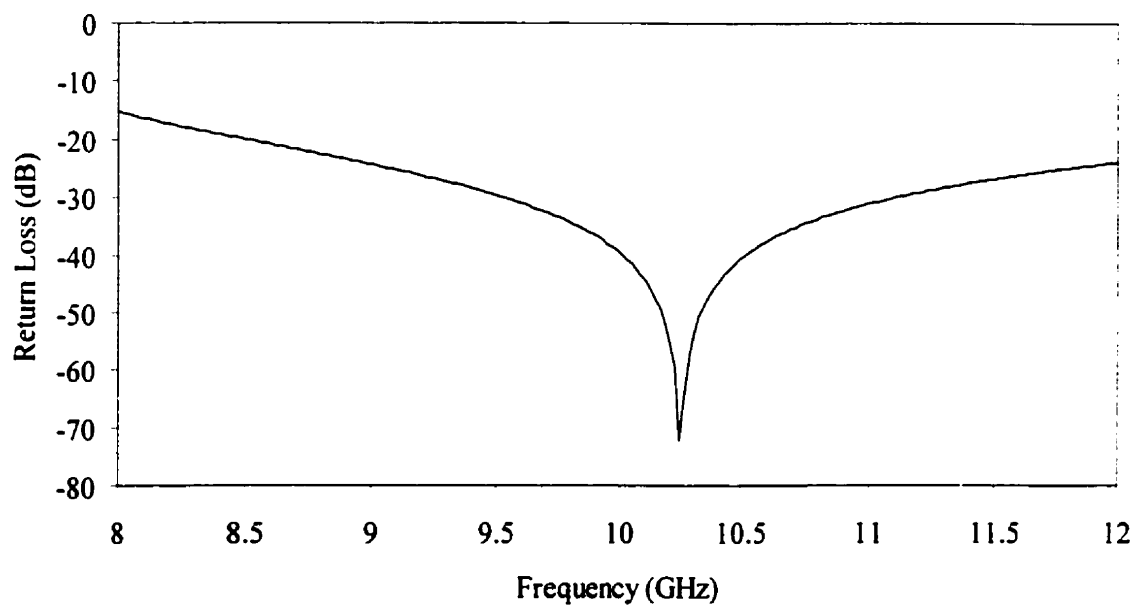


Figure B.6: Frequency response of the return loss for the miter bend with dielectric block at optimum 3, shown in Figure B.5.